Washington State Health Care Authority

Proton beam therapy – re-review

Final evidence report

April 15, 2019

Health Technology Assessment Program (HTA)
Washington State Health Care Authority
PO Box 42712
Olympia, WA 98504-2712
(360) 725-5126
www.hca.wa.gov/hta/
shtap@hca.wa.gov

Proton Beam Therapy – Re-review

Provided by:

Aggregate Analytics, Inc.

Prepared by:

Andrea C. Skelly, PhD, MPH
Erika D. Brodt, BS
Naomi Schwartz, MPH
Aaron J. R. Ferguson, BA
Shelby Kantner, BA

April 15, 2019

This technology assessment report is based on research conducted by a contracted technology assessment center, with updates as contracted by the Washington State Health Care Authority. This report is an independent assessment of the technology question(s) described based on accepted methodological principles. The findings and conclusions contained herein are those of the investigators and authors who are responsible for the content. These findings and conclusions may not necessarily represent the views of the HCA/Agency and thus, no statement in this report shall be construed as an official position or policy of the HCA/Agency.

The information in this assessment is intended to assist health care decision makers, clinicians, patients and policy makers in making sound evidence-based decisions that may improve the quality and cost-effectiveness of health care services. Information in this report is not a substitute for sound clinical judgment. Those making decisions regarding the provision of health care services should consider this report in a manner similar to any other medical reference, integrating the information with all other pertinent information to make decisions within the context of individual patient circumstances and resource availability.

Table of Contents

Tables

Figures

Abbreviations

AE: adverse event
CI: confidence interval
CGE: Cobalt Gray Equivalent (unit)
cGy: Centigray (unit)
CR: Complete Response
CSI: Craniospinal Irradiation
CSS: Cause-Specific Survival
CT: computed tomography
CTCAE: Common Terminology Criteria for Adverse Events
DFS: disease-free survival
DS-PBT: Double-Scattering Proton Beam Therapy
DSS: Disease-Specific Survival
EFS: event-free survival
FFDM: Freedom from Distant Metastases
FSRT: Fractionated stereotactic radiation therapy
F/U: follow-up
Gy: Gray (unit)
GyE: Gray-Equivalents (unit)
Gy (RBE): Gray Relative Biological Effectiveness (unit)
HART Hyper-fractionated acceleration radiotherapy
HR: hazard ratio
IMRT: Intensity Modulated Radiation Therapy
IMPT: Intensity Modulated Proton Therapy
MDASI: MD Anderson Symptom Inventory
MRI: Magnetic Resonance Imaging
NC: not calculable
NHL: Non-Hodgkin lymphoma
NR: not reported
NS: not statistically significant
OR: odds ratio
OS: overall survival
PBS: Pencil Beam Scanning
PBT: Proton Beam Therapy
PD: Progressive Disease
PFS: progression-free survival
PR: Partial Response
PS-PBT: Passive-Scatter Proton Beam Therapy
QoL: Quality of Life
RCT: randomized controlled trial
RD: risk difference
RFS: relapse- or recurrence-free survival
RN: Radiation Necrosis
RoB: risk of bias
RR: risk ratio
RT: radiation therapy

RTOG: Radiation Therapy Oncology Group
SBRT: Stereotactic Body Radiation Therapy
SE: standard error
SD: standard deviation
SD: Stable Disease
TACE: Trans arterial chemoembolization
TFS: Toxicity Free Survival
WHO: World Health Organization

Executive Summary

Introduction

Overall, it is estimated that 1.7 million new cases of cancer are diagnosed yearly and cancerous conditions are responsible for over half a million deaths per year.[4] Using incidence and survival data from the Surveillance, Epidemiology, and End Result (SEER) Program and population projections from the U.S Census Bureau, the National Cancer Institute (NCI) projects the total cost of cancer care in the United States in 2020 to be $174 billion.[1] Treatment options for cancerous and noncancerous conditions vary depending on the type, location and stage of the condition and can include radiation therapy, chemotherapy, targeted therapy (e.g. inhibitor drugs), immunotherapy (including monoclonal antibodies) and surgery, or combinations of these treatments. Radiation may be delivered systemically via radioactive drugs, however, the two most common forms of radiation therapy are external beam radiation therapy (EBRT) and brachytherapy (internal radiation therapy). Today, approximately 50% of all cancer patients benefit from radiation therapy in the management of their disease and it may be the sole therapy used.[24] The focus of this review will be to evaluate the safety and effectiveness of Proton Beam Therapy (PBT), a form of external beam radiation therapy compared with other forms of cancer treatment. The use of protons for radiotherapy has a history of over 60 years of clinical use. PBT use was initially directed towards conditions where sparing sensitive adjacent normal tissues was considered to be of utmost importance (such as cancerous or noncancerous malformations of the brain stem, eye, or spinal cord) or for many pediatric tumors because of the particular risk of pronounced acute and long-term toxicity in pediatric patients.[77] PBT may be most promising for tumors in moderate proximity to (>2 cm) organs at risk (OAR). In recent years the use of proton beam therapy (PBT) has expanded to include a variety of conditions including a number of cancer types, noncancerous brain tumors and cancerous conditions afflicting the central nervous system as well as eyes, lungs, liver, prostate, spine, and pelvis.

Radiation therapy (RT) involves high-energy radiation from gamma rays, electron beams, photon beams or proton beams that breaks the DNA of cancer cells, inhibiting their ability to proliferate. The radiation may also affect surrounding healthy tissues. Tumor types (and healthy tissues) vary with regard to their sensitivity to radiation. A goal of treatment planning is to damage cancer cells while minimizing damage to surrounding healthy cells including sensitive structures and organs at risk (OARs). Most often radiation is delivered using external beam radiation therapy (EBRT), a method of externally delivering radiation using a machine to aim high-energy beams directly at the tumor from outside the body. Classification of RT may be by the type of beam or particle used (i.e. electron, photon or proton) with photon RT being the most widely available and commonly used.[18] RT may be used for a variety of reasons including to cure a radiosensitive tumor, to shrink a tumor pre-operatively, to prevent recurrence or spread post-operatively (adjuvant treatment), to treat a recurrent tumor or as a palliative treatment. It may be combined with other treatments such as chemotherapy. Radiosensitive tumors for which RT may be curative include, but are not limited to, prostate cancers, head and neck cancers, and non-small cell lung cancer. RT in combination with other treatment regimens is commonly used to treat breast cancer, colon cancer, lung cancers, seminomas, and some cancers of the central nervous system, among others.

Side effects of radiation therapy occur when healthy tissues in the path of the radiation beam are damaged; the effects vary from person to person. A variety of factors impact the location, type, timing and severity of side effects including the type/method of delivery and dose of radiation, the area of the body that is exposed to radiation and a person's overall health. General short-term side effects of

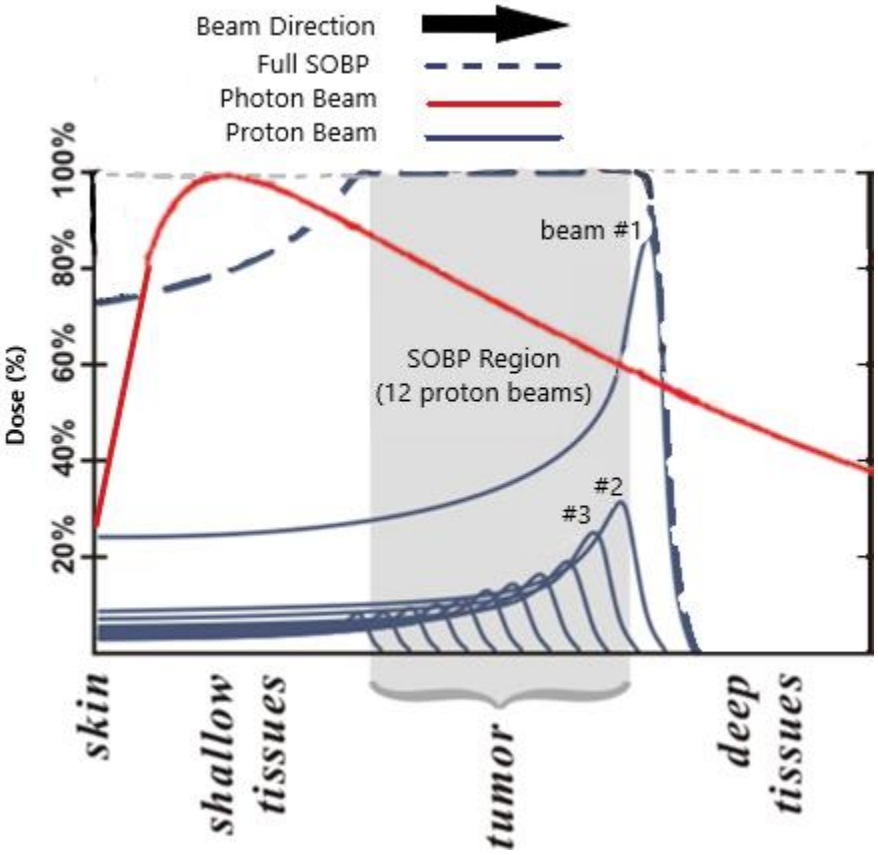
radiation therapy may include fatigue and skin irritation (radiation dermatitis) at the radiation site. These usually subside after treatment completion. Other side effects (short and longer term) depend on the site that was irradiated and the sensitivity of tissues surrounding the tumor and may range from mild to life-threatening. Long-term consequences to radiation therapy are generally rare. Radiation is a carcinogen and rarely, secondary cancers may occur in long-term cancer survivors who have had radiation therapy; this is of particular concern in patients receiving radiation at younger ages. The effects of radiation damage may be more nuanced in children, such as effects on neurocognitive development, especially when administered to children under 3 years of age.[83] Even lower-dose irradiation of normal tissue in pediatric patients can result in pronounced acute and long-term toxicity.[77] Thus, the opportunity to limit radiation exposure to normal and developing tissues is important and is part of radiation planning.

In its earliest applications, RT planning employed X-ray technology to take two-dimensional images (referred to as two-dimensional RT (2DRT) or Conventional RT (CRT) of the tumor location which were then used to determine how best to position the radiation beams in order to effectively treat the tumor. Major technological developments in computer and imaging technologies further improved upon the ability to deliver a consistent radiation dose to irregularly shaped tumors in difficult anatomic locations, while simultaneously sparing normal tissues from unnecessary radiation. Thus, 2DRT/CRT has largely been replaced by Three-dimensional Conformal Radiation Therapy (3DCRT), which uses three-dimensional imaging, such as Computed Tomography (CT) scans and Magnetic Resonance Imaging (MRI), to very accurately map the location and size of the tumor in three dimensions, as well as identify any critical OARs. Using these 3D images, beams are then matched very precisely to the shape of the tumor and delivered from all directions.[3,9] The development of linear accelerators (LINACs) (for delivering photons and electrons) and cyclotrons (for delivering protons and other heavy charged particles) has also contributed to the advancement of EBRT by allowing for the precise delivery of conventional photon or high-frequency accelerated particles directly to the tumor volume. Two of the most common applications are Intensity Modulated Radiation Therapy (IMRT) and Stereotactic Radiosurgery or Stereotactic Body Radiation Therapy (SBRT). IMRT is a further development of 3DCRT; it employs the same image planning and distribution techniques above but goes a step further by altering the intensity (strength) of the beams being delivered, usually lessening the intensity of the beam near OARs. This allows for more control of the level of radiation exposure to surrounding healthy tissues while delivering a high dose to the tumor volume.[24] Initially, this technique had only been applied to photon RT but more recently similar methods have been applied to PBT as well, which is often referred to as Intensity Modulated Proton Therapy (IMPT). In this review, IMPT was a common intervention for the treatment of head and neck cancers in adults and IMRT (with photons) was the most common comparator to PBT for the treatment of brain tumors, esophageal cancers, head and neck tumors, lung cancer, and prostate cancer. Stereotactic Radiosurgery and SBRT are similar to IMRT; however, the beams are delivered in fewer fractions (treatments) and at much higher doses than with IMRT. In addition to dose per fraction, the planning target volume margins are smaller with SBRT, requiring more rigid immobilization. Stereotactic radiosurgery, typically reserved for tumors in the brain and spine, is usually completed in a single session. SBRT is completed in 3 to 5 sessions and is normally used to treat larger tumors in areas of the body other than the brain.[5,24,69] These techniques are advantageous for patients who cannot tolerate surgery or have tumors in locations that are difficult to remove. Stereotactic Radiosurgery and SBRT can be delivered using photons, gamma rays or protons. In the United States, these techniques are most commonly used with photons and gamma rays. More recently, the use of these techniques with protons has emerged but is only offered at a few research centers in the United States. In this review, one study compared stereotactic radiosurgery to PBT for the treatment of ocular (uveal) melanoma.

With treatment planning and delivery techniques evolving similarly between varying types of EBRT, an important difference between modalities lies within the physical properties of each particle and how each reacts with tissue inside the body. Particles have different physical properties and thus their damaging effect on tissue varies.

Photons are uncharged and massless particles that reside within atoms and are characterized by a high deposit of energy near to the body surface with an exponential decrease of energy release as a function of depth.[24] As Figure 1 demonstrates, this has been a challenge for conventional photon therapy due to the amount of radiation deposited both before and after the target is reached. While the amount of photon radiation at entry into the body is much higher than at exit, photon beams typically "hit" normal tissues after leaving the target. In other words, photon beams contain an "exit dose" meaning that healthy tissue downstream from the tumor could be at an increased risk of exposure to unnecessary radiation.

Figure 1. Adapted figure from: Levin WP, Kooy H, Loeffler, DeLaney TF. Proton beam therapy. *Br J Cancer*. 2005;93(8):849-854.



This so-called "exit dose" is absent for protons, as tissue beyond the point of peak energy deposition receives little to no radiation.[42] Protons, heavy positively charged particles, can effectively treat cancerous cells at the end of their path while simultaneously lessening the damage to surrounding healthy tissues, possibly allowing for a greater dose of radiation to be delivered to the target neoplasm(s).[45] This phenomenon is referred to as the Bragg peak, and the total radiation dose, referred to as the "spread out Bragg peak" (SOBP) region is created by varying the energy of a proton beam,

creating a range of energies. For example, a shallower beam will have lower energy compared to a deeper beam (Figure 1). The large mass and acceleration applied to the protons provide each proton with a specific momentum that is mostly dispelled after traveling a defined distance. Protons are slowed down by interactions with their target which results in a sharp burst of energy deposited at the end of its path, followed by no further dose delivery ("exit dose").[78] This physical characteristic distinguishes PBT from other EBRT modalities such as photon RT. In theory, PBT offers physical advantages, though the technology is still new and more prospective clinical comparative evaluations still need to be completed.

It is generally assumed that the biological effects of protons are equivalent to that of photons, but recent studies have shown that the Relative Biological Effectiveness (RBE) of protons in relation to photons are not known with absolute certainty for all types of tissues and fractionation schemes, particularly in adult tumors.[58] However, RBE is dependent on several factors such as dose per fraction, Linear Energy Transfer (LET), tissue radio-sensitivity, particle speed, tissue type, and local microenvironments such as oxygen level.[25] One study identified situations in which RBE was found to be both larger and smaller than 1.1 and another found that ignoring possible variations in RBE could lead to suboptimal PBT treatment plans. The concern with assuming a 1.1 RBE for all tumor types treated with PBT is that it may result in treatment plans that deliver a lower biological dose to the target and a higher biological dose to the normal tissue.[26]

While the dose range is relatively certain for tumors that are close to the skin, there is more uncertainty around the end of the dose range when deep-seated tumors such as prostate cancer are considered.[27] Protons are also very sensitive to tissue heterogeneity, and the precision of the beam may be disturbed as it passes through different types of tissue.[80] Another concern is the effects of neutrons, which are produced by passively-scattered proton beams and result in additional radiation dose to the patient. The location of neutron production in a PBT patient and its biologic significance is currently a topic of significant debate.[33,37]

Policy context/Reason for selection

This topic was originally reviewed in 2014. It is being re-reviewed in 2018 due to newly available published evidence.

Objectives:

The aim of this report is to update the 2014 HTA on proton beam therapy (PBT) by systematically reviewing, critically appraising and analyzing new research evidence on the safety and efficacy of PBT, as a primary or as a salvage therapy (i.e., for recurrent disease or failure of initial therapy), for the treatment of multiple cancer types as well as selected noncancerous conditions in adults and children.

Key questions (Based on previous report):
Inclusion and exclusion criteria are summarized as follows and are detailed in the full report. Briefly, included studies met the following requirements with respect to participants, intervention, comparators, outcomes, and study design:

1. What is the comparative impact of proton beam therapy (PBT) treatment with curative intent on survival, disease progression, health-related quality of life, and other patient outcomes versus

radiation therapy alternatives and other cancer-specific treatment options (e.g., surgery, chemotherapy) for the following conditions:

a. Cancers
    i. Bone tumors
    ii. Brain, spinal, and paraspinal tumors
    iii. Breast cancer
    iv. Esophageal cancer
    v. Gastrointestinal cancers
    vi. Gynecologic cancers
    vii. Head and neck cancers (including skull base tumors)
    viii. Liver cancer
    ix. Lung cancer
    x. Lymphomas
    xi. Ocular tumors
    xii. Pediatric cancers (e.g., medulloblastoma, retinoblastoma, Ewing's sarcoma)
    xiii. Prostate cancer
    xiv. Soft tissue sarcomas
    xv. Seminoma
    xvi. Thymoma
    xvii. Other cancers
b. Noncancerous Conditions
    i. Arteriovenous malformations
    ii. Hemangiomas
    iii. Other benign tumors (e.g., acoustic neuromas, pituitary adenomas)

2. What is the comparative impact of salvage treatment (including treatment for recurrent disease) with proton beam therapy versus major alternatives on survival, disease progression, health-related quality of life, and other patient outcomes versus radiation therapy alternatives and other cancer-specific treatment options (e.g., surgery, chemotherapy) for the condition types listed in key question 1?
3. What are the comparative harms associated with the use of proton beam therapy relative to its major alternatives, including acute (i.e., within the first 90 days after treatment) and late (>90 days) toxicities, systemic effects such as fatigue and erythema, toxicities specific to each cancer type (e.g., bladder/bowel incontinence in prostate cancer, pneumonitis in lung or breast cancer), risks of secondary malignancy, and radiation dose?
4. What is the differential effectiveness and safety of proton beam therapy according to factors such as age, sex, race/ethnicity, disability, presence of comorbidities, tumor characteristics (e.g., tumor volume and location, proliferative status, genetic variation) and treatment protocol (e.g., dose, duration, timing of intervention, use of concomitant therapy)?
5. What is the comparative cost-effectiveness of proton beam therapy in the short- and long-term relative to other types of radiation therapy, radiation therapy alternatives or other cancer-specific treatment options (e.g., surgery, chemotherapy)?

Scope:

Population: Adults and children undergoing treatment of primary or recurrent disease, to include cancer types (bone cancer, brain, spinal, and paraspinal tumors, breast cancer, esophageal cancer, gastrointestinal cancer, gynecologic cancer, head and neck cancer, liver cancer, lung cancer, lymphomas, ocular tumors, pediatric cancers, prostate cancer, sarcomas, seminoma, thymoma,

other cancers) and noncancerous conditions (arteriovenous malformations, hemangiomas, other benign tumors).

Interventions: Proton beam therapy; all approaches were considered including monotherapy, use as a "boost" mechanism to conventional radiation, and combination therapy with other treatment modalities (e.g., chemotherapy, surgery).
Comparators: Primary comparators include other radiation alternatives (e.g., intensity-modulated radiation therapy (IMRT), stereotactic radiation techniques and other external beam therapies, and brachytherapy). Other treatment alternatives specific to each condition type treated, and may include chemotherapy, immunotherapy, surgical procedures, and other devices (e.g., laser therapy for ocular tumors).
Outcomes:
Primary Clinical outcomes:

- Overall survival/disease-free survival
- All-cause and/or disease-related mortality
- Direct measures of tumor regression, control or recurrence
- Incidence of metastases

Secondary or indirect (intermediate) outcomes

- Patient reported outcomes including health-related quality of life (HrQoL) using validated instruments
- Requirements for subsequent therapy
- Other outcomes specific to particular conditions (e.g., visual acuity for ocular tumors, shunt requirements for arteriovenous malformations)
- Intermediate measures of tumor recurrence such as biochemical measures

Safety outcomes:

- Treatment-related harms, to include generalized effects (e.g., fatigue, erythema) and localized toxicities specific to each condition (e.g., urinary incontinence in prostate cancer, pulmonary toxicity in lung or breast cancer); the primary focus is on adverse effects requiring medical attention
- Secondary malignancy risk due to radiation exposure

Economic outcomes:

- Long term and short term comparative cost-effectiveness measures (e.g. incremental cost-effectiveness ratio)

Studies:
The focus will be on high quality (low risk of bias) comparative studies (e.g., randomized controlled trials, comparative cohort studies with concurrent controls) will be considered for Key Questions 1-4. Comparative observational studies with long term clinical outcomes or safety will be considered for Key Questions 1-4. Case series will be considered but will not be the primary focus of evaluation for each key question. Dosimetry and planning studies will be included for context; they will be included as evidence if they directly answer the key questions. Full, comparative, formal economic studies (i.e., cost-effectiveness, cost-utility, cost-minimization, and cost-benefit studies) will be sought for Key Question 5; studies using modeling may be used to determine cost-effectiveness.

Methods

The draft key questions and scope are based on the 2014 report. They were available for public comment. All comments were considered in the finalization of the key questions. Responses to the public comments are posted on the Health Technology Assessment Program's website. Several

commenters provided suggested coverage policies. These are not included in this review as the evaluation or formulation of policy is not the purview of the evidence vendor.

A formal, structured systematic search of the peer-reviewed literature was performed across multiple databases was conducted to identify publications (including clinical guidelines) published subsequent to the original 2014 report, i.e., from November 2013 to December 2018. The search process is detailed in the main report and Appendix B. Reference lists of relevant studies and the bibliographies of systematic reviews were searched. Additionally, a total of 1,426 citations were received from comment received during the Topic Nomination and Draft Key Question public comment phase for this project, of which 390 remained after removal of duplicate citations and elimination of citations published prior to our specified search date range. These 390 studies were reviewed and compared alongside results from the search and included or excluded based on *a priori* criteria outlined in the report. All records were screened by two independent reviewers. Conference abstracts, non-English-language articles, duplicate publications that did not report different data or follow-up times, white papers, narrative reviews, preliminary reports, and incomplete economic evaluations were excluded. A list of excluded articles excluded at full text along with the reason for exclusion is available in Appendix C. Figure 3 in the full report outlines the results for the inclusion/exclusion process.

Consistent with the 2014 report, we focused on comparative studies performing a direct comparison of treatments in the same underlying patient population. Also consistent with the 2014 report, given uncertainties regarding proton physics and the relative biological effectiveness of PBT in all tissues, particularly in adults, only limited appraisal and abstraction of studies included dosimetry, planning and simulation studies included for context was done and focused on any clinical outcomes reported. Studies that did not report on clinical outcomes were not included.

Included studies reporting on primary outcomes of interest were critically appraised independently by two reviewers evaluating the methodological quality, study limitations and potential for bias based on study design as well as factors which may bias studies. Methods of assessing study quality are detailed in the full report. An overall Strength of Evidence (SOE) combined the appraisal of study limitations with consideration of the number of studies and the consistency across them, directness and precision of the findings to describe an overall confidence regarding the stability of estimates as further research is available. The SoE for all primary health outcomes was assessed by two researchers following the principles for adapting GRADE (Grades of Recommendation Assessment, Development and Evaluation) as outlined by the Agency for Healthcare Research and Quality (AHRQ).[2,8,31,32] The strength of evidence was based on the highest quality evidence available from comparative studies for a given outcome. In determining the strength of body of evidence regarding a given outcome, the following domains were considered:

- **Risk of bias**: the extent to which the included studies have protection against bias
- **Consistency:** the degree to which the included studies report results that are similar in terms of effect sizes, range and variability.
- **Directness**: describes whether the evidence is directly related to patient health outcomes or comparisons of interventions are direct (head to head).
- **Precision:** describes the level of certainty surrounding the effect estimates.
- **Publication or reporting bias:** is considered when there is concern of selective publishing or selective reporting. This is difficult to assess particularly for nonrandomized studies.

Bodies of evidence consisting of RCTs are initially considered as High strength of evidence. In general, the GRADE and AHRQ methodologies initially consider nonrandomized studies as Low strength of evidence as such studies typically are at higher risk of bias due to lack of randomization and inability of investigators to control for critical confounding factors. In some instances (e.g. rare conditions, pediatric populations), RCTs may be unavailable, not feasible, not ethical or not substantially applicable to the target populations to be treated and use of high quality nonrandomized observational studies may provide the "best evidence" and may be considered to substitute for RCT evidence.[65] This does not, however, imply that the quality of nonrandomized studies is elevated only that such studies represent the best available evidence and that decision makers need to accept and consider the greater uncertainty of such evidence; one should not have greater confidence in the effect estimates from such studies. Observational studies with few methodologic limitations which control for risk of bias via study conduct or analysis may be initially considered as moderate versus low, particularly for harms and outcomes when such studies may be at lower risk of bias due to confounding.[10] There are also situations where studies (particularly observational studies) could be upgraded if the study had large magnitude of effect or if a dose-response relationship is identified and there are no downgrades for the primary domains listed above and confounding is not a concern.

We compared overall conclusions and findings from the 2014 report with findings in this review to the extent possible based on general qualitative concepts of AHRQ guidance on signal updates for systematic reviews, primarily based on the Ottawa Method.[56,67,71] Considerations included general comparison of study quality for primary outcomes, of whether new evidence constitutes a major change in the evidence based on existence of opposing findings or major changes in effectiveness short of opposing findings based on the highest quality of evidence available, Substantial changes in effect size or changes in statistical significance beyond "borderline" changes, whether new evidence suggests substantial harm wherein risk of harm outweighs benefits and whether new evidence provides high quality data on clinically important expansion of treatment (e.g. to new subgroups of patients) or clinically important caveat.

Due to heterogeneity across studies with regard to designs, patient populations, treatments and clinical methods meta-analysis was not performed.

Results

Out of a total of 2328 citations retrieved by our search strategy, 215 met inclusion criteria. A total of 56 publications were in pediatric tumors, including 13 retrospective comparative cohorts[7,11,12,20,21,28,30,39,40,43,60,64,75], 41 case series, and 2 studies on cost-effectiveness.[35,50] The bulk of the evidence for this section was for the use of PBT in various pediatric brain tumors. A total of 155 publications were in adult tumors, including two RCTs (Liver and Lung cancer),[16,46] one quasi-RCT (Prostate cancer)[41], 33 retrospective comparative cohorts[6,13-15,17,19,22,23,29,34,36,38,47-49,52,53,55,57,59,61-63,66,70,72-74,76,79,81,82,84] and 115 case series; additionally four cost-effectiveness studies were identified.[44,51,54,68] The majority of the evidence in adults was for the following cancers: Esophageal, Head and Neck, Brain, Lung, Ocular, and Prostate. For a list of included case-series please see the full report; the Executive Summary is focused on comparative data only.

The overall quality of the available evidence base was considered poor. Comparative evidence for this report is primarily from retrospective, non-randomized (observational) studies which were considered

to be at moderately high risk of bias except where noted in the detailed description of results. Most studies were retrospective and a number of potential sources of bias must be considered when interpreting study findings. For purposes of this report, prospective comparative cohort studies which controlled for confounding and for which there was ≥ 80% follow-up and ≤10% difference in follow-up between treatments were considered "best evidence" in the absence of quality RCTs. Few studies met all of these criteria. In most instances, treatment groups were formed based on historical changes in methods of radiation therapy delivery, i.e. more conventional photon radiation therapy, including 3DCRT, was delivered to patients at a time prior to a switch to PBT as it became more the available. One consequence of the use of historically consecutive controls in these studies is differential length of follow-up by treatment group; historical groups receiving photon therapy had longer follow-up than those receiving PBT. Differences between treatment groups in patient characteristics, presentation, tumor stage, comorbidities, prior or concurrent treatments and surgical factors were noted in most studies. Although many studies evaluated possible confounding by such factors, there is the possibility of residual confounding or other biases that could influence results.

Comparison with 2014 report

The evidence base in the prior report primarily consisted of case series and focused on comparative studies for evaluation of benefits and harms as does this update. Comparative studies were primarily retrospective cohort studies. In general, the quality of comparative studies in the update report appears to be marginally better but varies somewhat by tumor category. Many studies published subsequent to the prior review had larger sample sizes, made direct comparisons of treatment groups and seemed to employ better methods for controlling for confounding and potential selection bias.

Many of the studies in the 2014 review used 3DCRT and some IMRT as a radiotherapy comparison with PBT; most of the studies in this update used IMRT and/or 3DCRT. The studies in the 2014 report included a variety of comparators, many of which were not represented in the studies included in this review. The prior report included carbon ion therapy as a comparator; it is not included in this review as it is not FDA approved. For some tumor categories, the comparators for studies included in the prior report were very different than comparators, which may reflect changes in clinical practice with time and may partially explain differences in findings between the 2014 report and this review. As an example, for ocular tumors, in the prior report, three studies compared PBT with surgical enucleation and one with transpupillary thermotherapy plus PBT. In this review, some less invasive treatments (brachytherapy and stereotactic radiosurgery) were the comparators employed by included studies. Similarly for hepatocellular carcinoma, the interim RCT analysis included in this review compared PBT with transarterial chemoembolization (TACE) whereas the in the prior review, PBT combined with chemotherapy and carbon ion therapy were the comparators employed in separate studies. Thus, in drawing conclusions across both reviews in such instances, these differences need to be considered. For few tumor classifications RCT data were available in the previous report, but no new RCTs were identified for this review. In addition to heterogeneity in study design and implementation/comparators between included studies for the 2014 and 2019 reviews, specific tumor types and or stages studied in a given classification of tumor may differ between the 2014 and 2019 reports; use of prior or concurrent chemotherapy and other treatments across included studies may also differ within and between reports. Differences in evidence base, comparators and other factors are described with bulleted summary findings for the various tumor classifications.

Table A below provides a broad overview of the strength of evidence and direction of benefits for the 2014 review (based in their table ES2) compared with this 2019 review. (This overview does *not* connote any recommendations for policy). While for many tumor categories, general conclusions regarding

benefits and harms are similar between the two reports, for some tumor types, general conclusions differ. These instances are described with the bulleted summary points for each tumor type.

Table A. Summary of strength of evidence, direction of benefit and general comparison of the 2014 and 2019 report

Condition	Incidence (per 100000)	Number of Publications		Net Health Benefit vs. Comparators Type of Net Benefit (B, H) SOE		Impact of new studies (focus on comparative studies)*
		2014 Report†	2019 Report‡	2014 Report	2019 Report	2014 versus 2019 Report
Adults						
Cancer						
Bladder	20.3	CS=1	CS=1	NR	Insufficient	Similar conclusions
Bone	0.9	CC=1; CS=4	CS=8	Insufficient Low	Insufficient	Similar conclusions
Brain/Spinal	6.5	CC=2; CS=6	CC=5; CS=6	Incremental B: = H: ↓ Low	<u>PBT vs. photon</u> Unclear B: ↑ H: NR Low (curative); <u>PBT boost + photon vs. photon alone</u> Comparable B: = H: = Low (curative) Insufficient (salvage)	3 new retrospective comparative cohorts [2 curative (1 case-matched, 1 large propensity score-matched database) and 1 salvage] of different interventions and tumor types vs. 2014 report. The net health benefit for PBT vs. photon is unclear from 1 large data base study which did not report harms. For PBT boost + photon vs. photon alone, 1 cohort lead to different conclusions regarding harms. Evidence was insufficient for salvage therapy from 1 small cohort.
Breast	124.7	CS=4; Econ=3	CC=2 CS=4; Econ=1	Insufficient none	Unclear B: = H: NR Low	The net health benefit is unclear (addition of 1 large retrospective database study which did not report harms.)
Esophageal	4.6	CC=2; CS=7	CC=5; CS=2	Insufficient none	Incremental B: ↑ H: = Low	New retrospective comparative evidence [5 cohorts (2 propensity score-matched)], leads to different conclusions

Condition	Incidence (per 100000)	Number of Publications		Net Health Benefit vs. Comparators Type of Net Benefit (B, H) SOE		Impact of new studies (focus on comparative studies)*
		2014 Report†	2019 Report‡	2014 Report	2019 Report	2014 versus 2019 Report
GI	100.6§	CS=7	CC=1; CS=2	Insufficient none	Insufficient	Similar conclusions (1 small retrospective comparative cohort, inadequate evidence)
Gynecological	49.8	CS=2	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Head/Neck (oropharyngeal, nasopharyngeal, paranasl sinus, and oral cancers)	17.2§**	CC=1; CS=15; Econ=2	CC=7; CS=14; Econ=1	Insufficient low	Comparable B: = H: = Low	6 additional, larger, retrospective comparative cohorts lead to different conclusions
Head/Neck (Chondro-sarcoma of the skull base)		CC=1 CS=15	CC=1 CS=9	Insufficient low	Insufficient	Similar conclusions (1 small retrospective comparative cohort, inadequate evidence)
Liver	8.1	CC=3; CS=26	RCT=1; CC=1 CS=12; Econ=1	Comparable B: = H: = Low	PBT vs. TACE Incremental B: = H: ↓ Moderate PBT vs. IMRT Incremental B: = H: ↓ Low	RCT interim results with different comparator (TACE). Hospitalization was used as a surrogate for toxicity (see full report). PBT vs. IMRT, larger retrospective comparative cohort. Net health benefit vs. comparators across both reports is unclear.
Lung	60.5	CC=4; CS=19; Econ=2	RCT=1; CC=6††; CS=12	Comparable B: = H: = Low‡‡	Comparable B: = H: = Low	Similar conclusions; addition of a RCT and 5 retrospective comparative cohorts (1 large propensity score-matched database study).
Lymphomas	22.4	CS=1	CS=3	Insufficient none	Insufficient	Similar conclusions
Mixed/Various	N/A§	CC=3; CS=12	CS=3	NR	Insufficient	Similar conclusions

Condition	Incidence (per 100000)	Number of Publications		Net Health Benefit vs. Comparators Type of Net Benefit (B, H) SOE		Impact of new studies (focus on comparative studies)*
		2014 Report†	2019 Report‡	2014 Report	2019 Report	2014 versus 2019 Report
Ocular	0.9	RCT=1; CC=8; CS=45	CC=3; CS=22; Econ=1	Superior (Incremental) §§ B: ↑ H: ↓ Moderate	PBT vs. BT alone Inferior B: ↓ H: = Low PBT + TSR vs. BT + TSR Incremental B: ↑ H: = Low PBT vs. SRS Insufficient	3 additional retrospective comparative cohorts (1 case-matched, and 1 large propensity score-matched database) with very different comparators. Prior report included primarily enucleation (4/7 studies) as comparator, also TTT (1 study); remaining 2 studies were indirect comparisons of case series. The net health benefit across all comparators (across both reports) is unclear.
Prostate	109.2	RCT=1; CC=9; CS=19; Econ=3	Quasi-RCT=1; CC=3; CS=11	Comparable B: = H: = Low‡‡	Comparable B: = H: = Low	Similar conclusions; addition of a quasi-RCT and 3 retrospective comparative cohorts (1 case-matched, 1 large propensity score-matched database)
Sarcomas	4.8§	CS=2	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Seminoma	4.0§	0	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Thymoma	0.2§	0	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Non-cancerous						
AVMs	1.0§	CS=6	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Hemangiomas	2.0§	CC=1; CS=3	CS=2	Comparable B: = H: = Low	Insufficient	Similar conclusions
Pituitary Adenoma	NR§	CS=2	CS=1	N/A	Insufficient	Similar conclusions
Meningioma	2.0§	CC=2; CS=8	CS=3	Insufficient none	Insufficient	Similar conclusions

Condition	Incidence (per 100000)	Number of Publications		Net Health Benefit vs. Comparators Type of Net Benefit (B, H) SOE		Impact of new studies (focus on comparative studies)*
		2014 Report†	2019 Report‡	2014 Report	2019 Report	2014 versus 2019 Report
Pediatric						
Cancer						
All Cancer Types* **	18.3	CC=1; CS=41; Econ=3	CC=13; CS=41; Econ=2	Incremental B: = H: ↓ Low‡‡	See below	See below
Brain	3.1	---	CC=11; CS=25 Econ=2	N/A***	Incremental B: = H: ↓ Low	No comparative studies in the 2014 report; 6 new retrospective cohorts and 2 new prospective cohorts (1 propensity score-matched) suggest incremental net benefit of PBT for pediatric brain tumors
Bone	0.9	---	CS=1	N/A***	Insufficient	N/A
Head/Neck	NR§	---	CC= 1; CS=3	N/A***	Insufficient	N/A
Ocular	0.4	---	CC=1; CS=2	N/A***	Insufficient	N/A
Lymphoma	2.4	---	CS=2	N/A***	Insufficient	N/A
Rhabdomyo-sarcoma	NR§	---	CS=6	N/A***	Insufficient	N/A
Mixed/Various	NR§	---	CS=2	N/A***	Insufficient	N/A

AVM = Arteriovenous Malformation; **B = Benefits**; CC = Comparative Cohort; CS = Case Series; **H = Harms;** N/A = not applicable; IMRT = intensity-modulated radiation therapy (photons). NR = not reported; RCT = Randomized Control Trial; TTT = transpupillary thermotherapy.

*Due to lack of clarity in reported totals of studies, the study totals for the 2014 report here are derived from study lists in the appendix, and may differ from reported totals in body of report.

†All included studies were published subsequent to the prior report. Only studies that provided data on efficacy, effectiveness, safety or cost-effectiveness are included in this table (i.e., contextual studies are not included here).

‡When possible, incidence statistics were updated with more recent data from the National Cancer Institute's Surveillance, Epidemiology, and End Results Program (SEER) database and the American Cancer Society Cancer Statistics Center. Footnoted conditions were either obtained from the incidence values reported in the prior report, not acquirable through the reviewed databases (NR) or not applicable (N/A) because they represented a mixed population.

§Incidence is for head and neck cancers to include skull-base tumors (e.g., chondrosarcoma).

**The comparative cohort count includes the nonrandomized group from the RCT (Liao 2018).

††The prior 2014 PBT report had discrepancies between Table ES2 and Table 3 regarding the Strength of Evidence for Lung Cancer, Prostate Cancer, and Pediatric Cancers. AAI has made the decision to use the Strength of Evidence reported in Table ES2.

‡‡Authors of the 2014 report list the net health benefit as "superior" in their executive summary table. In the report body authors state "Limited, low-quality evidence suggests comparable rates of harm for PBT relative to treatment alternatives in patients with ocular tumors" which suggests that the net health benefit may be more incremental.

§§In the 2014 report, assessment of pediatric cancer was not separated by cancer types.

Summary of Results

Pediatric Tumors

Key points across pediatric tumor categories

- **Pediatric brain tumors**:
  - The bulk of the comparative evidence from studies published subsequent to the 2014 report was for the use of PBT in various pediatric brain tumors. Eight comparative cohort studies at moderately high risk of bias compared PBT with treatment alternatives.
    - Three studies compared PBT with IMRT
    - Two studies compared patients who received PBT with those who received IMRT and/or 3DCRT
    - One study indicated PBT was compared to photon RT with no further specification and one study indicated that those in the comparison group received either 2DCRT or 3DCRT
    - One study compared craniospinal PBT and focal PBT with surgery.
  - Benefits in terms of OS, PFS and tumor recurrence were generally similar between PBT and other forms of radiation therapy across four comparative studies (Low SOE). Some differences may be clinically important.
  - Regarding toxicities and harms, hypothyroidism was less common with PBT versus other RT. Low SOE) Many other toxicities (including other endocrine-related toxicities) tended to be less frequent in those receiving PBT vs other RT, however statistical significance was generally not reached, likely due to study sample sizes and possibly residual confounding. (Low SOE) Some differences may be clinically important. One prospective cohort study reported declines for full scale intelligence quotient (FSIQ) and processing speed index scores when craniospinal PBT was compared with surgery but no differences between focal PBT and surgery for any score. The clinical relevance of the declines was not described. One retrospective cohort reported no difference between PBT and photon therapy for FSIQ scores (Low SOE for all outcomes.)
  - While two poor-quality full economics studies suggest that PBT may be cost-effective for treatment of pediatric brain or CNS tumors vs other types of radiation, the limitations of these studies need to be considered.
  - None of the included studies evaluated differential effectiveness or safety.
- **Other pediatric tumors:**
  - Evidence for effectiveness and safety was considered to be insufficient for all other pediatric tumors. Studies published subsequent to the 2014 report were identified for the following pediatric tumor categories: head and neck, soft tissue (rhabdomyosarcoma), ocular, lymphoma, bone and one study of mixed tumor types. Evidence was primarily from case series, with only two small comparative (one for salivary gland tumors, the other salvage treatment in ocular tumors) identified.
  - No full-economic studies or studies designed to evaluate differential effectiveness or safety were identified.

Adult Tumors/Conditions

Key points *across* adult tumor categories/conditions

Bladder cancer

- There is insufficient evidence from one case series to evaluate the effectiveness or safety of PBT for bladder cancer in adults.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.

Bone tumors

- There is insufficient evidence from seven case series to evaluate the effectiveness or safety of PBT for bone tumors in adults.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.

Brain, Spinal, Paraspinal tumors

- Results were inconsistent across two retrospective case-matched cohorts evaluating adult patients with different types of brain tumors undergoing treatment for curative intent. In one retrospective cohort, there was no statistical difference in the probability of 1-3 year OS and 1-2 year PFS following photon RT plus a PBT boost versus photon RT alone in patients with high-grade glioblastoma; those receiving PBT boost tended to have higher PFS but lower OS versus those receiving photon alone and differences may be clinically meaningful. One large database study of primarily high-grade glioma reported statistically higher 5-year overall survival following PBT alone versus photon RT alone. (Low SOE for both comparisons).
- One small retrospective cohort study in patients with metastatic CNS disease found no statistical difference between salvage PBT compared with photons in the probability of 6-month OS or of CNS relapse; at 1 year, OS was better in the PBT group but statistical testing was not done and sample size was small (Insufficient SOE).
- For safety, no statistical differences were seen between groups in the frequency of acute grade 3 toxicity across both studies or of radiation necrosis (1 study of curative intent) or severe CNS toxicity (1 study of salvage therapy) over the late term (Low SOE for curative intent; Insufficient SOE for salvage therapy).
- No studies meeting inclusion criteria were identified that evaluated differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Breast cancer

- There is low strength of evidence from one retrospective comparative database study that there is no statistical difference in the probability of OS at 5 years between PBT versus photon with or without electron boost therapy for treatment of breast cancer.
- One moderate quality cost-utility study (QHES 73/100) concluded that, compared with photon therapy, PBT was not cost effective in women without cardiac risk factors (CRF) or PBT mean heart radiation doses <5 Gy (RBE). PBT is more likely to be cost-effective for patients with higher risk of coronary heart disease (CHD) and for younger patients (40 or 50 years old versus 60 years old); authors indicate a societal perspective, however indirect societal costs were not described.

- No studies meeting inclusion criteria were identified that evaluated salvage therapy or differential effectiveness and safety in this population.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Esophageal tumors

- Five retrospective comparative cohort studies that evaluated the effectiveness and safety of PBT compared with photon RT for curative intent in adult patients with esophageal cancer that met inclusion criteria were identified.
- With the exception of OS at 1 year which was similar between groups, probabilities of OS and PFS/DFS were greater following PBT versus IMRT or 3D-CRT over 1 to 5 years follow-up in two studies; however, statistical significance was achieved in only the largest study (Low SOE).
- Mortality (as opposed to OS) was reported by two studies with no statistically significant differences seen between the PBT and the photon groups (IMRT, 3D-CRT, XRT) (Low SOE for the large, higher quality study; Insufficient SOE for the small, poorer-quality study).
- For the comparison of PBT versus IMRT, with the exception of grade 4 radiation-induced lymphopenia (2 studies) and any would event (1 study) which were less common with PBT, all other RT-related and treatment-related toxicities did not differ statistically between groups. For PBT versus 3DCRT or XRT, with the exception of GI events, PBT was associated with a statistically less treatment-related toxicity (i.e., pulmonary, cardiac, and wound events; grades ≥2 or not specified) across three studies (Low SOE for all).
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Gastrointestinal (pancreatic) tumors

- One small retrospective cohort study that compared PBT with hyper-fractionated accelerated radiotherapy (HART) for curative intent in adult patients with locally advanced and unresectable pancreatic adenocarcinoma reported no statistically significant differences between groups in the probability of 1- to 3-year OS, disease control/local progression or metastases or in the frequency of grade ≥3 radiation-related hematological or non-hematological toxicities which were rare; clinical importance of differences is unclear (Insufficient SOE).
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Head and Neck tumors (including skull-base)

- Across three retrospective cohort studies, the probabilities of 1- to 3-year OS and PFS (one case-matched study, primary oropharyngeal cancer), the incidence of all-cause mortality over a median 24 months (one small study, primary nasopharyngeal cancer), and 1-year OS (one small study, primary or metastatic salivary gland cancer) were not statistically different between PBT and IMRT groups. Clinical significance of differences is unknown. (Low SOE for primary oropharyngeal and nasopharyngeal cancer; Insufficient SOE for primary or metastatic salivary gland cancer).

- Across three retrospective comparative studies evaluating different tumor types (primary oropharyngeal; primary nasopharyngeal; and primary or metastatic salivary gland cancer), there were no statistically significant differences in the frequency of grade ≥3 acute or late toxicities or the incidence of ED visits/unplanned hospitalizations (1 study) following PBT versus IMRT (Low SOE based on largest, best quality study). A third retrospective comparative study in oropharyngeal cancer reported no statistical difference in the incidence of osteoradionecrosis after 6 months between PBT and IMRT (Insufficient SOE).
- Across five retrospective comparative cohorts evaluating different tumor types (2 primary oropharyngeal; 1 each of primary nasopharyngeal; primary nasopharyngeal or paranasal sinus; and primary or metastatic salivary gland cancer), gastrostomy tube dependence tended to be lower with PBT, however adjusted estimates from the largest study were not statistically significant, while smaller studies reported statistically significant differences. For the smallest study, the large confidence interval suggests instability of the effect estimate. Clinical significance of differences is unclear. It is unclear what role differences in study populations (including tumor characteristics, etc.) and possible residual confounding may play in these findings.
- One good quality cost-effectiveness analysis (QHES 90/100) took both societal and payer perspectives and concluded that, compared with IMRT, PBT was not cost-effective for patients with stage III-IV oropharyngeal squamous cell carcinoma using either perspective. However, at extremes of PBT superiority, it becomes cost-effective for younger human papilloma virus (HPV)-positive patients.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy (i.e., no comparative studies) or differential effectiveness and safety in this population.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Liver tumors

- No statistical differences were seen between PBT and transarterial chemoembolization (TACE) for the probabilities of 2-year OS, PFS, and local control in one small RCT of adult patients with unresectable hepatocellular carcinoma (HCC) treated with curative intent, though PFS and local control tended to be greater following PBT (Moderate SOE).
- OS was statistically higher following PBT versus intensity-modulated radiation therapy (IMRT) in one retrospective cohort study of adult patients with unresectable HCC but there was no difference in local and regional control between groups (Low SOE).
- Acute toxicity and serious complications were not well described in the RCT. Fewer patients who received PBT compared with TACE were hospitalized for a complications within 30 days of treatment, translating into fewer total days hospitalized for complications (Moderate SOE). In the retrospective cohort study, compared with IMRT, PBT was associated with a lower risk of nonclassic radiation-induced liver disease (RILD) (Low SOE) and death due to liver failure (Insufficient SOE).
- One poor quality cost-utility analysis (QHES 51/100) from Taiwan compared PBT with stereotactic body radiation therapy (SBRT) for a hypothetical cohort of patients with advanced, inoperable hepatocellular carcinoma using Markov modeling from a payer perspective and concluded that PBT is cost-effective for high risk patients at a willingness-to-pay threshold (WTP) of New Taiwan Dollars $2,157,024 per quality-life years (QALY) gained.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy (i.e., no comparative studies) or differential effectiveness and safety in this population.

- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Lung

- In one fair-quality RCT, no statistically significant differences were seen between PBT versus IMRT in the probability of OS at any timepoint up to 5 years or in the cumulative incidence of local failure in patients with non-small cell lung cancer being treated with curative intent (Moderate SOE). Findings from four retrospective comparative cohort studies were consistent with those of the RCT.
- For safety, no statistical differences were seen between PBT and IMRT in the frequency of grade ≥3 radiation pneumonitis at any timepoint up to 5 years in the fair-quality RCT (Moderate SOE). There was insufficient evidence from two retrospective cohort studies regarding grade ≥3 toxicities (radiation pneumonitis, radiation esophagitis, and radiation dermatitis) which did not differ statistically between PBT and IMRT; clinical importance of differences in unknown.
- The one comparative study of salvage PBT did not report survival or safety data; no studies that met inclusion criteria were identified that provided data on differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Lymphoma

- There is insufficient evidence from three case series to evaluate the effectiveness and safety of PBT for curative intent in adults (primarily) with Hodgkin or non-Hodgkin lymphoma.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.

Ocular tumors

- Across two retrospective cohort studies in patient with ocular tumors comparing PBT with brachytherapy or stereotactic radiosurgery (SRS) for curative intent, there were no statistically significant differences in OS at 2 years and mortality at 3 years; at 5-years PBT was associated with a statistically higher risk of mortality with PBT vs. brachytherapy in the larger, higher quality study (Low SOE).
- PBT was associated with a statistically lower frequency of local recurrence over 10 years compared with brachytherapy in one retrospective comparative cohort study (Low SOE). A second, poorer quality study comparing PBT versus stereotactic radiosurgery found no difference between groups in local recurrence at 3 years, however the strength of evidence was insufficient.
- With the exception of optic neuropathy which was statistically lower following PBT versus SRS in one study, no other statistical differences were seen in the frequency of adverse events (radiation retinopathy, enucleation, rubeosis of the iris, neovascular glaucoma, rubeotic glaucoma) over 3 years between PBT versus brachytherapy or SRS across two retrospective comparative cohort studies.
- One good quality (QHES 93/100) concluded that, compared to enucleation, PBT was not cost-effective for patients with intraocular melanoma using a WTP of $50,000/QALY based on a payer perspective. However, results ranged from cost-effective ($9,522/QALY) to very expensive ($441,750/QALY) in sensitivity analyses. PBT cost was a significant driver of the ICER.

- No studies meeting inclusion criteria were identified that evaluated salvage therapy (i.e., no comparative studies) or differential effectiveness and safety in this population.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Prostate tumors

- In one quasi-RCT, there were no statistically significant differences in the probabilities of 5- and 10-year overall survival and biochemical relapse-free survival between the combined photon and PBT boost group and the photon only group (Low SOE).
- The probabilities of acute and late grade 2 gastrointestinal (GI), but not genitourinary (GU), toxicity were significantly lower in patients who received the photons plus PBT boost versus photons only in one quasi-RCT; however, there were no statistically significant differences for grade 3 or 4 toxicities. Across three retrospective cohort studies comparing PBT with IMRT results regarding acute and late GU and GU toxicity differed, with two finding no statistical difference between groups and the third, a large database study, reporting lower cumulative incidences with PBT (to include erectile dysfunction) compared with IMRT; differences between groups were small and clinical significance is unknown (SOE Low for all).
- No studies that met inclusion criteria were identified that provided data on PBT for salvage therapy, differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

BENIGN TUMORS

Hemangiomas (Adults)

- There is insufficient evidence from two case series to evaluate the effectiveness and safety of PBT for curative intent in adults with hemangiomas.
- No studies meeting inclusion criteria were identified that evaluated salvage PBT, differential effectiveness and safety or cost-effectiveness.

Other Benign Tumors

- There is insufficient evidence from three case series to evaluate the effectiveness and safety of PBT for other non-cancerous tumors (i.e., meningioma, pituitary adenoma).
- No studies meeting inclusion criteria were identified that evaluated salvage PBT (i.e., no comparative studies), differential effectiveness and safety or cost-effectiveness.

Mixed/Various Tumors

- There is insufficient evidence from three case series to evaluate the effectiveness and safety of PBT for mixed tumor populations.
- No studies meeting inclusion criteria were identified that evaluated salvage PBT, differential effectiveness and safety, or cost-effectiveness in mixed tumor populations.

Strength of Evidence Summary for Pediatric Brain, Spinal, and Paraspinal Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrade	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
KQ 1 Curative intent					
Survival outcomes					
Probability, overall survival	3 year	Sato 2017 † (N=79) Ependymoma	**Consistency** Unknown (different tumor types) **Serious Imprecision** Yes[3] (-1)	97% (83%-99%) vs. 81% (63%-90%); NR; p=0.08	PBT resulted in similar (3 studies, Bishop, Kopecky, Eaton) or slightly greater (2 studies, Sato, Gunther) OS compared with IMT or CRT however statistical significance was not reached in any study at any time; sample sizes may play a role. ⊕⊕◯◯ LOW
		Bishop 2014 (N=52) Craniopharyngioma		94.1% (NR) vs. 96.8% (NR); NR; p=0.742	
	4 year	Gunther 2015† (N=72) Ependymoma		87.5% (51.6% - 97.3%) vs. 78.8% (60.6% - 89.3%); NR; p=0.21	
	5 year	Kopecky 2017§ (N=783) Medulloblastoma		%NR HR 0.99 (0.41 to 2.4); p=0.98 (PBT vs. CRT)	
	6 year	Sato 2017† (N=79) Ependymoma		88% (NR) vs. 70% (NR)‡ NR	
		Eaton 2016a,b (N = 88); Medulloblastoma		82.0% (65.4% - 91.1%) vs. 87.6% (72.7% - 94.7%); adj. HR, 2.17 (0.66 to 7.16)	
Probability, Progression free or relapse free survival	3 year	Sato 2017† (N=79) Mod high Ependymoma	**Consistency** Unknown (different tumor types) **Serious Imprecision** Yes[3] (-1)	PFS: 82% (64%-92%) vs. 60% (42%-74%); HR (vs IMRT), 0.42 (0.16-1.10)	At 3 and 6 years, PFS in patients with ependymoma who received PBT tended to have longer PFS vs. IMRT, but differences were not statistically significant at 3 years. RFS was similar between groups in patients with medulloblastoma ⊕⊕◯◯ LOW
	6 year	Eaton 2016a,b (N=88); Medulloblastoma		RFS: 78.8% (63% - 89%) vs. 76.5% (60.6% - 86.6%); adj. HR 1.31 (0.5 to 3.41)	
		Sato 2017† (N=79) Ependymoma		PFS: 82% (NR) vs. 38% (NR) p=NR	
Other Primary					
Any recurrence or relapse	74.4 mos. vs. 85 mos.	Eaton 2016a (N=88) Medulloblastoma	**Consistency**	22.2% (10/45) vs. 23.3% (10/43); NR	Recurrence was similar between groups in patients

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrade	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
	31.2 vs. 58.8 mos.	Sato 2017 † (N = 79) Ependymoma	Unknown (different tumor types) **Serious Imprecision** Yes[3] (-1)	17% (7/41) vs. 55% (21/38); RR 0.31 (0.15 to 0.64)	with medulloblastoma however was significantly less common in patients with ependymoma ⊕⊕◯◯ LOW
KQ 3: Safety Outcomes					
Hypothyroidism	56.4 mos. vs. 121.2 mos.	Bielamowicz (N=84) Medulloblastoma	**Serious Imprecision** Yes[3] (-1)	**Hypothyroidism (any):** 19% vs. 46.3%; adj. HR 1.85 (0.8 to 4.2) **Primary hypothyroidism:** 7.3% vs. 20.4%; adj HR 2.1 (0.6 to 7.7) **Central hypothyroidism:** 9.8% vs. 24.0% ; adj HR 2.2 (0.7 to 6.6)	Across 2 studies, hypothyroidism was less common with PBT statistical differences were only seen in one study ⊕⊕◯◯ LOW
	69.6 mos. vs. 84 mos.	Eaton 2016b (N=77) Medulloblastoma		**Hypothyroidism:** 22.5% (9/40) vs 64.9% (24/37); adj OR: 0.13 (0.04 to 0.41)	
Other Endocrine toxicities	33.1 mos. vs. 106.1 mos.	Bishop 2014 (N=52) Craniopharyngioma	**Consistency** Unknown **Serious Imprecision** Yes[3] (-1)	**Panhypopituitarism:** 33% (7/21) vs. 55% (17/31); RR 0.61 (0.31, 1.2) **Other** endocrinopathy: 43% (9/21) vs. 23% (7/31); RR 1.9 (0.84, 4.3)	Other specific endocrinopathies across the two studies tended to be less common in PBT recipients compared with other forms of radiation therapy; however, statistical significance was only achieved for sex hormone deficiency. Endocrine replacement therapy was less common in those receiving PBT vs. photon RT.
	69.6 mos. vs. 84 mos.	Eaton 2016b (N=77) Medulloblastoma	**Consistency** Unknown **Serious Imprecision** Yes[3] (-1)	**Growth hormone deficiency:** 52.5% (21/40) vs. 56.76% (21/37); adj OR 0.81 (0.26 to 2.59) **Sex hormone deficiency**: 2.5%	

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrade	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
				(1/40) vs. 18.92% (7/37); adj OR 0.06 (0.01 to 0.55) **Endocrine replacement therapy**: 55% (22/40) vs. 78.38% (29/37) adj OR 0.30 (0.09 to 0.99)	⊕⊕◯◯ LOW
Changes in IQ score changes per year	32.4 mos. vs. 64.8 mos.	Kahalley (N=150) Various brain tumors	**Consistency** Unknown **Serious Imprecision** Yes[3] (-1)	**FSIQ (adjusted beta coefficient, 95%CI)** **PBT vs. Photon RT**** **All patients** -0.7 (-1.6 to 0.2) vs. -1.1 (-1.8 to -0.4; p= 0.51 **CSI: -** 0.8 vs. -0.9 (CIs NR); p = 0.89 **Focal RT:** 0.6 (-2.0 to 0.8) vs. -1.6 (-3.0 to -0.2); p = 0.34	There were no differences between PBT and photon radiation in with regard to changes in IQ scores. ⊕⊕◯◯ LOW
	33.6 mos. to 37.2mos. post-treatment	Kahalley 2019 (N=93) Various brain tumors	**Consistency** Unknown **Serious Imprecision** Yes[3] (-1)	**<u>Focal PBT vs. surgery</u>** NS differences FSIQ or for any subscale (all p-values >0.05); scores remained stable for both groups over time. **<u>CSI PBT vs. surgery</u> (adjusted beta coefficient, 95% CI)**** FSIQ: -2.1 (-3.8 to -0.3), p = 0.020 PSI; -2.6 (-4.7 to -0.3), p = 0.019. NS differences for any other subscales (all p-values >0.05)	There were no differences between focal PBT and surgery in changes in FSIQ or subscores after adjustments for baseline differences. CSI PBT was associated with a decline in FSIQ and PSI with time compared with surgery. The clinical significance of the changes is not described. ⊕⊕◯◯ LOW
Other Late toxicities or adverse events	PBT 33.1 mos. vs. 106 mos.	Bishop 2014 (N=52) Craniopharyngioma	**Consistency** Unknown	**Vascular Injury (on imaging),** 10% (2/21) vs. 10% (3/31);	Risk of vascular injury, hearing loss and radiation necrosis were similar

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrade	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
(Median f/u by treatment)			**Serious Imprecision** Yes[3] (-1)	**Vision changes:** 5% (1/21) vs. 13% (4/31); RR 0.37 (0.04, 3.07) **Hypothalamic obesity:** 19% (4/21) vs. 29% (9/31); RR 0.66 (0.23, 1.9)	between PBT and other types of RT; although risk of vision changes and hypothalamic obesity were somewhat lower for PBT in one study, groups were not statistically different. ⊕⊕◯◯ LOW
	55.5 mos. vs.65.5 mos.	Paulino 2018 (N=84) Medulloblastoma	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Serious Imprecision** Yes[3] (-1)	**Hearing Loss (worse ear)** *Grade 3:* 26.3% (10/38) vs. 21.7% (10/46) *Grade 4:* 2.6% (1/38) vs. 6.5% (3/46) *Grade 3 and 4:* 29.9% (11/38) vs. 28.3% (13/46), p=1.0	
	31.2 mos. vs. 58.8 mos.	Sato 2017 (N = 79) Ependymoma	**Consistency** Unknown **Serious Imprecision** Yes[3] (-1)	**All events:** 7.3% (3/41) vs. 13.2% (5/38); RR 0.56 (0.14, 2.17) **Radiation Necrosis**: 7.3% (3/41) vs. 7.9% (3/38) **Stroke**: 0% (0/41) vs. 2.6% (1/38) **Cavernoma:** 0% (0/41) vs. 2.6% (1/38)	
Acute Toxicities	Acute	Song 2014 (n=30 PBT, n=13 photon) Various tumors	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Serious Imprecision** Yes[3] (-1)	**Leukopenia** Grade 3: 57% (14/30) vs. 46% (6/13) Grade 4: 7% (2/30) vs. 31% (4/13) **Grade 3 or 4** RR: 0.68 (0.44, 1.08) **Anemia** Grade 3: 0% (0/30) vs. 15% (2/13); p=0.493 Grade 4: 0% (0/30) vs. 0% (0/13) **Thrombocytopenia:** - Grade 3: 20% (6/30) vs. 31% (4/13) - Grade 4: 3% (1/30) vs. 23% (3/13); **Grade 3 or 4 RR**: 0.43 (0.19, 0.98)	Frequency of acute Grade 3 or 4 hematological toxicities was lower with PBT vs. photon RT, however the overall sample size is small, particularly in the photon group. There is insufficient evidence to draw conclusions. ⊕◯◯◯ INSUFFICIENT

adj RR= adjusted risk ratio; CI = Confidence Interval; f/u = follow-up; FSIQ = Full Scale Intelligence Quotient; HR = Hazard Ratio; IMRT = Intensity Modulated Radiation Therapy; IQ = Inteligence Quotient; NR = Not Reported; NS = Not significant; OR = Odds ratio; OS = Overall Survival; PBT = Proton Beam Therapy; PFS = Progression Free Survival; PSI = Processing Speed Index; RFS = Recurrence Free Survival; RR = crude Risk Ratio; RT = Radiation Therapy; SOE = Summary of Evidence
* PBT was compared with IMRT in Bishop, Gunther and Sato; IMRT or 3DCRT was used in Eaton; Kopecky had 3 arms; PBT, IMRT and 2D/3D CRT but effect sizes were only reported for PBT vs. 2D/3D CRT not for PBT vs. IMRT;
† Sato and Gunther report on the same underlying patient population. Sato 6 year estimates from author's graph
‡PBT was done as "definitive" treatment in 13% and post-operative/adjuvant treatment in 44%, salvage treatment in 42%
§ 517 pts (of the 1300 identified) diagnosed after 2009 were excluded from survival analysis leaving 783 for survival analysis across three treatment groups but authors do not specify to which treatment group they belong or the number of patient with PBT and CRT which were compared in survival analysis
** Authors do not provide mean changes only beta coefficients and p-values; Beta coefficients represent the increase or if negative, decrease in points per year on each index by treatment group. Inclusion of 0 in the confidence interval signifies results are not statistically significant.

Reasons for downgrade:
1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Pediatric Head and Neck Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrade	PBT vs. other RT * Effect estimate (95% CI)	Conclusion Quality (SoE)
Head, Neck					
Toxicity	Acute	Grant (N=24) 1 Retro cohort (N=24) Salivary Gland tumors	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Serious Imprecision** Yes[3] (-1)	Grade 2/3 acute toxicities: Dysphagia (0 vs. 3/11), Otitis externa (1/13 vs. 2/11), Mucositis (6/13 vs. 10/11, RR 0.51 (0.27, 0.94).	Mucositis may be less common following adjuvant PBT vs. adjuvant photon RT; risk of other toxicities was similar between groups. ⊕◯◯◯ INSUFFICIENT
Ocular (Salvage)					
Effectiveness	Last f/u 3 yrs. PBT, 10 yrs. RT	Agarwal 2016 (N=39 patients, 47 eyes) Retinoblastoma	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Serious Imprecision** Yes[3] (-2)	OS: 97.4% across groups Enucleation-free survival: 38.5% vs. 54.5% Enucleation performed: 37.5% (6/16 eyes) vs. 29.6% (8/27 eyes)	No comparative data reported for OS. Enucleation-free survival was lower with PBT, however small sample size, may preclude detection of statistical difference ⊕◯◯◯ INSUFFICIENT
Toxicity	Acute Late			**Acute Toxicity:** PBT 93.8% vs. ERT 74.1%; p =0.22 (mostly skin erythema) **Late/long-term** (number of eyes): **PBT vs. ERT** Any (≥1 event): 62.5% (10/16 eyes) vs. 55.6% (15/27 eyes); p=0.275 **PBT vs. Other Tx** Cataract: 5 vs. 10 Vitreous hemorrhage: 3 vs. 4 Radiation retinopathy: 2 vs. 3 Visual acuity Δ: 0 vs. 4 Strabismus: 1 vs. 2	Although acute toxicities were more common with PBT vs. ERT, differences were not statistically significant. Evidence is limited ⊕◯◯◯ INSUFFICIENT

adj RR= adjusted risk ratio; CI = Confidence Interval; f/u = follow-up; ERT= electron beam radiation therapy; HR = Hazard Ratio; IMRT = Intensity Modulated Radiation Therapy; NR = Not Reported; NS = Not significant; OR = Odds ratio; OS = Overall Survival;

PBT = Proton Beam Therapy; PFS = Progression Free Survival; RFS = Recurrence Free Survival; RR = crude Risk Ratio; RT = Radiation Therapy; SOE = Summary of Evidence

* Grant compared PBT (passive scatter n =8, intensity modulated n=5) vs. other RT (electron bean n=8, IMRT n=3); Agarwal compared PBT (passive scatter, n= 16 eyes) vs. photon or electron RT (n=27 eyes) and brachytherapy (n= eyes).

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary: Adult Tumors

Strength of Evidence Summary for Adult Brain, Spinal, Paraspinal Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon* Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)					
Survival outcomes					
Probability, overall survival (OS)†	1-3 years	Adeberg 2017 (N=132) Retro case-matched cohort Glioblastoma (high-grade)	**Inconsistency** Unknown **Imprecision** Yes[3] (-1)	*PBT boost + photon vs. photon alone:* 1 year: 75% vs. 85% 2 years: 40% vs. 43% 3 years: 12% vs. 28% p=NS at all timepoints	Results across studies and tumors types are inconsistent For those with high-grade glioblastoma, PBT boost tended to result in lower OS but higher PFS probability versus photon alone; results were not statistically significant but may be clinically meaningful. In the large database study of primarily high-grade glioma, statistically higher 5-year overall survival was reported following PBT versus photon RT. Of note, the median follow-up period was significantly shorter in the PBT group (50.3 vs. 62.3 months). There is the potential for misclassification in database studies. ⊕⊕◯◯ LOW
	5-years	Jhaveri 2018 (N=49,575) Retro comparative database study; propensity-score matched cohort (n=322) Glioma (91% high-grade)	**Inconsistency** Unknown	*PBT vs. any photon, entire cohort:* adj. HR 0.66, 95% CI (0.53 to 0.83); favors PBT *PBT vs. any photon, propensity-score matched:* 46.1% vs. 35.5%, p=0.009 vs. IMRT: p=0.01 vs. 3D-CRT: p=0.007	
Probability, Progression free survival (PFS)†		Adeberg 2017 (N=132) Retro case-matched cohort Glioblastoma (high-grade)	**Inconsistency** Unknown **Imprecision** Yes[3] (-1)	*PBT boost + photon vs. photon alone:* 1 year: 31% vs 21% 2 years: 8% vs 2% p=NS at both timepoints	
Salvage therapy (KQ2)					
Survival and recurrence outcomes					
Probability, overall survival	6 mos. to 1 year	Gunther 2017 (N=37) Retro cohort CNS involvement in lymphoma or leukemia (pre-SCT)	**Risk of Bias** Yes[1] (-1) **Inconsistency** Unknown **Imprecision** Yes[3] (-1)	*PBT vs. Photon:* 6 mos.: 78.6% vs. 69.6%, p=0.15 1 year: 70% vs. 38%,‡ p=NR	No statistical difference between groups in OS at 6 months, statistical testing not reported at 1 year; no statistical difference in CNS relapse risk. Sample size

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon* Effect estimate (95% CI)	Conclusion Quality (SoE)
CNS relapse	5 mos.			*PBT vs. Photon:* 7% (1/14)§ vs. 0% (0/23); p=1.0	may have played a role in these findings. ⊕◯◯◯ INSUFFICIENT
Safety (KQ3)					
Acute Toxicity (≤3 mos.)	Median 15 mos.	Adeberg 2017 (N=132) Retro case-matched cohort Primary Glioblastoma (high-grade)	**Imprecision** Yes[3] (-1)	*PBT boost + photon vs. photon alone:* Grade ≥2: 9% (6/66) vs. 14% (9/66), p=NR; Grade 3: 0% (0/66) vs. 7.5% (5/66), p<0.1	No statistical differences between groups; unclear if differences may be clinically important. Sample size may have played a role in these findings. ⊕⊕◯◯ LOW
	During CSI	Gunther 2017 (N=37) Retro cohort CNS-involvement in leukemia/ lymphoma Salvage therapy (pre-SCT RT)	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	*PBT vs. Photon* Mucositis, Grade 3: 7% (1/14) vs. 9% (2/23), p=0.1; Mucositis, any Grade: 7% (1/14) vs. 44% (10/23); RR 0.16 (0.02 to 1.15)**; Gastrointestinal (Grade NR): 29% (4/14) vs. 30% (7/23), p=1.0; CNS (Grade NR): 21% (3/14) vs. 13% (3/23), p=0.65	PBT resulted in a lower frequency of mucositis (any grade); no other differences were seen over acute or late term. Sample size may have played a role in these findings. ⊕◯◯◯ INSUFFICIENT
	"Late"			*PBT vs. Photon* Severe CNS neurotoxicity††: 7% (1/14) vs. 0% (0/23), p=NS	
Radiation necrosis (outside of treatment field)	Median 15 mos.	Adeberg 2017 N=132) Retro case-matched cohort Primary Glioblastoma (high-grade)	**Inconsistency** Unknown **Imprecision** Yes[3] (-1)	*PBT boost + photon vs. photon alone:* 0% (0/66) vs 0% (0/66)	No cases of radiation necrosis outside the treatment field in either group. Sample size may have played a role in the findings. ⊕⊕◯◯ LOW
Change in symptomology, % (n/N)	Median 15 mos.	Adeberg 2017 N=132)	**Inconsistency** Unknown **Imprecision** Yes[3] (-1)	*PBT boost + photon vs. photon alone:* Neurocognitive deficits‡‡	No statistical differences between groups in the proportion of patients experiencing either

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon* Effect estimate (95% CI)	Conclusion Quality (SoE)
		Retro case-matched cohort Primary Glioblastoma (high-grade)		**Worse** (vs. baseline): 3% (2/66) vs. 6% (4/66) **New**: 9% (6/66) vs. 2% (2/66) Sensorimotor deficits‡‡ **Worse** (vs. baseline): 3% (2/66) vs. 5% (3/66) **New**: 11% (7/66) vs. 14% (9/66) Seizures‡‡ **Worse** (vs. baseline): 0% (0/66) vs. 0% (0/66) **New**: 2% (1/66) vs. 6% (4/66) p=NS for all	worsening of preexisting symptoms or new deficits following treatment ⊕⊕◯◯ LOW

CNS = central nervous system; CI = confidence interval; CSI = craniospinal irradiation; KQ = Key Question; NR = not reported; NS = not statistically significant; PBT = proton beam therapy; Retro = retrospective; SCT = stem cell transplantation; SOE = strength of evidence.

* **Adeberg 2017**: Photon + PBT boost vs. Photon alone.

Gunther 2017: PBT (passive scatter) vs. Photon.

Jhaveri 2018: PBT vs. photons (IMRT, 3D-CRT, and other photon not specified).

†All data estimated from graphs provided by authors.

‡Estimated from graph in article.

§Also had concurrent systemic relapse and died from disease.

**Crude RR calculated by AAI using exact methods in Stata.

††Characterized by diffuse demyelination and necrosis, neurocognitive impairment, lower extremity weakness, incontinence, difficulty swallowing

‡‡ Authors describe these as/along with toxicity. As baseline in the PBT vs. photon groups, neurocognitive deficits, sensorimotor deficits, and seizures were presents in 30% (20/66) vs. 42% (28/66), 39% (26/66) vs. 30% (20/66), and 6% (4/66) vs. 3% (2/66), respectively. The majority of patients with pre-therapeutic deficits showed a stable deficit level after radiotherapy.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Breast Cancer for Effectiveness

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon/Electron Boost* Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)					
Survival outcomes					
Probability, overall survival (OS)	5 years	Chowdhary 2019 (N=724,492) Retro comparative database study (NCDB)	**Inconsistency** Unknown	91.9% vs. 88.9% (unadjusted probabilities) Adjusted HR† 0.85 (95% CI, 0.68 to 1.07), p=0.12 A second additional multivariate analysis conducted after stratifying for factors associated with increase heart doses also showed no difference.	No statistical difference between PBT versus photon/electron boost therapy for the probability of OS at 5 years. ⊕⊕◯◯ LOW

CI = confidence interval; KQ = Key Question; NCDB = National Cancer Data Base; PBT = proton beam therapy; Retro = retrospective; SOE = strength of evidence.

*Aside from the breast, additional lymph node irradiation was indicated in 22% of patients. Other treatments received included chemotherapy in 46% and endocrine therapy in 69%.

†In multivariate analysis, adjusted for: race, Charlson-Deyo comorbidity score, facility (academic vs. nonacademic), household income, regional location, residence (urban vs. rural), laterality, pT-stage, pN-stage, receptor status, receipt of chemotherapy, receipt of endocrine therapy, type of surgery, and year of diagnosis.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Esophageal Cancer for Effectiveness

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
Probability, overall survival (OS)	1 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	88% vs. 85%‡ Log-rank, p=0.01	Probabilities of OS at 1 year were similar, however, over subsequent years OS was better following PBT vs. IMRT or 3DCRT across both studies. However, statistical significance was achieved in only the largest study. ⊕⊕◯◯ LOW
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	80% vs. 78%‡ Log-rank, p=0.10	
	2 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	70% vs. 50%‡ Log-rank, p=0.01	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	66% vs. 49%‡ Log-rank, p=0.10	
	3 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	55% vs. 39%‡ Log-rank, p=0.01	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	48% vs. 38%‡ Log-rank, p=0.10	
	4 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	44% vs. 35%‡ Log-rank, p=0.01	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	42% vs. 30%‡ Log-rank, p=0.10	

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
	5 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	41.6% vs. 31.6%; adj. HR 1.45 (1.09 to 1.94) p=0.010 *Stage III only*: 34.6% vs. 25.0%, p=0.04	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	42% vs. 19%; adj. HR 1.48 (0.93 to 2.35), p=0.10 *All patients*: HR 0.82 (0.56 to 1.20), p=0.30	
Mortality, % (n/N)	3 months	Lin 2017 (N=580) Retro cohort AC (92%) or SCC (8%) Stage III/IV (63%)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	1 mo. post-op: 0% vs. 1.5% (7/469), p=0.425 2 mos. post-op: 0.9% (1/111) vs. 2.6% (12/469), p=0.59 3 mos. post-op: 0.9% (1/111) vs. 4.3% (20/469), p=0.26	No statistically differences; per authors, the difference at 3 months may be clinically meaningful. ⊕⊕◯◯ LOW
	Median 22 months	Makishima 2015 N=44 SCC (100%) Stage III (52%); Stage I/II (48%)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	20% (5/25) vs. 31.6% (6/19), p=NR	No statistically significant differences; sample sizes are small. ⊕◯◯◯ INSUFFICIENT
Probability, Progression-free survival (PFS) or Disease-free survival (DFS)	1 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	PFS: 62% vs. 50%, p=0.001	At all timepoints, PFS/DFS was better following PBT vs. IMRT or 3DCRT across both studies. However, statistical significance was achieved in only the largest study. ⊕⊕◯◯ LOW
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	DFS: 55% vs. 45%, p=0.11	

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
	2 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	PFS: 50% vs. 33%, p=0.001	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	DFS: 45% vs. 26%, p=0.11	
	3 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	PFS: 42% vs. 28%, p=0.001	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	DFS: 41% vs. 23%, p=0.11	
	4 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	PFS: 39% vs. 24%, p=0.001	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	DFS: 41% vs. 23%, p=0.11	
	5 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	PFS: 34.9% vs. 20.4%; adj. HR 1.56 (95% CI 1.19-2.05), p=0.001 *Stage III*: 33.5% vs. 13.2%, p=0.005	

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	DFS: 41% vs. 18%, adj. HR 1.42 (0.92 to 2.19) p=0.11	

adj. = adjusted; 3D-CRT = 3-dimensional conformal radiation therapy; AC = adenocarcinoma; CI = confidence interval; HR = hazard ratio; KQ = Key Question; PBT = proton beam therapy; IMRT = intensity-modulated radiation therapy; NS = not statistically significant; Retro = retrospective study design; SCC = squamous cell carcinoma; XRT = X-ray radiation therapy.
***Fang 2018**: PBT vs. IMRT
Lin 2017: PBT vs. IMRT and vs. 3D-CRT
Makishima 2015: passive scatter PBT vs. XRT
Shiraishi 2018: PBT vs. IMRT
Xi 2017: PBT vs. IMRT
†If no 95% CI is provided in the table, the authors did not report one; log-rank p-values.
‡Estimated from graphs in articles.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Esophageal Cancer for Safety

Outcome	Time	Studies, Year, N, RoB Tumor Indication	Reason for Downgrading	PBT vs. various photon Effect estimate (95% CI)	Conclusion Quality (SoE)
RT-related toxicities					
Radiation pneumonitis, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	None	PBT vs. IMRT: 1.5% (2/132) vs. 2.8% (6/211), p=NS	For PBT versus IMRT, with the exception of grade 4 radiation-induced lymphopenia (2 studies) and wound events (1 study) which were less common with PBT, the frequency of all other RT-related and treatment-related toxicities and adverse events did not differ statistically between groups. For PBT vs. 3DCRT or XRT, with the exception of GI events, PBT was associated with a statistically lower frequency of any treatment-related toxicity (i.e., pulmonary, cardiac, and wound events; grades ≥2 or not specified) across three studies. There were no differences in the frequency of grade ≥3 radiation pneumonitis and pleural effusion between PBT vs. XRT in one small study. ⊕⊕◯◯ LOW
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	PBT vs. XRT: 0% (0/25) vs. 5.3% (1/19), p=NS	
Radiation esophagitis, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%)	None	PBT vs. IMRT: 11.4% (15/132) vs. 14.2% (30/211), p=NS	
Radiation induced lymphopenia, grade 4	Acute (during RT; timing NOS)	Fang 2018 (N=220) Retro propensity-score matched cohort AC (74%) or SCC (26%)	**Imprecision** Yes[3] (-1)	PBT vs. IMRT: 31% (34/110) vs. 47% (52/110); adj. OR 0.47 (0.26 to 0.84) p=0.01	
	Acute (during RT; timing NOS)	Shiraishi 2018 (N=272) Retro propensity-score matched cohort AC (97%) or SCC (3%)	**Imprecision** Yes[3] (-1)	PBT vs. IMRT: 17.6% (24/136) vs. 40.4% (55/136); adj. OR 0.29 (0.16 to 0.52) p<0.0001	
Treatment-related toxicity*					
Esophageal fistula, Esophageal stricture, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%)	None	PBT vs. IMRT: Esophageal fistula: 0% (0/132) vs. 1.4% (3/211) Grade 5: 0% (0/132) vs. 0.5% (1/211) Esophageal stricture: 9.8% (13/132) vs. 8.1% (17/211) Grade 5: 0% (0/132) vs. 0.5% (1/211) p=NS for all	

Outcome	Time	Studies, Year, N, RoB Tumor Indication	Reason for Downgrading	PBT vs. various photon Effect estimate (95% CI)	Conclusion Quality (SoE)
Any pulmonary event	Acute (post-op)†	Lin (2017) (N=580) Retro cohort AC (92%) or SCC (8%)	**Imprecision** Yes[3] (-1)	*Grade NR* **PBT**: 16.2% (18/111) **IMRT**: 24.2% (62/255) **3DCRT**: 39.5% (85/214) PBT *vs. IMRT*: adj. OR 0.58 (95% CI 0.32 to 1.05), p=0.08; PBT *vs. 3D-CRT*: adj. OR 0.34 (95% CI 0.19 to 0.61), p<0.001	
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	PBT vs. XRT: Grade ≥2: 0% (0/25) vs. 42.1% (8/19), p<0.001	
Pleural effusion, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%)	None	PBT vs. IMRT: 0.8% (1/132) vs. 1.9% (4/211), p=0.19	
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	PBT vs. XRT: 0% (0/25) vs. 5.3% (1/19), p=NS	
Any cardiac event	Acute (post-op)†	Lin (2017) (N=580) Retro cohort AC (92%) or SCC (8%)	**Imprecision** Yes[3] (-1)	*Grade NR* PBT: 11.7% (13/111) IMRT: 11.7% (30/255) 3DCRT: 27.4% (59/214) PBT *vs. IMRT*: adj. OR 0.87 (95% CI 0.42 to 1.77), p=0.70; PBT *vs. 3D-CRT*: adj. OR 0.34 (95% CI 0.17 to 0.66), p=0.002	
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	PBT vs. XRT: Grade ≥2: 4% (1/25) vs. 52.6% (10/19), p<0.001 RR 0.08 (0.01 to 0.54)‡	
Pericardial effusion, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%)	None	PBT vs. IMRT: 0.8% (1/132) vs. 2.4% (5/211), p=0.19	

Outcome	Time	Studies, Year, N, RoB Tumor Indication	Reason for Downgrading	PBT vs. various photon Effect estimate (95% CI)	Conclusion Quality (SoE)
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	PBT vs. XRT: 0% (0/25) vs. 0% (0/19), p=NS	
Any GI event, any wound event	Acute (post-op)†	Lin (2017) (N=580) Retro cohort AC (92%) or SCC (8%)	**Imprecision** Yes[3] (-1)	*Grade NR* GI event PBT: 18.9% (21/111) IMRT: 23.0% (59/255) 3DCRT: 20.9% (45/214) Chi Squared p-value: p=0.656 Wound event PBT: 4.5% (5/111) IMRT: 14.1% (36/255) 3DCRT: 15.3% (33/214) PBT *vs. IMRT*: adj. OR 0.28 (95% CI 0.11 to 0.73), p=0.009 PBT vs. *3D-CRT*: OR 0.26 (95% CI 0.10 to 0.68), p=0.006	
Readmission within 60 days or death during same hospitalization	2 mos.†	Lin (2017) (N=580) Retro cohort AC (92%) or SCC (8%)	**Imprecision** Yes[3] (-1)	PBT: 17.1% (19/111) IMRT: 15.6% (40/255) 3DCRT: 23.7% (51/214) Chi Squared p-value: p=0.070	

3D-CRT: 3-dimensional conformal radiation therapy; AC: adenocarcinoma; CI: confidence interval; PBT: proton beam therapy; IMRT: intensity-modulated radiation therapy (photons); NOS: not otherwise specified; NS: not statistically significant; OR: odds ratio; post-op: post-operative; Retro: retrospective study design; RR: risk ratio; SCC: squamous cell carcinoma; XRT: X-ray radiation therapy.

*Not directly stated by authors as related to RT – called "treatment-related"; because all patients were receiving concurrent or adjuvant chemotherapy is it unclear the degree to which PBT directly affected these outcomes.

†All patients in the study were treated with neoadjuvant concurrent chemotherapy and radiation therapy followed by surgical resection (most commonly esophagectomy 84%); follow-up period post-op is unclear though appears to be up to 3 months. Postoperative complications were identified from hospital notes, discharge summary, and/or from a prospectively collected surgical database.

‡Crude RR calculated by AAI.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of

overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).

3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Gastrointestinal (Pancreas) Cancer for Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT (spot scanning) vs. HART Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)					No statistically significant differences were seen between PBT and HART for any primary outcome (OS, disease control, local progression, and metastasis) or for any acute RT-related toxicity (hematological and non-hematological); clinical importance of differences is unclear. The sample size was very small. ⊕◯◯◯ INSUFFICIENT
Probability, overall survival (OS)	1-3 years	Maemura 2017 (N=25) Retro cohort Adenocarcinoma (locally advanced and unresectable)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	1-year: 80% vs. 86.7% 2-year: 45% vs. 33.3% 3-year: 22.5% vs. 26.6% p=NS at all timepoints	
Disease control, % (n/N)	NR			80% (8/10) vs 93% (14/15), p=NR; RR 0.86 (0.61 to 1.20)*	
Local progression, % (n/N)	NR			40% (4/10) vs 60% (9/15), p=NR; RR 0.60 (0.26 to 1.39)*	
Metastasis, % (n/N)	NR			Any: 30% (3/10) vs. 20% (3/15) • Lung: 10% (1/10) vs 0% (0/15) • Liver: 30% (3/10) vs 6.7% (1/15) • Peritoneum: 10% (1/10) vs 13.3% (2/15) p=NR	
Safety (KQ3) (Curative intent only)					
Acute Toxicity (≤3 mos.)	NR	Maemura 2017 (N=25) Retro cohort Adenocarcinoma (locally advanced and unresectable)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	RT-related Toxicities, % (n/N) <u>Hematological</u> Leukopenia • Grade 2: 10% (1/10) vs. 13% (2/15) • Grade 3: 0% (0/10) vs. 20% (3/15) Thrombocytopenia: • Grade 2: 10% (1/10) vs. 20% (3/15) • Grade 3: 0% (0/10) vs. 6.7% (1/15) Neutropenia; Anemia: • Grade 2 or 3: 0% (0/10) vs. 0% (0/15)	

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT (spot scanning) vs. HART Effect estimate (95% CI)	Conclusion Quality (SoE)
				Non-hematological Ulcer: • Grade 2: 10% (1/10) vs 0% (0/15) • Grade 3: 10% (1/10) vs 0% (0/15) Nausea: • Grade 2: 0% (0/10) vs. 7% (1/15) • Grade 3: 0% (0/10) vs. 0% (0/15) Anorexia: • Grade 2: 0% (0/10) vs. 20% (3/15) • Grade 3: 0% (0/10) vs. 0% (0/15) Malaise • Grade 2 or 3: 0% (0/10) vs. 0% (0/15) *No grade 4 toxicities occurred in either group*	

CI = confidence interval; HART = Hyper-fractionated accelerated RT; KQ = Key Question; NR = not reported; PBT = proton beam therapy; Retro = retrospective study design; RR = risk ratio; SOE = strength of evidence.
*Crude RR calculated by AAI.

Reasons for downgrade:
1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Head and Neck Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)					
Survival outcomes					
Probability, overall survival (OS)	1-year	Romesser 2016 (N=41) Retro cohort Salivary gland cancer (primary or metastasis)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	83.3% vs. 93.3%, p=0.08	Regardless of tumor types, no statistically significant differences were seen between PBT and IMRT in the probability of 1-3 year OS (2 studies) or 3-year PFS (1 study) or in the incidence of all-cause mortality (1 study). Clinical significance of differences is unclear. ⊕⊕◯◯ LOW for primary oropharyngeal and nasopharyngeal cancer ⊕◯◯◯ INSUFFICIENT for salivary cancer (primary or metastatic)
	3-years	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	94.3% vs. 89.3%; adj. HR 0.55 (95% CI 0.1 to 2.5), p=0.44	
Probability, progression free survival (PFS)	3-years	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	86.4% vs. 85.8%; adj. HR 1.0 (95% CI 0.4 to 2.6), p=0.99	
All-cause mortality, % (n/N)	Median 24 mos.	Holliday 2015 (N=30) Retro case-matched cohort Nasopharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	10% (1/10) vs. 5% (1/20), p=NS	
Safety (KQ3) (Curative intent only)					
Toxicities and other adverse events					
Acute toxicity grade ≥3	≤3 mos.	Romesser 2016 (N=41) Retro cohort Salivary gland cancer (primary or metastasis)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	Dermatitis: 27.8% (5/18) vs. 34.8% (8/23) Mucositis: 0% (0/18) vs. 8.7% (2/23) Nausea, Dysphagia, Dysgeusia, Fatigue: no	There were no statistically significant differences in the frequency of grade ≥3 acute or late toxicities following PBT versus IMRT across three studies. Clinical significance of differences is unclear. Sample size and residual confounding and/or tumor

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
				events in either group p=NS for all no *grade 4* events in either group	type and stage may have played a role in some of these findings. ⊕⊕◯◯ LOW for acute (based on highest quality studies) and late toxicity
		Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	Dermatitis: Data NR, p=0.15 Mucositis: Data NR, p=0.90 Weight loss (>20% vs. baseline): 8.3% (4/48) vs. 13.5% (13/98); adj OR 0.64 (95 CI 0.19 to 2.11) Fatigue (grade 2 or 3): 40.8% (20/49) vs. 36.2% (34/94); adj OR 1.1 (95% CI 0.53 to 2.27) Xerostomia (grade 2 or 3): 42% (21/50) vs. 61.2% (60/98); adj OR 0.38 (95% CI 0.18 to 0.79)	
		Holliday 2015 (N=30) Retro case-matched cohort Nasopharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	Any Grade 3 event: 50% (5/10) [9 events] vs. 90% (18/20) [30 events]; RR 0.56 (95% CI 0.29 to 1.05)† Dermatitis (Grade 3): 40% (4/10) vs. 25% (5/20); RR 1.6 (0.55 to 4.68)† Any Grade 4/5 events: 0% vs. 0% Swallowing dysfunction: 0%	

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
				(0/10) vs. 15% (3/20), p=0.175 Mean percentage (IQR) body weight lost from pre to post RT: 5.7% (4.5% to 11.2%) vs. 7.6% (6.1% to 12.1%), p=0.333	
Late toxicity grade ≥3	1 year	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	Weight loss (>20% vs. baseline): 6.7% (3/45) vs. 19.3% (17/88); adj OR 0.28 (95 CI 0.08 to 1.05) Fatigue (grade 2 or 3): 14.6% (7/48) vs. 22.1% (17/77); adj OR 0.5 (95% CI 0.18 to 1.36) Xerostomia (grade 2 or 3): 42% (21/50) vs. 47.2% (42/89); adj OR 0.63 (95% CI 0.30 to 1.33)	
	NR (median 24 mos.)	Holliday (N=30) Retro case-matched cohort Nasopharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	Any Grade 3 event: 30% (3/10) [5 events] vs. 15% (3/20) [3 events]; RR 2.0 (95% CI 0.49 to 8.18)†	
Gastrostomy tube dependence	Acute	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	≤3 months: 12% (6/50) vs. 23% (23/100); adj OR 0.43 (95% CI 0.16 to 1.17)	GT dependence tended to be lower with PBT, however adjusted estimates from the largest study were not statistically significant, while smaller studies in different cancer

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
		Holliday (N=30) Retro case-matched cohort Nasopharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	During or after RT: 20% (2/10) vs. 65% (13/20), p=0.02; adj. OR 9.33 (95% CI 1.74 to 75.96), p=0.008	types reported statistically significant differences. For the smallest study, the large confidence interval suggest instability of the effect estimate. Clinical significance of differences is unclear. It is unclear what role differences in study populations (including tumor characteristics, etc.) and possible residual confounding may play in these findings. ⊕⊕◯◯ LOW
		McDonald 2016 (N=40) Retro comparative cohort Nasopharynx, nasal cavity or paranasal sinus cancers (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	End of RT: adj. OR 0.03 (95 % CI <0.01 to 0.15), p<0.001 1 month post-RT: adj. OR 0.11 (95% CI <0.01 to 0.61), p=0.028	
		Romesser 2016 (N=41) Retro cohort Salivary gland cancer (primary or metastasis)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	≤3 months: 0% vs. 0% (reactive gastrostomy tube or tracheostomy)	
	Late	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	1 year: 2% (1/50) vs. 7.8% (7/90); adj OR 0.16 (95% CI 0.02 to 1.37)	
		Sharma 2018 (N=64) Prospective cohort Oropharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	6 months: 0% vs. 0%	
ED visit or hospital-ization	During RT	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	ED visit: 32%(16/50) vs. 32% (32/100); adj. OR 0.95 (95% CI 0.45 to 2.0) Unscheduled hospitalization: 20% (10/50) vs. 21% (21/100);	No statistically significant differences in the frequency of ED visits or unplanned hospitalizations following PBT versus IMRT. ⊕⊕◯◯ LOW

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
				adj OR 0.92 (95% CI 0.39 to 2.2)	
Osteoradio-necrosis	Median 34 mos.	Zhang 2017 (N=584) Retro cohort Oropharyngeal cancer (primary)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	Any grade: 2% (1/50) vs. 7.7% (41/534); RR 0.26 (0.04 to 1.85)† Grade 3: 0% (0/50) vs. 0.9% (5/534) Grade 4: 0% (0/50) vs. 2.2% (12/534) Grade 3 or 4: 0% (50) vs. 3.2% (17/534) p=NS for all	No statistically significant differences in the frequency of osteoradionecrosis following PBT versus IMRT. The small number of patients for PBT may preclude identification of rare events and residual confounding may have played role in some of these findings. ⊕◯◯◯ INSUFFICIENT

adj. = adjusted; CI = confidence interval; ED = emergency department; HR = hazard ratio; KQ = Key Question; OR = odds ratio; PBT = proton beam therapy; IMRT = intensity-modulated radiation therapy; NS = not statistically significant; Retro = retrospective study design; RT = radiation therapy.

* **Blanchard 2016:** intensity modulated spot-scanning PBT vs. IMRT
Holliday 2015: intensity modulated spot-scanning PBT vs. IMRT
McDonald 2016: 3D conformal PBT vs.IMRT
Romesser 2016: Uniform scanning-beam PBT vs. IMRT
Sharma 2018: Adjuvant pencil beam scanning PBT vs. IMRT via volumetric modulated arc therapy (VMAT) following transoral robotic surgery and selective neck dissection
Zhang 2017: intensity modulated spot-scanning PBT vs. IMRT

†Crude RR calculated by AAI. The small number of patients for PBT may preclude identification of rare events.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Skull-base Head and Neck Cancer for Effectiveness

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	Surgery + adjuvant PBT vs. Surgery alone RR (95% CI)*	Conclusion Quality (SoE)
Curative intent (KQ1)					
Survival and tumor control outcomes					
Probability, disease-specific survival (DSS)	5-, 10-years	Simon 2018 N=47 (n=34 petroclival only) Retro comparative cohort Chondro-sarcoma (grade II)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	*5- and 10-year DSS in:* **All patients** 100% vs. 89.8% (76.2% to 100%), p=0.138 **Petroclival patients only** 100% vs. 76.4% (46.1% to 100%), p=0.028	The probability of PFS, but not DSS, at 5 and 10 years was statistically better following surgery with adjuvant PBT versus surgery alone. PBT resulted in improved DSS and PFS at both time points for the subgroup of patients with petroclival tumors. Local control was statistically better following adjuvant PBT. ⊕◯◯◯ INSUFFICIENT
Probability, progression-free survival (PFS)	5-, 10-years			**All patients** *5-year*: 100% vs. 67.8% (47.7% to 88.0%) *10-year*: 87.5% (64.6% to 100%) vs. 58.2% (33.5% to 82.8%) p=0.006 **Petroclival patients only** *5-year*: 100% vs. 50% (15.4% to 84.6%) *10-year*: 85.7% (59.8% to 100%) vs. 50.0% (15.4% to 84.6%) p=0.001	
Proportion of patients experiencing local relapse, or regional or distant metastases% (n/N)	Median 7.5 years			**Local relapse:** 4.3% (1/23) vs. 33% (8/24); RR 0.13, 95% CI 0.02 to 0.96, p=0.01 (5/9 patients went on to receive secondary proton therapy) **Regional or distant metastases:** 0% vs. 0%	
Safety					
Any complication, % (n/N)	Median 7.5 years	Simon 2018 N=28 for PBT and 47 for surgery† Retro comparative cohort Chondro-sarcoma (grade II)	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	68% (19/28) vs. 26% (12/47), RR 2.7 (1.5 to 4.6)	Unadjusted estimates of treatment-related death and severe complications (grade ≥3 toxicity) did not differ statistically between groups, however, patients who received adjuvant PBT had a higher risk of experiencing any
Any grade ≥3 toxicity, % (n/N)				25% (7/28) vs. 11% (5/47), p=0.10	
Treatment-related death, % (n/N)				0% (0/28) vs. 2% (1/47), p=0.44	
Hearing loss and dizziness, % (n/N)				**Sensorineural hearing loss:** 39% (11/28) vs. 6%	

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	Surgery + adjuvant PBT vs. Surgery alone RR (95% CI)*	Conclusion Quality (SoE)
				(3/47), RR 6.2 (1.9 to 20.2) **Severe hearing loss:** 21% (6/28) vs. 4% (2/47), RR 5.0 (1.1 to 23.3) **Conductive hearing loss:** 11% (3/28) vs. 4% (2/47), p=0.28 **Dizziness**: 14% (4/28) vs. 0% (0/47), p=0.008	complication, specifically sensorineural and severe hearing loss. However, confidence intervals were wide suggesting instability of the effect estimate. ⊕◯◯◯ INSUFFICIENT
Other complications from PBT, % (n/N)				**Vision loss:** 11% (3/28) **Hypopituitarism**: 18% (5/28) **Temporal lobe necrosis**: 18% (5/28)	

CI = confidence interval; KQ = Key Question; NCDB = National Cancer Data Base; PBT = proton beam therapy; Retro = retrospective; SOE = strength of evidence.
*Crude RRs and 95% CIs were calculated by AAI.
†All patients were included in evaluation of complications due to surgery and 28 total patients were included in the evaluation of complications due to PBT (23 primary treatment and 5 secondary PBT treatment follow-up local relapse).

Reasons for downgrade:
1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Liver Tumors for Efficacy and Safety

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT (passive scatter) vs. TACE (RCT) or vs. IMRT (Observational study) Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)					
Randomized controlled trial					
Probability, overall survival (OS)	2-year	Bush 2016 (N=69) RCT Moderately low RoB HCC	**Consistency** Unknown **Imprecision** Yes[3] (-1)	All patients: 59% (NR) Patients receiving liver transplant post-treatment (n=22): 82% (NR) p=NS for both, data not provided	No significant difference between groups in the probability of 2-year OS; patients who received PBT tended to have improved probability of 2-year PFS and local tumor control compared with TACE patients, although the difference did not reach statistical significance. Results are from interim analysis of an ongoing trial. ⊕⊕⊕◯ MODERATE
Probability, progression free survival (PFS)	2-year			48% (NR) vs. 31% (NR); p=0.06	
Probability, local control (LC)	2-year			88% (NR) vs. 45% (NR); p=0.06	
Observational study					
Probability, overall survival (OS)	2-year	Sanford 2019 (N=133) Retrospective cohort study Moderately high RoB HCC	**Inconsistency** Unknown **Imprecision** Yes[3] (-1)	59.1% vs. 28.6%; adj. HR 0.47 (95% CI 0.27 to 0.82)	OS was significantly higher following PBT vs. IMRT but there was no difference in local and regional control between groups. ⊕⊕◯◯ LOW
Probability, local and locoregional control	2-year			**Local control** (cumulative incidence): 93% (NR) vs. 90% (NR); HR for cumulative incidence of *local failure* 0.74 (95% CI 0.18 to 3.01) **Locoregional recurrence** (cumulative incidence): 53% vs. 42%; adjusted HR 0.98 (95% CI 0.54 to 1.75).	

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT (passive scatter) vs. TACE (RCT) or vs. IMRT (Observational study) Effect estimate (95% CI)	Conclusion Quality (SoE)
Safety (KQ3) (Curative intent only)					
Randomized controlled trial					
Acute Toxicity (≤3 mos.)	NR	Bush 2016 (N=69) RCT Moderately low RoB HCC	**Consistency** Unknown **Imprecision** Yes[3] (-1)	Acute toxicity was generally limited to the following, which were experience by most patients (no data provided): PBT: fatigue and radiation skin reaction TACE: abdominal pain and nausea Authors state that serious complications from PBT were uncommon events (no data provided).	Limited information provided on acute toxicity. Significantly fewer patients who received PBT required hospitalization in the month following treatment compared with TACE patients; total days hospitalized was also significantly less in the PBT vs. the TACE group. Results are from interim analysis of an ongoing trial. ⊕⊕⊕◯ MODERATE
Proportion of patients hospitalized for an acute complication, % (n/N)	≤1 mo.			6.1% (2/33) vs. 41.7% (15/36); p<0.001	
Total days hospitalized within 1 month of treatment	≤1 mo.			**Overall**: 24 (0.73 days per patient) vs. 166 (4.6 days per patient); p<0.001 *for routine observation*: 0 vs. 53 *for complications:* 24 vs. 113	
Observational study					
Incidence of nonclassic radiation-induced liver disease (RILD)*	3 mos.	Sandford 2019 (N=100)† Retrospective cohort study Moderately high RoB HCC	**Consistency** Unknown **Imprecision** Yes[3] (-1)	adj. OR 0.26 (95% CI 0.08 to 0.86) (PBT, n=4 patients; IMRT, n=17 patients) Authors also report that the development of RILD at 3 months was associated with significantly worse OS (HR 3.83; 95% CI 2.12 to 6.92).	Lower risk of RILD in the acute period with PBT versus IMRT ⊕⊕◯◯ LOW

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT (passive scatter) vs. TACE (RCT) or vs. IMRT (Observational study) Effect estimate (95% CI)	Conclusion Quality (SoE)
Death due to liver failure	NR (median f/u 14 mos.)	Sandford 2019 (N=36)‡ Retrospective cohort study Moderately high RoB HCC	**Consistency** Unknown **Imprecision** Yes[3] (-1)	53% (8/15) vs. 91% (19/21); RR 0.59 (95% CI 0.36 to 0.97)§	Lower risk of death due to liver failure with PBT versus IMRT; however data was from a small subset of patients. ⊕◯◯◯ INSUFFICIENT

HCC = hepatocellular carcinoma; IMRT = intensity-modulated radiation therapy (photons); NR = not reported; PBT = proton beam therapy; RCT = randomized controlled trial; RoB = risk of bias; SOE = strength of evidence; TACE = Transarterial chemoembolization

*RILD was defined as worsening of Child-Pugh score by ≥2 points compared with baseline. At baseline, patients treated with photons had worse baseline child-Pugh score (median 6 vs. 5, p=0.008), however, this variable was included in and controlled for via multivariate analyses.

†RILD was calculated in 100 (of 133) patients for whom data was available; denominators for this subset of patients by treatment group were not provided.

‡Death due to liver failure was reported only among the 36 patients (15 PBT, 21 IMRT) without disease progression.

§RR and 95% CI calculated by AAI.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Lung Cancer for Efficacy/Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
Curative intent (KQ1)					
Randomized controlled trials					
Probability, overall survival (OS)‡	1-5 year	Liao 2018 N=173 (ITT) RCT NSCLC	**Consistency** Unclear **Imprecision** Yes[3] (-1)	1-year: 75% vs. 82% 2-year: 56% vs. 60% 3-year: 26% vs. 37% 4-year: 38% vs. 32% 5-year: 24% vs. 32% p=0.30	No statistically significant differences between groups in the probability of OS or the cumulative incidence of local failure at any timepoint measured. ⊕⊕⊕◯ MODERATE
Cumulative incidence of local failure (%)‡				1-year: 9% vs. 10% 2-year: 27% vs. 26% 3-year: 37% vs. 37% 4-year: 37% vs. 32% 5-year: 37% vs. 39% p=0.99	
Observational studies					
Probability, overall survival (OS)	1-year	Liao 2018§ N=39 Pro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	69% vs. 57% p=0.97	No statistically significant differences between groups in the probability of OS over 1-5 years (across 4 studies) or LRFS at 1 or 2 years (1 study) or in the incidence of local failure at 2 or 3 years (2 studies) ⊕⊕◯◯ LOW for OS ⊕◯◯◯ INSUFFICIENT for LRFS and local failure
		Remick 2017 N=61 Retro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	85.2% (72.8%–99.7%) vs. 82.4% (70.5%–96.2%) p=0.65	
		Higgins 2017 N=1850 (propensity-matched) Retro database NSCLC	**Imprecision** Yes[3] (-1)	62.0% (56.2%–67.2%) vs. 54.2% (51.6%–56.7%) p=NR	
	2-year	Liao 2018§ N=39 Pro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	43% (NR) vs. 43% (NR) p=0.97	
		Remick 2017 N=61 Retro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	77.8% (63.6%–95.2%) vs. 73.2% (59.6%–89.9%) p=0.65	
		Tucker 2016 N=468 Retro cohort NSCLC	**Imprecision** Yes[3] (-1)	PBT: 56% (40%–69%) IMRT: 52% (45%–58%) 3DCRT: 39% (32%–46%) p=NS, PBT vs. IMRT p=0.015, PBT vs. 3DCRT	

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
	3-year	Liao 2018§ N=39 Pro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	25% (NR) vs. 32.5% (NR) Log-rank p=0.97	
	5-year	Higgins 2017 N=1850 (propensity-matched) Retro database NSCLC	**Imprecision** Yes[3] (-1)	5:1 matching: 22.3% (16.3%–28.9%) vs. 15.7% (13.5%–18.1%) adj. HR 1.18 (95% CI 1.02 to 1.37) *a-priori* 1:1 matching: adj. HR 1.16 (95% CI 0.97 to 1.39)	
Probability, Local Recurrence-Free Survival (LRFS)	1-2 year	Remick 2017 N=61 Retro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Consistency** Unclear **Imprecision** Yes[3] (-1)	1-year: 92.3% (82.5%–100%) vs. 93.3% (84.8%–100%) 2-year: 93.1% vs. 85.7% p=0.82	
Local Failure	1-2 years	Liao 2018§ N=39 Pro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Consistency** Unclear **Imprecision** Yes[3] (-1)	Cumulative incidence‡: 1-year: 6% vs. 3% 2-year: 6% vs. 3% 3-year: 26% vs. 26% p=0.93	
	2-years	Remick 2017 N=61 Retro cohort NSCLC		11.1% (3/27) vs. 5.9% (2/34), p=NS	
Safety (KQ3) (all curative intent)					
Randomized controlled trials					
Rate of radiation pneumonitis, Grade ≥3‡	1-5 years	Liao 2018 N=173 (ITT) RCT NSCLC	**Consistency** Unclear **Imprecision** Yes[3] (-1)	8% vs. 7% at 1, 2, 3, 4 and 5 years; p=0.58	No statistically significant differences between groups. ⊕⊕⊕◯ MODERATE
Observational studies					
Radiation esophagitis	NR (median 26 months)	Remick 2017 N=61 Retro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	Grade 2: 18.5% (5/27) vs. 29.4% (10/34), p=NR Grade 3: 3.7% (1/27) vs. 11.8% (4/34), p=NR	No statistically significant differences between groups for any grade 3 outcome; however
	NR	Niedzielski 2017		Grade 2: 59.2% (29/49) vs. 54.1% (46/85), p=NS	

Outcome	Time	Studies, Year, N, Tumor	Reason for Downgrading	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
		N=134 Retro cohort NSCLC		Grade 3: 22.4% (11/49) vs. 17.6% (15/85); OR 1.4 (0.7 to 2.9), p=0.37	differences may be clinically important. ⊕◯◯◯ INSUFFICIENT
Radiation pneumonitis	NR (median 26 months)	Remick 2017 N=61 Retro cohort NSCLC	**Risk of Bias** Yes[1] (-1) **Imprecision** Yes[3] (-1)	Grade 2: 3.7% (1/27) vs. 8.8% (3/34), p=NR Grade 3: 3.7% (1/27) vs. 2.9% (1/34), p=NR	
Radiation dermatitis				Grade 2: 37% (10/27) vs. 12% (4/34), p=NR Grade 3: 0% (0/27) vs. 0% (0/34), p=NR	

3D-CRT = Three-dimension conformal radiation therapy; adj. = adjusted; CI = confidence interval; HR = hazard ratio; IMRT = intensity-modulated radiation therapy; ITT = intention-to-treat analysis; KQ = Key Question; NR = not reported; NS = not statistically significant; NSCLC = non-small cell lung cancer; RCT = randomized controlled trial; Retro = retrospective study design; Pro = prospective study design.

***Liao 2018** (RCT and observational): passive scatter PBT vs. IMRT

Higgins 2017: PBT vs. various photon (external beam, 3D-conformal, IMRT, "photons")

Niedzielski 2017: passively scattered PBT vs. IMRT

Remick 2017: double scatter or pencil beam PBT vs. IMRT

Tucker 2016: pencil beam PBT vs. IMRT vs. 3DCRT

†If no 95% CI is provided in the table, the authors did not report one; log-rank p-values.

‡Estimated from figures/graphs in article.

§This cohort is comprised of patients from the RCT who could not be randomized because their PBT or IMRT plans did not allow for random assignment (i.e., did not meet prespecified dose-volume constraints developed for photon radiation); they were followed as an observational cohort.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Ocular Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT vs. Brachytherapy or Stereotactic Radiosurgery* Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)					
Survival and tumor recurrence outcomes					
Probability of overall survival (OS)	2, 5, years	Lin 2017 (N=452) Retro propensity-score matched cohort (NCD) Choroid melanoma	**Consistency** Unknown **Imprecision** Yes[3] (-1)	2-year OS: 93% vs. 97%, p=NS 5-year OS: 51% vs. 77% adj. HR for *risk of mortality*: 1.89, 95% CI 1.24 to 2.95	Similar OS/mortality at 2 and 3 years for PBT vs. brachytherapy or SRS in 2 studies of choroid and uveal melanoma. In the larger database study of choroid melanoma, PBT was associated with a statistically higher risk of mortality at 5 years vs. brachytherapy. ⊕⊕◯◯ LOW
Mortality, % (n/N)	3 years	Sikuade 2015 (N=191) Retro cohort Uveal Melanoma	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	13% (14/106) vs. 16% (14/85), p=NS	
Local recurrence	3, 5, 10 years	Böker (2018), N=140 Retro case-matched cohort Large Uveal Melanoma	**Consistency** Unknown **Imprecision** Yes[3] (-1)	Rate (95% CI) 3-years: 4% (1.2% to 17.8%) vs. 24.6% (15.8% to 37.1%), p<0.001 5-years: 9.1% (2.9% to 27.3%) vs. 27.5% (17.8% to 41.1%), p<0.001 10-years: 9.1% (2.8% to 27.3%) vs. 36.5% (20.7% to 59.1%); adj. HR 7.69 (95% CI 2.22 to 26.06) *for brachytherapy*	PBT was associated with a statistically lower frequency of local recurrence over 10 years compared with brachytherapy (+TSR for both). ⊕⊕◯◯ LOW
	Mean 3 years	Sikuade 2015 (N=191) Retro cohort Uveal Melanoma	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	2.8% (3/106) vs. 0% (0/85), p=NS	No statistical difference in local recurrence between PBT versus SRS ⊕◯◯◯ INSUFFICIENT
Safety (KQ3) (Curative intent only)					
	Mean 3.3 years	Böker (2018), N=140 Retro case-matched cohort Large Uveal Melanoma	**Consistency** Unknown **Imprecision** Yes[3] (-1)	Enucleation: 8.5% (6 eyes) vs. 15.7% (11 eyes), p=0.196 Rubeosis of the iris: 1.4% (1/70) vs. 0% (0/70), p=0.316 Neovascular glaucoma: 1.4% (1/70) vs. 1.4% (1/70), p=NS	With the exception of optic neuropathy which was statistically lower following PBT versus SRS in one study of uveal melanoma, no other statistical

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT vs. Brachytherapy or Stereotactic Radiosurgery* Effect estimate (95% CI)	Conclusion Quality (SoE)
	Mean 3 years	Sikuade 2015 (N=191) Retro cohort Uveal Melanoma	**Risk of Bias** Yes[1] (-1) **Consistency** Unknown **Imprecision** Yes[3] (-1)	Enucleation: 1.9% (2/106) vs. 2.4% (2/85), p=NS Rubeotic glaucoma: 4.7% (5/106)† vs. 11% (9/85)†, p=NS Radiation retinopathy: 30% (31/106) vs. 24% (20/85), p=NS Optic Neuropathy: 13% (14/106) vs. 28% (23/85); RR=0.49 (0.27 to 0.89)‡	differences were seen in the frequency of adverse events over 3 years between PBT versus brachytherapy or SRS. ⊕⊕◯◯ LOW

adj. = adjusted; CI = confidence interval; HR = hazard ratio; KQ = Key Question; NCD = National Cancer Database; NS = not statistically significant; PBT = proton beam therapy; Retro = retrospective study design; RR = risk ratio; SRS = stereotactic radiosurgery; TSR = transscleral resection.

*__Boker 2018__: Neoadjuvant PBT + TSR vs. Adjuvant Brachytherapy + TSR

__Lin 2017__: PBT vs. Brachytherapy

__Sikuade 2015__: PBT vs. SRS

†*Requiring enucleation:* 1.9% (2/106) [40% (2/5) with rubeotic glaucoma] vs. 2.4% (2/85) [22% (2/9) with rubeotic glaucoma].

‡Calculated by AAI.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Strength of Evidence Summary for Adult Prostate Cancer for Effectiveness and Safety

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)					
Survival outcomes – *quasi-RCT*					
Probability, overall survival (OS)	5-year	Khmelevsky 2018 quasi-RCT (N=289) Moderately high RoB Prostate Cancer Risk: High (53%), Intermediate (42%), Low (5%)	**Consistency** Unclear **Imprecision** Yes[3] (-1)	74% ± 5.0% vs. 78.8% ± 4.1%, p=NS	No statistically significant differences between Photon plus PBT boost vs. Photon alone in the probability of 5- and 10-year OS or BRFS ⊕⊕◯◯ LOW
	10-year			55.9% ± 9.0% vs. 60.6% ± 5.7%, p=NS	
Probability, Biochemical Relapse Free Survival (BRFS)	5-year			60% ± 5.4% vs. 61.9% ± 4.4%, p=NS	
	10-year			45.5% ± 8.5% vs. 42.8% ± 7.1%, p=NS	
Safety (KQ3) (curative intent only)					
Quasi-RCT					
GI toxicity, probability	Acute	Khmelevsky 2018 quasi-RCT (N=289) Moderately high RoB Prostate Cancer Risk: High (53%), Intermediate (42%), Low (5%)	**Consistency** Unknown **Imprecision** Yes[3] (-1)	**Grade 2**: 54.4% ± 5.4% vs. 69.2% ± 5.7%, p<0.01 **Grade 3 or 4**: 0% vs. 0%	There were no statistically significant differences in the probabilities of grade 3 or 4 toxicities; however, acute and late Grade 2 GI, but not GU, toxicity, were significantly lower in patients who received the PBT boost versus photons only. The actuarial frequency of grade ≥3 GI and GU toxicities was lower in the PBT boost group but statistical testing was not done. ⊕⊕◯◯ LOW
	Late			**Grade 2**: 10.2% ± 5.5% vs. 34.8% ± 7.4%, p<0.01 **Grade 3 or 4**: 0.9% ± 1.7% vs. 1.3% ± 1.8%, p=NS	
GU toxicity, probability	Acute			**Grade 2**: 33.3% ± 4.6% vs. 36.1% ± 3.5%, p=NS **Grade 3 or 4**: PBT: 0% vs. 1.9% ± 1.8%, p=NS	
	Late			**Grade 2**: 8.3% ± 5.0% vs. 9.1% ± 4.5%, p=NS **Grade 3 or 4**: 2.8% ± 2.6% vs. 3.8% ± 3.0%, p=NS	
Actuarial frequency of GI and GU toxicities, Grade ≥3	10 years			1.7% vs. 8.7%, p=NR	

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
Observational studies					
GI toxicity	Acute	Dutz 2019 (N=58) Retro propensity score-matched cohort Prostate Cancer Risk: Low (3%), Intermediate (78%), High (19%)	**Imprecision** Yes[3] (-1)	*Proportion of patients* **Grade 1:** 48% (14/29) vs. 38% (11/29); RR 1.27 (95% CI 0.70 to 2.32)† **Grade 2:** 14% (4/29) vs. 17% (5/29); RR 0.80 (95% CI 0.24 to 2.68)† **Grade 3:** 3% (1/29) vs. 0% (0/29), p=0.60	In the two clinical studies, there were no statistical difference between PBT and IMRT in acute or late toxicity (GI or GU). In the large database study, PBT resulted in lower cumulative incidences of any grade GI and GU toxicity and erectile dysfunction compared with IMRT; differences between groups were small and clinical significance is unknown. However, only the incidence of urethral stricture remained significant in a sensitivity analysis using validated diagnosis and procedure codes for severe toxicities post-pelvic radiation. ⊕⊕◯◯ LOW
		Fang 2015 (N=188) Retro case-matched cohort Prostate Cancer Risk: Low (55%), Intermediate (31%), High (7%)	**Imprecision** Yes[3] (-1)	*Proportion of patients* **Grade 0 to 1:** 95.7% (90/94) vs. 86.2% (81/94) **Grade 2 to 3:** 4.3% (4/94) vs. 13.8% (13/94); adj. OR 0.27 (0.06 to 1.24); p=0.09	
	Late	Dutz 2019 (N=58) Retro propensity score-matched cohort Prostate Cancer Risk: Low (3%), Intermediate (78%), High (19%)	**Imprecision** Yes[3] (-1)	*Proportion of patients* **Grade 1:** 9% (2/22) vs. 27% (6/22); RR 0.33 (95% CI 0.08 to 1.47)† **Grade 2:** 9% (2/22) vs. 9% (2/22) **Grade 3:** 5% (1/22) vs. 0% (0/22), p=0.32	
		Fang 2015 (N=188) Retro case-matched cohort Prostate Cancer Risk: Low (55%), Intermediate (31%), High (7%)	**Imprecision** Yes[3] (-1)	*Proportion of patients* **Grade 0 to 1:** 87.2% (82/94) vs. 88.3% (83/94) **Grade 2 to 3:** 12.8% (12/94) vs. 10.8% (10/94); adj. HR 1.24 (0.53 to 2.94) p=0.62	
		Pan 2018 (N=4158) Retro propensity-score matched database study‡ Prostate Cancer Risk: NR	**Imprecision** Yes[3] (-1)	*Cumulative incidence, any bowel toxicity (any grade)* **6-months**: 1.6% (n=693) vs. 3.2% (n=3465) **12-months**: 7.4% (n=572) vs. 7.7% (n=2862) **24-months**: 19.5% (n=341) vs. 15.4% (n=1718) **36-months**: 24.9% (n=205) vs. 19.2% (n=1003)	

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
				HR 1.27 (1.05 to 1.55); p=0.02 *Sensitivity analysis* based on validated diagnosis and procedure codes for severe toxicities post-pelvic radiation showed no difference in rectal complications between groups at 24 months (1.5% vs. 2.0%; HR 1.19, 95% CI 0.62 to 2.30)	
GU toxicity	Acute	Dutz 2019 (N=58) Retro propensity score-matched cohort Prostate Cancer Risk: Low (3%), Intermediate (78%), High (19%)	**Imprecision** Yes[3] (-1)	*Proportion of patients* **Grade 1:** 66% (19/29) vs. 45% (13/29); RR 1.46 (95% CI 0.90 to 2.37)† **Grade 2:** 24% (7/29) vs. 41% (12/29); RR 0.58 (95% CI 0.27 to 1.27)† **Grade 3:** 3% (1/29) vs. 3% (1/29)	
		Fang 2015 (N=188) Retro case-matched cohort Prostate Cancer Risk: Low (55%), Intermediate (31%), High (7%)	**Imprecision** Yes[3] (-1)	*Proportion of patients* **Grade 0 to 1:** 78.7% (74/94) vs. 71.3% (67/94) **Grade 2 to 3:** 21.3% (20/94) vs. 28.7% (27/94); adj. OR 0.69 (0.32 to 1.51); p= 0.36	
	Late	Dutz 2019 (N=58) Retro propensity score-matched cohort Prostate Cancer Risk: Low (3%), Intermediate (78%), High (19%)	**Imprecision** Yes[3] (-1)	*Proportion of patients* **Grade 1:** 23% (5/22) vs. 32% (7/22); RR 0.71 (95% CI 0.27 to 1.91)† **Grade 2:** 23% (5/22) vs. 27% (6/22); RR 0.83 (95% CI 0.30 to 2.33)† **Grade 3:** 0% (0/22) vs. 5% (1/22), p=0.32	
		Fang 2015 (N=188) Retro case-matched cohort Prostate Cancer Risk: Low (55%), Intermediate (31%), High (7%)	**Imprecision** Yes[3] (-1)	*Proportion of patients* **Grade 0 to 1:** 87.2% (82/94) vs. 80.9% (76/94) **Grade 2 to 3:** 12.8% (12/94) vs. 18.3% (17/94); adj. HR 0.56 (0.22 to 1.41); p=0.22	

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
		Pan 2018 (N=4158) Retro propensity-score matched database study‡ Prostate Cancer Risk: NR	**Imprecision** Yes[3] (-1)	*Cumulative incidence, any urinary toxicity (any grade)* **6-months**: 12.1% (n=693) vs. 21.5% (n=3465) **12-months**: 23.1% (n=572) vs. 31.6% (n=2862) **24-months**: 33.3% (n=341) vs. 42.2% (n=1718) **36-months**: 39.1% (n=205) vs. 48.3% (n=1003) HR 0.72 (0.63 to 0.83); p<0.001 *Sensitivity analysis* based on validated diagnosis and procedure codes for severe toxicities post-pelvic radiation showed less urethral stricture with PBT at 24 months (1.3% vs. 0%; HR 0.12, 95% CI 0.02 to 0.86); no differences in cystitis, ureteral stricture, or urinary/rectal fistula.	
Erectile dysfunction (cumulative incidence)	36 mons.	Pan 2018 (N=4158) Retro propensity-score matched database study‡ Prostate Cancer Risk: NR	**Imprecision** Yes[3] (-1)	**6-months**: 5.0% (n=693) vs. 9.7% (n=3465) **12-months**: 10.6% (n=572) vs. 18.1% (n=2862) **24-months**: 20.7% (n=341) vs. 27.8% (n=1718) **36-months**: 28.6% (n=205) vs. 34.3% (n=1003) HR 0.71 (0.59 to 0.84); p=0.001 *Sensitivity analysis* using procedure codes only (as surrogate for toxicity severity), 24 month incidence: 2.0% vs. 3.1%,	

Outcome	Time	Studies, Year, N, RoB Tumor	Reason for Downgrading	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
				HR 0.63, 95% CI 0.36 to 1.10	

adj. = adjusted; CI = confidence interval; GI = gastrointestinal; GU = genitourinary; HR = hazard ratio; IMRT = intensity-modulated radiation therapy; KQ = Key Question; NR = not reported; NS = not statistically significant; OR = odds ratio; PBT = proton beam therapy; Retro = retrospective study design; RR = risk ratio

* **Khmelevsky 2018**: Photon (standard conformal) + PBT boost vs. Photon (standard conformal) alone.

Dutz 2019: PBT (passive scatter) vs. IMRT

Fang 2015: PBT (passive scatter) vs. IMRT

Pan 2018: PBT vs. IMRT

†RR and 95% CI were calculated by AAI.

‡MarketScan Commercial Claims and Encounters database.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded

Executive Summary References

1. Cancer Prevalence and Cost of Care Projections. National Cancer Institute. at https://costprojections.cancer.gov/.)
2. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.
3. Evolution of Cancer Treatments: Radiation. American Cancer Society, 2014. at https://www.cancer.org/cancer/cancer-basics/history-of-cancer/cancer-treatment-radiation.html.)
4. Cancer Statistics. National Cancer Institute, 2018. at https://www.cancer.gov/about-cancer/understanding/statistics.)
5. Stereotactic radiosurgery. Mayo Clinic, 2018. at https://www.mayoclinic.org/tests-procedures/stereotactic-radiosurgery/about/pac-20384526.)
6. Adeberg S, Bernhardt D, Harrabi SB, et al. Sequential proton boost after standard chemoradiation for high-grade glioma. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;125:266-72.
7. Agarwal A, Thaker NG, Tawk B, et al. The evolution of radiation therapy for retinoblastoma: the MD Anderson Cancer Center experience. International Journal of Particle Therapy 2016;2:490-8.
8. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401-6.
9. Baskar R, Lee KA, Yeo R, Yeoh KW. Cancer and radiation therapy: current advances and future directions. International journal of medical sciences 2012;9:193-9.
10. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD)2013.
11. Bielamowicz K, Okcu MF, Sonabend R, et al. Hypothyroidism after craniospinal irradiation with proton or photon therapy in patients with medulloblastoma. Pediatric hematology and oncology 2018;35:257-67.
12. Bishop AJ, Greenfield B, Mahajan A, et al. Proton beam therapy versus conformal photon radiation therapy for childhood craniopharyngioma: multi-institutional analysis of outcomes, cyst dynamics, and toxicity. International journal of radiation oncology, biology, physics 2014;90:354-61.
13. Blanchard P, Garden AS, Gunn GB, et al. Intensity-modulated proton beam therapy (IMPT) versus intensity-modulated photon therapy (IMRT) for patients with oropharynx cancer - A case matched analysis. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;120:48-55.
14. Boker A, Pilger D, Cordini D, et al. Neoadjuvant proton beam irradiation vs. adjuvant ruthenium brachytherapy in transscleral resection of uveal melanoma. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie 2018.

15. Bronk JK, Guha-Thakurta N, Allen PK, Mahajan A, Grosshans DR, McGovern SL. Analysis of pseudoprogression after proton or photon therapy of 99 patients with low grade and anaplastic glioma. Clinical and translational radiation oncology 2018;9:30-4.
16. Bush DA, Smith JC, Slater JD, et al. Randomized Clinical Trial Comparing Proton Beam Radiation Therapy with Transarterial Chemoembolization for Hepatocellular Carcinoma: Results of an Interim Analysis. International journal of radiation oncology, biology, physics 2016;95:477-82.
17. Chowdhary M, Lee A, Gao S, et al. Is Proton Therapy a "Pro" for Breast Cancer? A Comparison of Proton vs. Non-proton Radiotherapy Using the National Cancer Database. Frontiers in oncology 2018;8:678.
18. Connell PP, Hellman S. Advances in radiotherapy and implications for the next century: a historical perspective. Cancer research 2009;69:383-92.
19. Dutz A, Agolli L, Baumann M, et al. Early and late side effects, dosimetric parameters and quality of life after proton beam therapy and IMRT for prostate cancer: a matched-pair analysis. Acta oncologica (Stockholm, Sweden) 2019:1-10.
20. Eaton BR, Esiashvili N, Kim S, et al. Endocrine outcomes with proton and photon radiotherapy for standard risk medulloblastoma. Neuro-oncology 2016b;18:881-7.
21. Eaton BR, Esiashvili N, Kim S, et al. Clinical Outcomes Among Children With Standard-Risk Medulloblastoma Treated With Proton and Photon Radiation Therapy: A Comparison of Disease Control and Overall Survival. International journal of radiation oncology, biology, physics 2016a;94:133-8.
22. Fang P, Mick R, Deville C, et al. A case-matched study of toxicity outcomes after proton therapy and intensity-modulated radiation therapy for prostate cancer. Cancer 2015;121:1118-27.
23. Fang P, Shiraishi Y, Verma V, et al. Lymphocyte-Sparing Effect of Proton Therapy in Patients with Esophageal Cancer Treated with Definitive Chemoradiation. Int J Part Ther 2018;4:23-32.
24. Garibaldi C, Jereczek-Fossa BA, Marvaso G, et al. Recent advances in radiation oncology. Ecancermedicalscience 2017;11:785.
25. Giap H, Giap B. Historical perspective and evolution of charged particle beam therapy. Translational Cancer Research 2012;1:127-36.
26. Giovannini G, Bohlen T, Cabal G, et al. Variable RBE in proton therapy: comparison of different model predictions and their influence on clinical-like scenarios. Radiation oncology (London, England) 2016;11:68.
27. Goitein M. Magical protons? International Journal of Radiation Oncology• Biology• Physics 2008;70:654-6.
28. Grant SR, Grosshans DR, Bilton SD, et al. Proton versus conventional radiotherapy for pediatric salivary gland tumors: Acute toxicity and dosimetric characteristics. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2015;116:309-15.
29. Gunther JR, Rahman AR, Dong W, et al. Craniospinal irradiation prior to stem cell transplant for hematologic malignancies with CNS involvement: Effectiveness and toxicity after photon or proton treatment. Pract Radiat Oncol 2017;7:e401-e8.

30. Gunther JR, Sato M, Chintagumpala M, et al. Imaging Changes in Pediatric Intracranial Ependymoma Patients Treated With Proton Beam Radiation Therapy Compared to Intensity Modulated Radiation Therapy. International journal of radiation oncology, biology, physics 2015;93:54-63.
31. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383-94.
32. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 2011;64:380-2.
33. Hashimoto T, Isobe T, Hashii H, et al. Influence of secondary neutrons induced by proton radiotherapy for cancer patients with implantable cardioverter defibrillators. Radiation oncology 2012;7:10.
34. Higgins KA, O'Connell K, Liu Y, et al. National Cancer Database Analysis of Proton Versus Photon Radiation Therapy in Non-Small Cell Lung Cancer. International journal of radiation oncology, biology, physics 2017;97:128-37.
35. Hirano E, Fuji H, Onoe T, Kumar V, Shirato H, Kawabuchi K. Cost-effectiveness analysis of cochlear dose reduction by proton beam therapy for medulloblastoma in childhood. J Radiat Res 2014;55:320-7.
36. Holliday EB, Garden AS, Rosenthal DI, et al. Proton therapy reduces treatment-related toxicities for patients with nasopharyngeal cancer: a case-match control study of intensity-modulated proton therapy and intensity-modulated photon therapy. International journal of particle therapy 2015;2:19-28.
37. Jarlskog CZ, Paganetti H. Sensitivity of different dose scoring methods on organ-specific neutron dose calculations in proton therapy. Physics in Medicine & Biology 2008;53:4523.
38. Jhaveri J, Cheng E, Tian S, et al. Proton vs. Photon Radiation Therapy for Primary Gliomas: An Analysis of the National Cancer Data Base. Frontiers in oncology 2018;8:440.
39. Kahalley LS, Ris MD, Grosshans DR, et al. Comparing Intelligence Quotient Change After Treatment With Proton Versus Photon Radiation Therapy for Pediatric Brain Tumors. J Clin Oncol 2016;34:1043-9.
40. Kahalley LS, Ris MD, Mahajan A, et al. Prospective, Longitudinal Comparison of Neurocognitive Change in Pediatric Brain Tumor Patients Treated with Proton Radiotherapy versus Surgery Only. Neuro-oncology 2019.
41. Khmelevsky EV, Kancheli IN, Khoroshkov VS, Kaprin AD. Morbidity dynamics in proton-photon or photon radiation therapy for locally advanced prostate cancer. Reports of practical oncology and radiotherapy : journal of Greatpoland Cancer Center in Poznan and Polish Society of Radiation Oncology 2018;23:21-7.
42. Kjellberg R, Sweet W, Preston W, Koehler A. The Bragg peak of a proton beam in intracranial therapy of tumors. Transactions of the American Neurological Association (US) 1962;87.
43. Kopecky AS, Khan AJ, Pan W, Drachtman R, Parikh RR. Outcomes and patterns of care in a nationwide cohort of pediatric medulloblastoma: Factors affecting proton therapy utilization. Adv Radiat Oncol 2017;2:588-96.
44. Leung HW, Chan AL. Cost-utility of stereotactic radiation therapy versus proton beam therapy for inoperable advanced hepatocellular carcinoma. Oncotarget 2017;8:75568.

45. Levin W, Kooy H, Loeffler J, DeLaney T. Proton beam therapy. British journal of Cancer 2005;93:849.

46. Liao Z, Lee JJ, Komaki R, et al. Bayesian Adaptive Randomization Trial of Passive Scattering Proton Therapy and Intensity-Modulated Photon Radiotherapy for Locally Advanced Non-Small-Cell Lung Cancer. J Clin Oncol 2018;36:1813-22.

47. Lin AJ, Rao YJ, Acharya S, Schwarz J, Rao PK, Grigsby P. Patterns of care and outcomes of proton and eye plaque brachytherapy for uveal melanoma: Review of the National Cancer Database. Brachytherapy 2017;16:1225-31.

48. Lin SH, Merrell KW, Shen J, et al. Multi-institutional analysis of radiation modality use and postoperative outcomes of neoadjuvant chemoradiation for esophageal cancer. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;123:376-81.

49. Maemura K, Mataki Y, Kurahara H, et al. Comparison of proton beam radiotherapy and hyper-fractionated accelerated chemoradiotherapy for locally advanced pancreatic cancer. Pancreatology 2017;17:833-8.

50. Mailhot Vega R, Kim J, Hollander A, et al. Cost effectiveness of proton versus photon radiation therapy with respect to the risk of growth hormone deficiency in children. Cancer 2015;121:1694-702.

51. Mailhot Vega RB, Ishaq O, Raldow A, et al. Establishing cost-effective allocation of proton therapy for breast irradiation. International Journal of Radiation Oncology* Biology* Physics 2016;95:11-8.

52. Makishima H, Ishikawa H, Terunuma T, et al. Comparison of adverse effects of proton and X-ray chemoradiotherapy for esophageal cancer using an adaptive dose-volume histogram analysis. J Radiat Res 2015;56:568-76.

53. McDonald MW, Liu Y, Moore MG, Johnstone PA. Acute toxicity in comprehensive head and neck radiation for nasopharynx and paranasal sinus cancers: cohort comparison of 3D conformal proton therapy and intensity modulated radiation therapy. Radiation oncology (London, England) 2016;11:32.

54. Moriarty JP, Borah BJ, Foote RL, Pulido JS, Shah ND. Cost-effectiveness of proton beam therapy for intraocular melanoma. PloS one 2015;10:e0127814.

55. Mozes P, Dittmar JO, Habermehl D, et al. Volumetric response of intracranial meningioma after photon or particle irradiation. Acta oncologica (Stockholm, Sweden) 2017;56:431-7.

56. Newberry SJ, Ahmadzai N, Motala A, et al. AHRQ Methods for Effective Health Care. Surveillance and Identification of Signals for Updating Systematic Reviews: Implementation and Early Experience. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013.

57. Niedzielski JS, Yang J, Mohan R, et al. Differences in Normal Tissue Response in the Esophagus Between Proton and Photon Radiation Therapy for Non-Small Cell Lung Cancer Using In Vivo Imaging Biomarkers. International journal of radiation oncology, biology, physics 2017;99:1013-20.

58. Paganetti H, Niemierko A, Ancukiewicz M, et al. Relative biological effectiveness (RBE) values for proton beam therapy. International Journal of Radiation Oncology* Biology* Physics 2002;53:407-21.

59. Pan HY, Jiang J, Hoffman KE, et al. Comparative Toxicities and Cost of Intensity-Modulated Radiotherapy, Proton Radiation, and Stereotactic Body Radiotherapy Among Younger Men With Prostate Cancer. J Clin Oncol 2018;36:1823-30.

60. Paulino AC, Mahajan A, Ye R, et al. Ototoxicity and cochlear sparing in children with medulloblastoma: Proton vs. photon radiotherapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2018;128:128-32.

61. Remick JS, Schonewolf C, Gabriel P, et al. First Clinical Report of Proton Beam Therapy for Postoperative Radiotherapy for Non-Small-Cell Lung Cancer. Clinical lung cancer 2017;18:364-71.

62. Romesser PB, Cahlon O, Scher E, et al. Proton beam radiation therapy results in significantly reduced toxicity compared with intensity-modulated radiation therapy for head and neck tumors that require ipsilateral radiation. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;118:286-92.

63. Sanford NN, Pursley J, Noe B, et al. Protons versus Photons for Unresectable Hepatocellular Carcinoma: Liver Decompensation and Overall Survival. International journal of radiation oncology, biology, physics 2019.

64. Sato M, Gunther JR, Mahajan A, et al. Progression-free survival of children with localized ependymoma treated with intensity-modulated radiation therapy or proton-beam radiation therapy. Cancer 2017;123:2570-8.

65. Schünemann HJ, Tugwell P, Reeves BC, et al. Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions. Research synthesis methods 2013;4:49-62.

66. Sharma S, Zhou O, Thompson R, et al. Quality of Life of Postoperative Photon versus Proton Radiation Therapy for Oropharynx Cancer. International Journal of Particle Therapy 2018;5:11-7.

67. Shekelle PG, Newberry SJ, Wu H, et al. AHRQ Methods for Effective Health Care. Identifying Signals for Updating Systematic Reviews: A Comparison of Two Methods. Rockville (MD): Agency for Healthcare Research and Quality (US); 2011.

68. Sher DJ, Tishler RB, Pham N-L, Punglia RS. Cost-Effectiveness Analysis of Intensity Modulated Radiation Therapy Versus Proton Therapy for Oropharyngeal Squamous Cell Carcinoma. International journal of radiation oncology, biology, physics 2018;101:875-82.

69. Radiation Therapy: Which type is right for me? OncoLink Penn Medicine, 2018. at https://www.oncolink.org/cancer-treatment/radiation/introduction-to-radiation-therapy/radiation-therapy-which-type-is-right-for-me.)

70. Shiraishi Y, Fang P, Xu C, et al. Severe lymphopenia during neoadjuvant chemoradiation for esophageal cancer: A propensity matched analysis of the relative risk of proton versus photon-based radiation therapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2018;128:154-60.

71. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Annals of internal medicine 2007;147:224-33.

72. Sikuade MJ, Salvi S, Rundle PA, Errington DG, Kacperek A, Rennie IG. Outcomes of treatment with stereotactic radiosurgery or proton beam therapy for choroidal melanoma. Eye (London, England) 2015;29:1194-8.

73. Simon F, Feuvret L, Bresson D, et al. Surgery and protontherapy in Grade I and II skull base chondrosarcoma: A comparative retrospective study. PLoS One 2018;13:e0208786.

74. Sio TT, Lin HK, Shi Q, et al. Intensity Modulated Proton Therapy Versus Intensity Modulated Photon Radiation Therapy for Oropharyngeal Cancer: First Comparative Results of Patient-Reported Outcomes. International journal of radiation oncology, biology, physics 2016;95:1107-14.

75. Song S, Park HJ, Yoon JH, et al. Proton beam therapy reduces the incidence of acute haematological and gastrointestinal toxicities associated with craniospinal irradiation in pediatric brain tumors. Acta oncologica (Stockholm, Sweden) 2014;53:1158-64.

76. Teichman SL, Do S, Lum S, et al. Improved long-term patient-reported health and well-being outcomes of early-stage breast cancer treated with partial breast proton therapy. Cancer Med 2018;7:6064-76.

77. Thorp N. Proton therapy for children. RAD Magazine 2010;36:17-8.

78. Tian X, Liu K, Hou Y, Cheng J, Zhang J. The evolution of proton beam therapy: Current and future status. Mol Clin Oncol 2018;8:15-21.

79. Tucker SL, Liu A, Gomez D, et al. Impact of heart and lung dose on early survival in patients with non-small cell lung cancer treated with chemoradiation. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;119:495-500.

80. Unkelbach J, Chan TC, Bortfeld T. Accounting for range uncertainties in the optimization of intensity modulated proton therapy. Physics in Medicine & Biology 2007;52:2755.

81. Wang XS, Shi Q, Williams LA, et al. Prospective Study of Patient-Reported Symptom Burden in Patients With Non-Small-Cell Lung Cancer Undergoing Proton or Photon Chemoradiation Therapy. J Pain Symptom Manage 2016;51:832-8.

82. Xi M, Xu C, Liao Z, et al. Comparative Outcomes After Definitive Chemoradiotherapy Using Proton Beam Therapy Versus Intensity Modulated Radiation Therapy for Esophageal Cancer: A Retrospective, Single-Institutional Analysis. International journal of radiation oncology, biology, physics 2017;99:667-76.

83. Yock TI, Tarbell NJ. Technology insight: Proton beam radiotherapy for treatment in pediatric brain tumors. Nature Reviews Clinical Oncology 2004;1:97.

84. Zhang W, Zhang X, Yang P, et al. Intensity-modulated proton therapy and osteoradionecrosis in oropharyngeal cancer. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;123:401-5.

1 Appraisal

1.1 Background and Rationale

Overall, it's estimated that 1.7 million new cases of cancer are diagnosed yearly and cancerous conditions are responsible for over half a million deaths per year. Treatment options for cancerous and noncancerous conditions vary depending on the type and stage of cancer and can include radiation therapy, chemotherapy, targeted therapy (e.g. inhibitor drugs), immunotherapy (including monoclonal antibodies) and surgery. In recent years the use of proton beam therapy (PBT) has expanded to include a variety of conditions including a number of cancer types, noncancerous brain tumors and cancerous conditions afflicting the central nervous system as well as eyes, lungs, liver, prostate, spine, and pelvis. The use of protons for radiotherapy has a history of over 60 years of clinical use. In conventional radiotherapy, photons deliver radiation across tissue depths on the way toward the target tumor and beyond. In contrast, PBT, which is a form of external beam radiotherapy, deposits peak radiation energy more precisely at or around the target followed by sharp decline in energy output to deeper tissues via a phenomenon known as the Bragg peak.[155] Because the proton beam is focused on a specific area, a greater dose of radiation may be delivered to the target neoplasm(s) while mitigating unwanted radiation delivered to surrounding tissue.[160] PBT use was initially directed towards conditions where sparing sensitive adjacent normal tissues was considered to be of utmost importance (such as cancerous or noncancerous malformations of the brain stem, eye, or spinal cord) or for many pediatric tumors because of the particular risk of pronounced acute and long-term toxicity in pediatric patients.[279] PBT may be most promising for tumors in moderate proximity to (>2 cm) to organs at risk (OAR).
In the past two decades the number of centers offering PBT has increased to over 20, with more planned or under construction, even given the high cost of facility construction and operation. Despite increasing availability of PBT and its potential for precise delivery of radiation therapy, its effectiveness compared with other forms of therapy and with the emerging techniques, such as intensity modulated radiation therapy (IMRT), is evolving and currently is unclear for some conditions.

Policy Context
This topic was originally reviewed in 2014. It is being re-reviewed in 2019 due to newly available published evidence.

Objectives
The aim of this report is to update the 2014 HTA on proton beam therapy (PBT) by systematically reviewing, critically appraising and analyzing new research evidence on the safety and efficacy of PBT, both as a primary or as a salvage therapy (i.e., for recurrent disease or failure of initial therapy), for the treatment of multiple types of cancer as well as selected noncancerous conditions in adults and children.

1.2 Key Questions

1. What is the comparative impact of proton beam therapy treatment with curative intent on survival, disease progression, health-related quality of life, and other patient outcomes versus radiation therapy alternatives and other cancer-specific treatment options (e.g., surgery, chemotherapy) for the following conditions:
    a. Cancers
        i. Bone tumors
        ii. Brain, spinal, and paraspinal tumors
        iii. Breast cancer
        iv. Esophageal cancer
        v. Gastrointestinal cancers
        vi. Gynecologic cancers
        vii. Head and neck cancers (including skull base tumors)
        viii. Liver cancer
        ix. Lung cancer
        x. Lymphomas
        xi. Ocular tumors
        xii. Pediatric cancers (e.g., medulloblastoma, retinoblastoma, Ewing's sarcoma)
        xiii. Prostate cancer
        xiv. Soft tissue sarcomas
        xv. Seminoma
        xvi. Thymoma
        xvii. Other cancers
    b. Noncancerous Conditions
        iv. Arteriovenous malformations
        v. Hemangiomas
        vi. Other benign tumors (e.g., acoustic neuromas, pituitary adenomas)
2. What is the comparative impact of salvage treatment (including treatment for recurrent disease) with proton beam therapy versus major alternatives on survival, disease progression, health-related quality of life, and other patient outcomes versus radiation therapy alternatives and other cancer-specific treatment options (e.g., surgery, chemotherapy) for the condition types listed in key question 1?
3. What are the comparative harms associated with the use of proton beam therapy relative to its major alternatives, including acute (i.e., within the first 90 days after treatment) and late (>90 days) toxicities, systemic effects such as fatigue and erythema, toxicities specific to each cancer type (e.g., bladder/bowel incontinence in prostate cancer, pneumonitis in lung or breast cancer), risks of secondary malignancy, and radiation dose?
4. What is the differential effectiveness and safety of proton beam therapy according to factors such as age, sex, race/ethnicity, disability, presence of comorbidities, tumor characteristics (e.g., tumor volume and location, proliferative status, genetic variation) and treatment protocol (e.g., dose, duration, timing of intervention, use of concomitant therapy)?
5. What is the comparative cost-effectiveness of proton beam therapy in the short- and long-term relative to other types of radiation therapy, radiation therapy alternatives or other cancer-specific treatment options (e.g., surgery, chemotherapy)

Inclusion and exclusion criteria are summarized as follows and are detailed in the full report. Briefly, incl uded studies met the following requirements with respect to participants, intervention comparators, outcomes, and study design:

- **Population**: Adults and children undergoing treatment of primary or recurrent disease, to include cancer types (bone cancer, brain, spinal, and paraspinal tumors, breast cancer, esophageal cancer, gastrointestinal cancer, gynecologic cancer, head and neck cancer, liver cancer, lung cancer, lymphomas, ocular tumors, pediatric cancers, prostate cancer, sarcomas, seminoma, thymoma, other cancers) and noncancerous conditions (arteriovenous malformations, hemangiomas, other benign tumors).
- Interventions: Proton beam therapy; all approaches were considered including monotherapy, use as a "boost" mechanism to conventional radiation, and combination therapy with other treatment modalities (e.g., chemotherapy, surgery).
- Comparators: Primary comparators include other radiation alternatives (e.g., intensity-modulated radiation therapy (IMRT), stereotactic radiation techniques and other external beam therapies, and brachytherapy). Other treatment alternatives specific to each condition type treated, and may include chemotherapy, immunotherapy, surgical procedures, and other devices (e.g., laser therapy for ocular tumors).

- **Outcomes:**

Primary Clinical outcomes:

- Overall survival/disease-free survival
- All-cause and/or disease-related mortality
- Direct measures of tumor regression, control or recurrence
- Incidence of metastases

Secondary or indirect (intermediate) outcomes

- Patient reported outcomes including health-related quality of life (HrQoL) using validated instruments
- Requirements for subsequent therapy
- Other outcomes specific to particular conditions (e.g., visual acuity for ocular tumors, shunt requirements for arteriovenous malformations)
- Intermediate measures of tumor recurrence such as biochemical measures

Safety outcomes:

- Treatment-related harms, to include generalized effects (e.g., fatigue, erythema) and localized toxicities specific to each condition (e.g., urinary incontinence in prostate cancer, pulmonary toxicity in lung or breast cancer); the primary focus is on adverse effects requiring medical attention
- Secondary malignancy risk due to radiation exposure

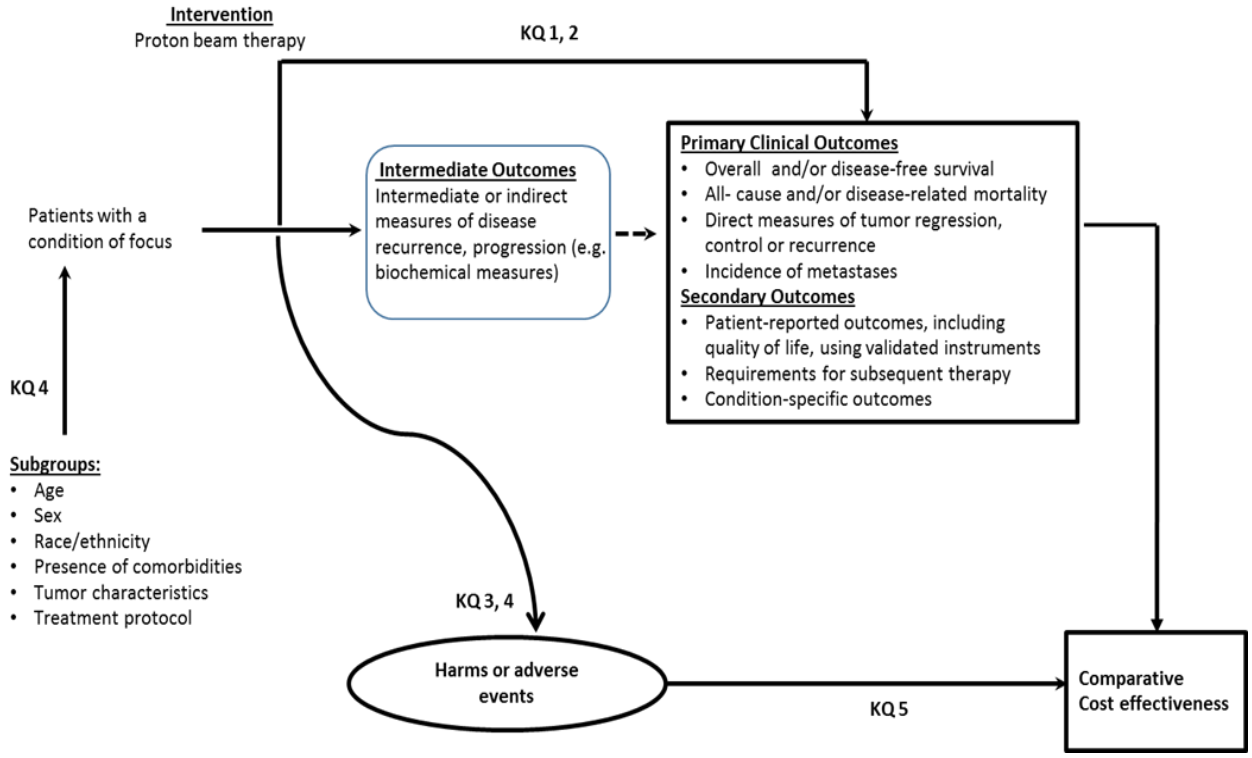
Economic outcomes:

- Long term and short term comparative cost-effectiveness measures

- **Studies:** The focus will be on high quality (low risk of bias) comparative studies (e.g., randomized controlled trials, comparative cohort studies with concurrent controls) will be considered for Key Questions 1-4. Comparative observational studies with long term clinical outcomes or safety will be considered for Key Questions 1-4. Case series will be considered but will not be the primary focus of evaluation for each key question. Dosimetry and planning studies will be included for context; the will be included as evidence if they directly answer the key questions.

Full, comparative, formal economic studies (i.e., cost-effectiveness, cost-utility, cost-minimization, and cost-benefit studies) will be sought for Key Question 5; studies using modeling may be used to determine cost-effectiveness.

Figure 1. Analytic Framework



1.3 Outcomes Assessed

The primary outcomes of interest for this report are listed below.

- Overall survival (OS)
- Progression-free survival (PFS) or Local control (LC)
  - Disease-free survival (DFS)
  - Relapse-free survival (RFS)
- Treatment-related toxicity (as reported specific to PBT when possible) and secondary malignancy risk due to radiation exposure

OS and PFS were stated *a priori* as primary outcomes of interest. Some of the included studies also reported DFS and RFS. Excluding OS, definitions of these outcomes varied slightly between the studies. Other outcomes reported included health-related quality of life (based on validated instruments), incidence of metastases, and other outcomes specific to particular conditions.

Outcomes are detailed in the evidence tables in the appendices and/or the body of the report. Summary tables for case series are also found in the appendices.

Strength of evidence was assessed for the primary clinical outcomes only.

Table 1. Outcome measures reported on in included studies

Outcome measure	Assessed By	Components	Score range	Interpretation	MCID*
MD Anderson Symptom Inventory-Head and Neck Cancer (MDASI-HN)[51,240,262]	Patient	22 items (0 to 10 points each) that are grouped into three separate domains: Interference items • Walking • Activity • Working (including housework) • Relations with other people • Enjoyment of life • Mood Core symptoms (13 items) • Pain • Fatigue • Nausea • Disturbed sleep • Distress/feeling upset • Shortness of breath • Difficulty remembering • Lack of appetite • Drowsiness • Dry mouth • Sadness • Vomiting	0 to 10 points	0: not present 10: as bad as you can imagine	1.16 points

Outcome measure	Assessed By	Components	Score range	Interpretation	MCID*
		• Numbness/tingling Head and neck cancer module items (9 items) • Mucus in the mouth and throat • Difficulty swallowing or chewing • Choking or coughing • Difficulty with voice or speech • Skin pain/burning/rash • Constipation • Problems with tasting food • Mouth/throat sores • Problems with teeth or gums			
MD Anderson Symptom Inventory (MDASI)[51,52,262]	Physician	13 symptom items and 6 interference items (0 to 10 points each) Interference items • Walking • Activity • Working (including housework) • Relations with other people • Enjoyment of life • Mood Symptom Items • Pain • Fatigue • Nausea • Disturbed sleep • Distress/feeling upset • Shortness of breath • Difficulty remembering • Lack of appetite • Drowsiness • Dry mouth • Sadness • Vomiting • Numbness/tingling	0 to 10 points	0: not present 10: as bad as you can imagine	0.98 points
Leiter International	Computer	10 subsets organized into four domains designed to assess non-verbal IQ	NR	NR	NR

Outcome measure	Assessed By	Components	Score range	Interpretation	MCID*
Performance Scale[237]		• Fluid Intelligence • Visualization • Memory • Attention			
Wechsler Adult Intelligence Scale (WAIS)[162]	Physician	There are four index scores representing major components of intelligence • Verbal Comprehension Index (VCI) • Perceptual Reasoning Index (PRI) • Working Memory Index (WMI) • Processing Speed Index (PSI) Two broad scores, which can be used to summarize general intellectual abilities, can also be derived: • Full Scale IQ (FSIQ), based on the total combined performance of the VCI, PRI, WMI, and PSI • General Ability Index (GAI), based only on the six subtests that the VCI and PRI comprise.	0 to 130 (100 as an average score with a standard deviation of 15)	• Below Average: standard score below 79 • Low Average: standard score 80 to 89 • Average: 90 to 109 • High Average: 110 to 119 • Superior: 120 to 129 • Very Superior: above 130	NR
Wechsler Intelligence Scale for Children (WISC)[314]	Physician	There are five primary index scores • Verbal Comprehension Index (VCI) • Visual Spatial Index (VSI) • Fluid Reasoning Index (FRI) • Working Memory Index (WMI) • Processing Speed Index (PSI) One broad score, which can be used to summarize general intellectual abilities, can also be derived: • Full Scale IQ (FSIQ), based on the total combined performance	0 to 130 (100 as an average score with a standard deviation of 15)	• Below Average: standard score below 79 • Low Average: standard score 80 to 89 • Average: 90 to 109 • High Average: 110 to 119 • Superior: 120 to 129 • Very Superior: above 130	NR

Outcome measure	Assessed By	Components	Score range	Interpretation	MCID*
		of the VCI, VSI, FRI, WMI, and PSI			
Woodcock-Johnson Tests of Cognitive Ability[315]	Physician	20 tests consisting of two batteries: Cognitive Ability and Tests of Achievement • Comprehension-Knowledge - Verbal Comprehension - General Information • Long-Term Retrieval - Visual-Auditory Learning - Retrieval Fluency • Visual Processing - Spatial Relations - Picture Recognition • Auditory Processes - Sound Blending - Auditory Attention • Fluid Reasoning - Concept Formation - Analysis-Synthesis • Processing Speed - Visual Matching - Decision Speed • Short-Term Memory - Numbers Reversed - Memory for Words • Incomplete Words • Auditory Working Memory • Visual-Auditory Learning – Delayed • Rapid Picture Naming • Planning • Pair Cancellation	Range: <69 to >131 points 1 point awarded for correct answers, 0 points awarded for incorrect answers Age or Grade Equivalents: Reflects age or grade level at which average score is same as subject's raw score Raw score: Number correct Relative Proficiency Index (RPI): Ranges from 0/90 to 100/90. RPI predicts a student's level of proficiency on tasks that typical age- or grade-peers would perform with 90% proficiency.	• Very low: ≤69 • Low: 70 to 79 • Low Average: 80 to 89 • Average: 90 to 110 • High Average: 110 to 119 • Superior: 120 to 129 • Very Superior: ≥131	NR
Pediatric Quality of Life (PedsQL)[287,288,290]	Patient	23 items grouped into 4 domains • Physical Functioning (8 items) • Emotional Functioning (5 items) • Social Functioning (5 items)	0 to 100 points	Higher scores indicate better health related quality of life	• Total Score: 4.36 • Physical Health: 6.66 • Psychosocial Health: 5.3 • Emotional Functioning: 8.94

Outcome measure	Assessed By	Components	Score range	Interpretation	MCID*
		• School Functioning (5 items) [Emotional Functioning, Social Functioning, and School Functioning Scores can be used to acquire an overall Psychosocial Health score]			• Social Functioning: 8.36 • School Functioning: 9.12
Pediatric Quality of Life – Parent Proxy (PedsQL – Parent Proxy)[287-290]	Parent of patient	23 items grouped into 4 domains • Physical Functioning (8 items) • Emotional Functioning (5 items) • Social Functioning (5 items) • School Functioning (5 items) [Emotional Functioning, Social Functioning, and School Functioning Scores can be used to acquire an overall Psychosocial Health score]	0 to 100 points	Higher scores indicate better health related quality of life	• Total Score: 4.5 • Physical Health: 6.92 • Psychosocial Health: 5.49 • Emotional Functioning: 7.79 • Social Functioning: 8.98 • School Functioning: 9.67
Modified Epworth Sleepiness Scale[126]	Parent of patient or patient themselves (depending on age)	8 items (0 to 3 points each) Chance of dozing during following activities • Sitting and reading • Sitting and watching TV or video • Sitting in a classroom at school during the morning • Sitting and riding in a car or bus for about 30 minutes • Sitting and talking to someone • Sitting quietly by yourself after lunch • Sitting and eating a meal	0 to 3 points	• 0=no chance of dozing • 3=high chance of dozing • Impaired: total score >10 • Unimpaired: total score ≤9 (Per Jacola 2016)	NR in patient population
Mental Development Index (MDI) derived from the Bayley Scales of Infant	Physician	178 items (0 or 1 points each) addressing 5 different developmental areas • Cognitive • Language • Motor	0 points for incorrect answers 1 point for correct answers	Raw scores (the total number of correct answers) are used to calculated the Mental Development Index	NR

Outcome measure	Assessed By	Components	Score range	Interpretation	MCID*
Development[1] [7,27]		• Social-Emotional • Adaptive Behavior		Higher scores indicate an increased level of mental development. A standardized mean score is 100	
Scales of Independent Behavior-Revised (SIB-R)[38]	Parents	14 subscales, organized into 4 adaptive behavior clusters • Motor Skills • Personal Living Skills • Social Interaction and Communication Skills • Community Living Skills	Unclear	Lower scores indicate lower functioning or greater problems	NR
World Health Organization (WHO) Performance Status/Eastern Cooperative Oncology Group (ECOG) Score[210]	Physician	A single score rating from 0 to 5 that measures a patients performance status	0 to 5	• 0: Asymptomatic (Fully active, able to carry on all pre-disease activities without restriction) • 1: Symptomatic but completely ambulatory (Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature. For example, light housework, office work) • 2: Symptomatic, <50% in bed during the day (Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours) • 3: Symptomatic, >50% in bed, but not bedbound (Capable of only limited self-care, confined to bed or chair 50% or more of waking hours) • 4: Bedbound (Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair) • 5: Death	NR
Continuous Performance Test, 2nd Edition (CPT-II)[55]	Computer	The test is taken at a computer. The participant presses the space bar or clicks the mouse button when a letter other than X shows up onscreen. Letters	Not Applicable	Provides an estimate of the probability that a given child's performance resembles that of a child with clinically significant attention problems.	NR

Outcome measure	Assessed By	Components	Score range	Interpretation	MCID*
		appear on the screen with different time intervals between each one. Exactly 14 minutes is required for completion.			
Behavior Rating Inventory of Executive Function (BRIEF)[86]	Parent	86 question questionnaire Each question uses a three point scale representing Never (1), Sometimes (2), and Often (3)	0 to 3	This measure provides a T-score with a mean of 50 and a SD of 10; higher scores indicate more problems with Executive Functions.	NR
Behavior Assessment System for Children, 2nd Edition (BASC-2, Attention Subscale)[234]	Parent	134 to 160 items in which parents or caregivers rate the frequency of the child's behavior.	0 to 4 Likert scale ranging from "never occurs" to "almost always occurs"	This measure provides a T-score; higher scores indicate more attention problems.	NR
American Urological Association (AUA) Symptom Index[25]	Patient	7 questions addressing frequency, nocturia, weak urinary stream, hesitancy, intermittence, incomplete emptying, and urgency	0 to 5 points	Lower scores represent less presence of symptoms while higher scores represent	5 points or more
Expanded Prostate Cancer Index Composite (EPIC) Quality of Life Survey[128,261,311]		EPIC assesses the disease-specific aspects of prostate cancer and its therapies and comprises four summary domains (Urinary, Bowel, Sexual and Hormonal) and is made up of 50 prostate-targeted items	0 to 100 points	Higher scores represent better Health Related Quality of Life. Lower numbers corresponded to worsening function and increased bother.	*All scores are representative of the mean difference from baseline* <u>Per Norman 2003</u> Half a standard deviation <u>Per Jeldres 2015</u> Sexual Function: 11 Sexual Bother: 14 Urinary Function: 5 Urinary Bother: 8 Bowel Function: 4 Bowel Bother: 5 Hormone Function: 5 Hormone Bother: 4 <u>Per Skolarus 2015</u> Urinary Incontinence: 6 to 9 points

Outcome measure	Assessed By	Components	Score range	Interpretation	MCID*
					Urinary Irritative/Obstructive: 5 to 7 points Bowel Summary: 4 to 6 Sexual Summary: 10 to 12 Hormonal: 4 to 6

NR = Not reported

1.4 Washington State Utilization Data

Populations

The **Proton Beam Therapy** analysis includes member utilization and cost data from the following agencies: Public Employees Benefit Board Uniform Medical Plan (PEBB/UMP); PEBB Medicare; Medicaid Managed Care; and Medicaid Fee-for-Service. The Department of Labor and Industries (LNI) had no proton beam therapy claims.

The analysis period was five (5) calendar years, 2013 - 2017. Primary population inclusion criteria included incurring a paid claim(s) comprised of at least one of the targeted CPT/HCPCS codes from Table I. Initial analysis focused on diagnosis from Table IIA. Additional analysis lead to an expanded range of diagnoses codes (see Table IIB). Final data evaluation included all diagnoses for individuals undergoing proton beam therapy. Denied claims were excluded from the analysis.

Methods

First, all paid patient claims (children and adults) with a targeted CPT procedure (Table I) were identified. Second, those same claims underwent examination to identify those that also included targeted primary diagnoses codes from Table IIA (later expanded to Table IIB). Final data evaluation included examining utilization by member; by age range; analysis of individual and aggregate ICD-9 and ICD-10 codes and by paid claims' costs.

Table I
Targeted CPT Descriptions

CPT	Procedure Code Description
77520	Proton treatment delivery; simple, without compensation
77522	Proton treatment delivery; simple, with compensation
77523	Proton treatment delivery; intermediate
77525	Proton treatment delivery; complex

Table IIA
Target Diagnosis Codes: Initial

Specific Cancer Diagnosis Codes and Descriptions		
	ICD-9	ICD-10
Lung	162.0-162.9	C34.90 - C34.92
Prostate	185.0	C61
Eye	190.0-190.9	C69-C69.92
Brain	191 - 191.9	C71 - C71.9
Spinal	192.2-192.3	C72 - C72.9

Table IIB
2013 – 2017
Neoplasm ICD-9 and ICD-10 Diagnosis Codes
PEBB/UMP, Medicare/UMP, Medicaid Managed Care, Medicare Fee-for-Service
Utilization: Proton Beam Therapy

Range of codes utilized for analyzing claims*

ICD-10	ICD-10 Description/ICD-9 Description	ICD-9
C00-C14	Malignant Neoplasm of Lip, Oral cavity, and Pharynx	140-149
C15-C26	Malignant Neoplasm of Digestive Organs	150-159
C30-C39	Malignant Neoplasm of Respiratory and Intrathoracic	160-165
C40-C41	Malignant Neoplasm of Bone and Articular Cartilage	170-176
C43-C44	Malignant Neoplasm of Skin	170-176
C45-C49	Malignant Neoplasm of Mesothelial and Soft Tissue	170-176
C50	Malignant Neoplasm of Breast	170-176
C51-C63	Malignant Neoplasm of Genital organs	179-189
C64-C68	Malignant Neoplasm of Urinary Tract	190
C69-C72	Malignant Neoplasm of Eye, Brain, CNS	191-192
C73-C75	Malignant Neoplasm of Endocrine	194
C76-C80	Malignant Neoplasm Ill Defined, Secondary (and Other)	195
C81-C96	Malignant Neoplasm of Lymphoid	196, 200-208
D37-D48, D49	Neoplasm uncertain or unspecific behavior	235-239
D10-D36, D3A	Benign tumors	210-229

*1) Not all diagnoses codes were represented in the data.

2) Utilization and cost analyses contain V and/or Z codes (Supplementary Classification of Factors Influencing Health Status and Contact with Health Services) when substituted for a primary diagnosis.

Chart I
2013 – 2017
Aggregate Utilization
PEBB/UMP and Medicare/UMP & Medicaid Managed Care and Fee-for-Service
Distribution of Patients Receiving Proton Beam Therapy by Primary Cancer Diagnosis
N=246

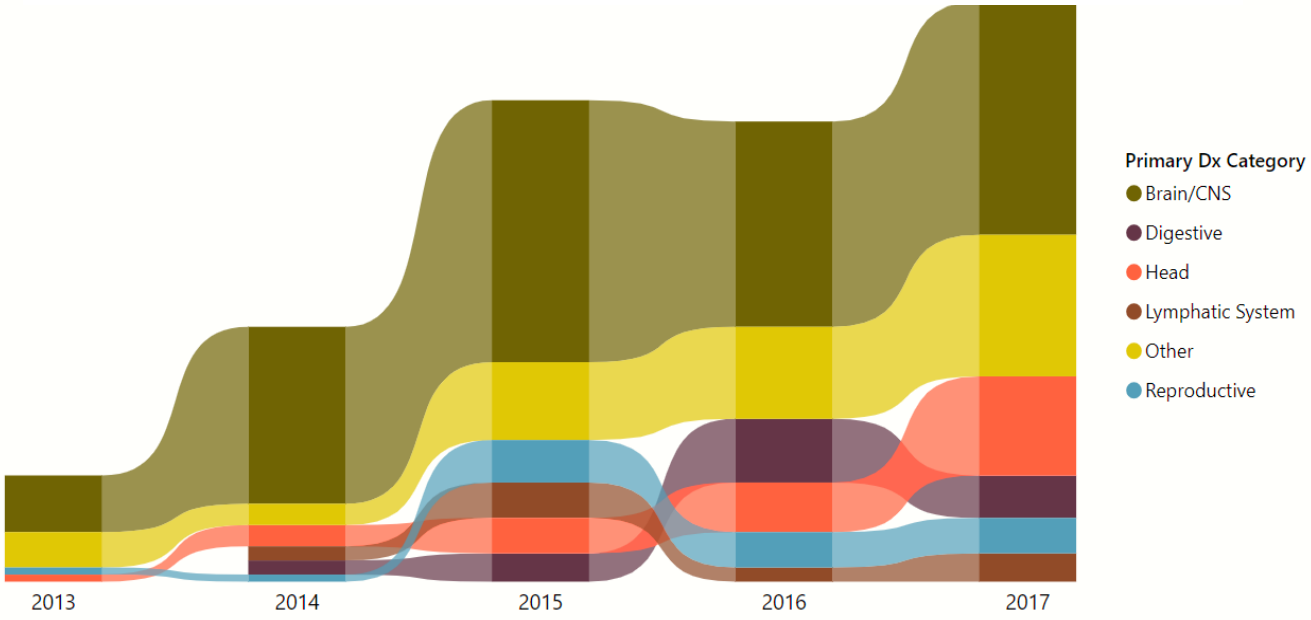
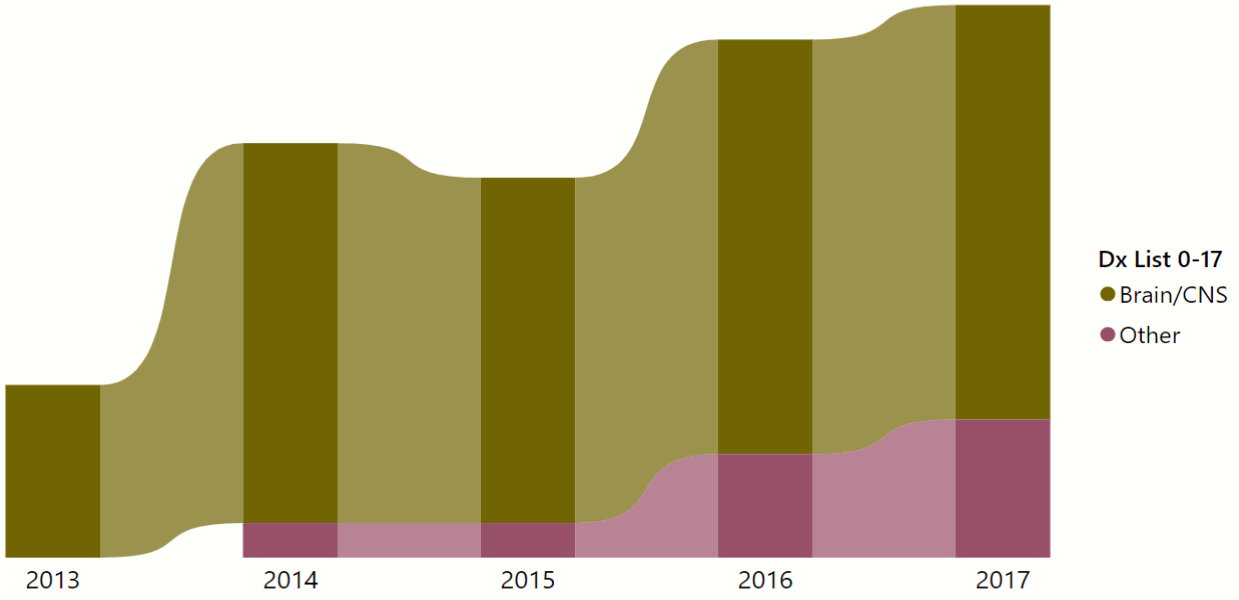


Chart II
2013 – 2017
Aggregate Utilization
PEBB/UMP and Medicare/PEBB
2013 – 2017
Medicaid Manage Care and Fee-for-Service
Distribution of Patients Receiving Proton Beam Therapy by Primary Cancer Diagnosis



Final

Chart III
2013 – 2017
PEBB/UMP and Medicare/UMP & Medicaid Manage Care and Fee-for-Service
Distribution of Patients Receiving Proton Beam Therapy by Age Range by Year
N = 246

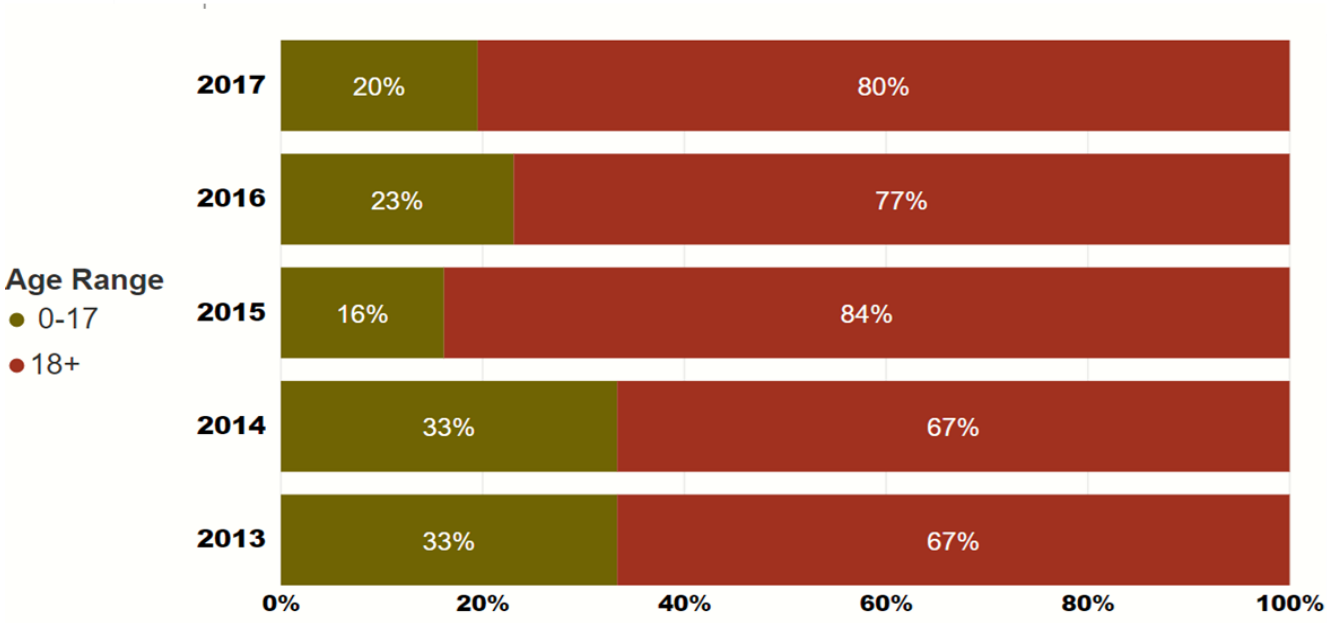


Table III
2013 – 2017
PEBB/UMP and Medicare/UMP
Utilization: Proton Beam Therapy - Outpatient
*Therapy** = Proton Treatment Delivery 77520, 77522, 77523, 77525 and Stereoscopic x-ray guidance 77424/G6002
NOTE: PEBB pays secondary to Medicare

PEBB/UMP and Medicare/UMP N = 63		2013-14	2015	2016	2017
Unique Individuals*		11	18	13	21
Patients		11	18	15	23
Average Paid Dollars/ *Therapy*	Medicare/UMP	$235	$227	$225	$220
	PEBB/UMP	$4,648	$4,683	$2,365	$2,474
Total Paid Dollars for All *Therapy*	Medicare/UMP	$39,193	$79,709	$65,528	$90,884
	PEBB/UMP	$144,095	$538,587	$208,164	$378,455

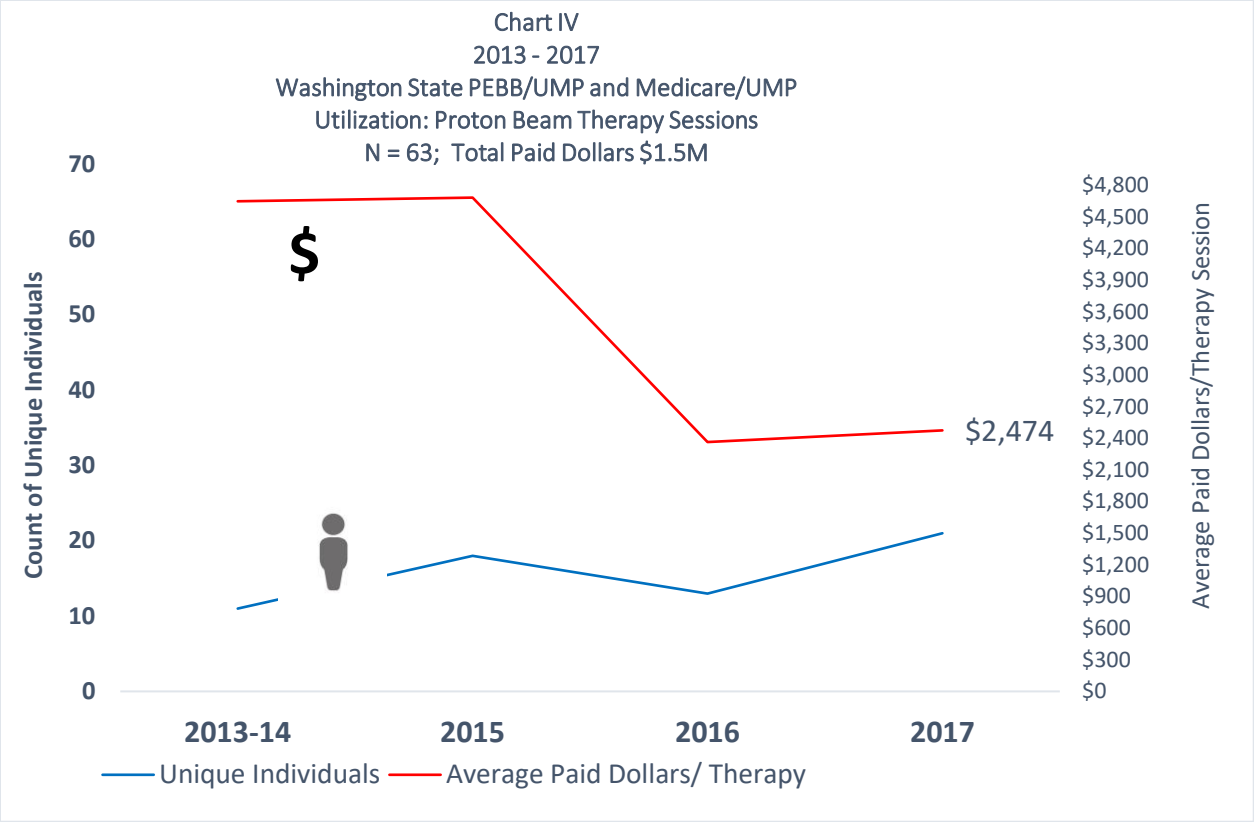


Chart V
2013 – 2017
PEBB/UMP and Medicare/UMP
Distribution of Patients Receiving Proton Beam Therapy by Primary Cancer Diagnosis

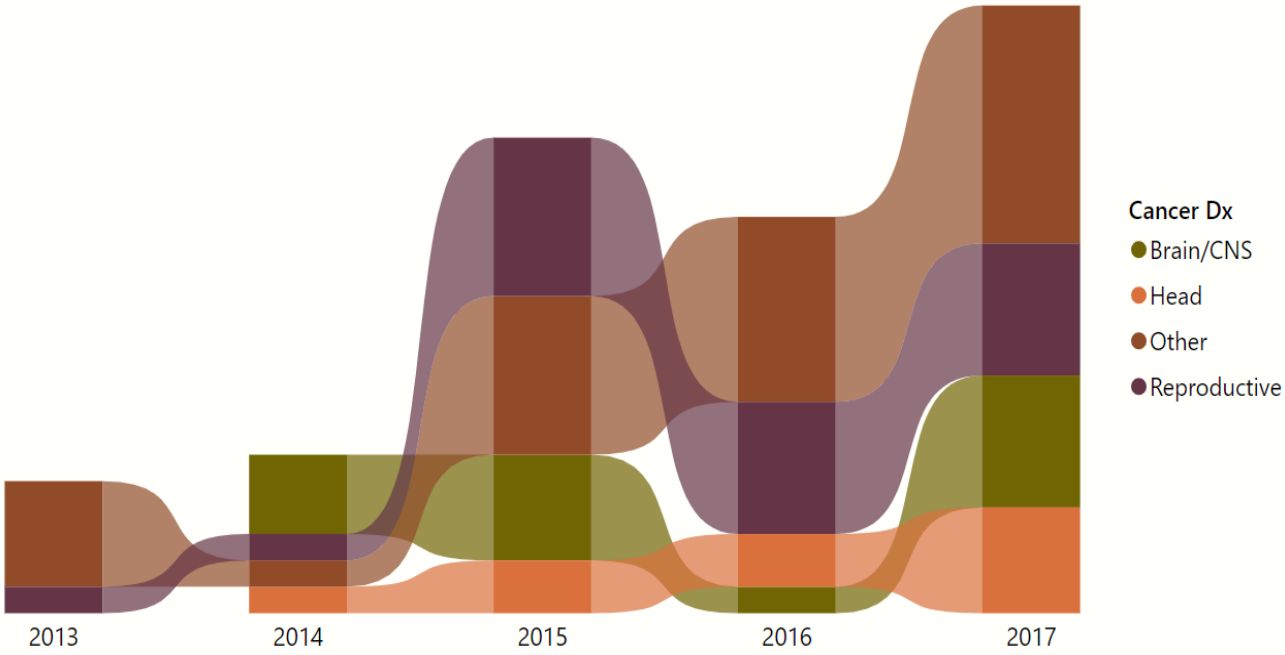


Table IV
2013 – 2017
Medicaid Managed Care and Fee-for-Service
Utilization: Proton Beam Therapy - Outpatient
*Therapy** = Proton Treatment Delivery 77520, 77522, 77523, 77525 and Stereoscopic x-ray guidance 77424/G6002

MEDICAID MCO/FFS N = 179	2013	2014	2015	2016	2017
Unique Individuals*	---	29	51	50	57
Patients	10	31	54	54	66
Average Paid Dollars[#]/ *Therapy*	$1,772	$607	$663	$667	$570
Range Paid	$607- $2,525	$607	$525-$680	$525-$680	$525-$680
Total Paid Dollars for All *Therapy*	$467,727	$401,473	$739,164	$649,384	$722,941
Total Paid Dollars Day of *Therapy*	$504,781	$522,230	$871,882	$854,998	$1,036,237

* Between 2013 and 2017, 9% (19) of patients received Therapy services paid by MCO and by FFS during the same year.

[#] "Paid dollars" uses Line Paid Amount from Claims. Patients are individuals with more than a signal category of cancer diagnoses.

--- Masked due to small numbers.

NOTE: PEBB pays secondary to Medicare

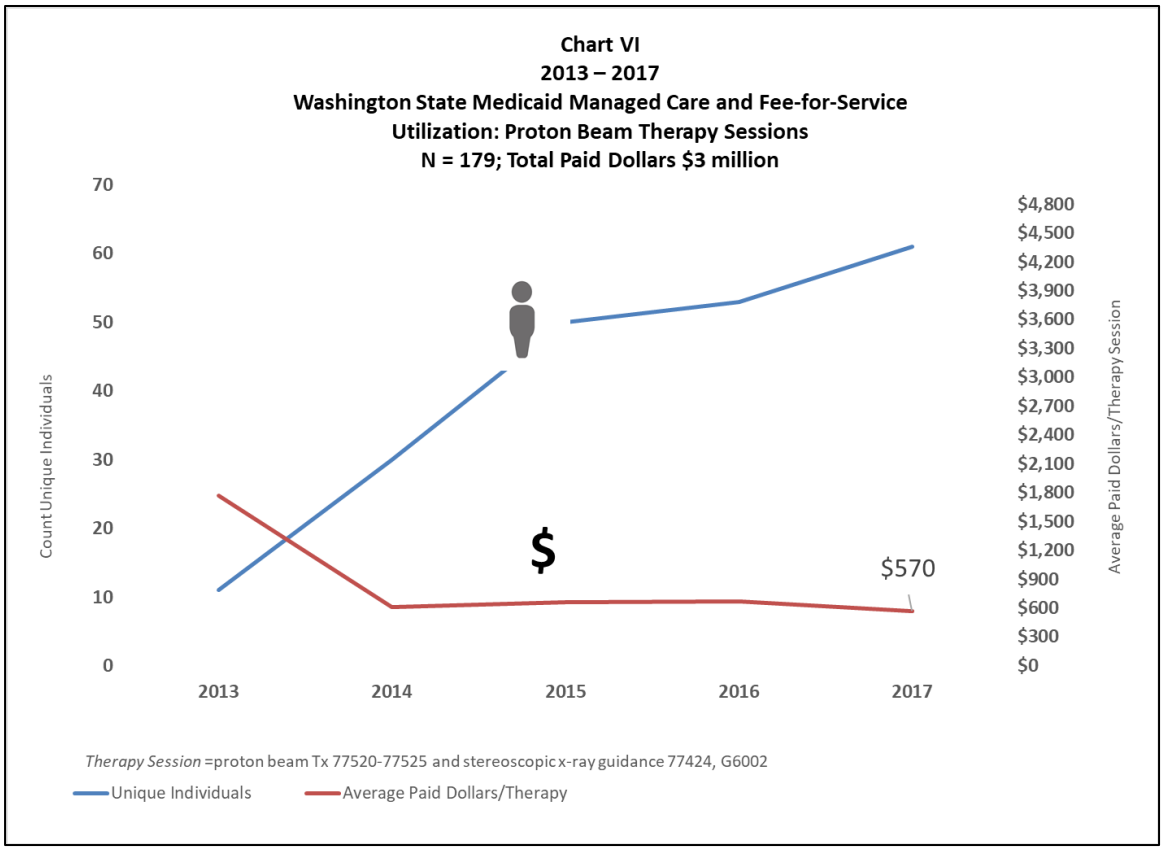
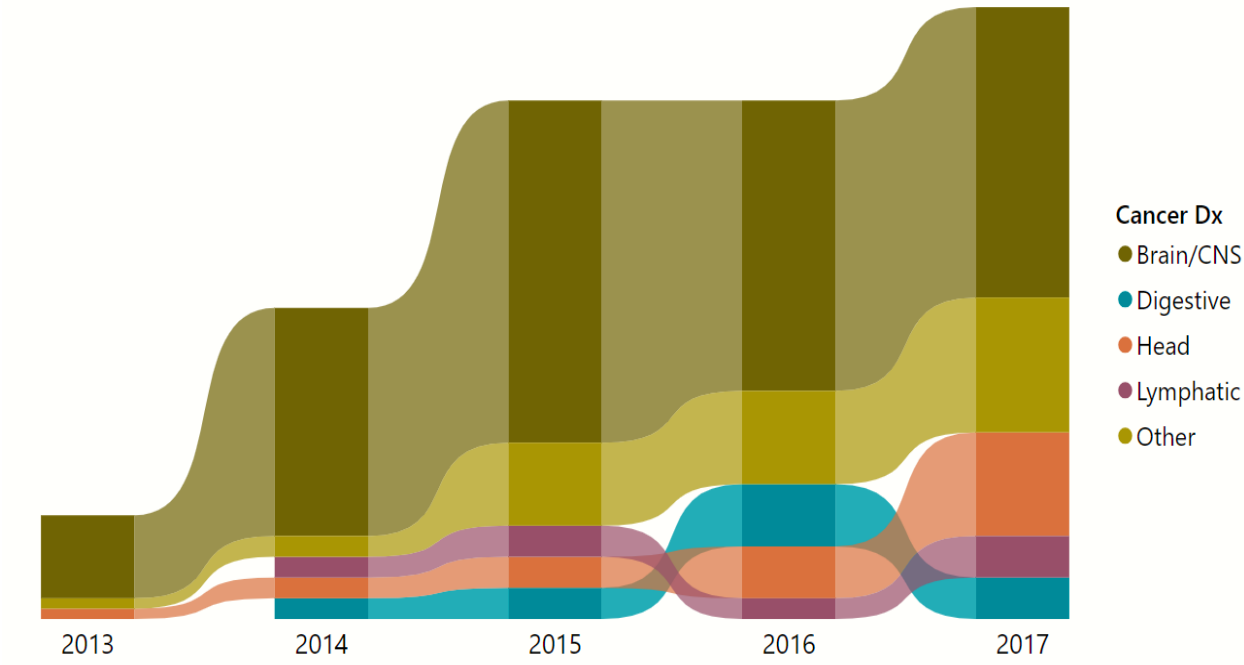


Chart VII
2013 – 2017
Medicaid Manage Care and Fee-for-Service
Distribution of Patients Receiving Proton Beam Therapy by Primary Cancer Diagnosis



Final

2 Background

2.1 Epidemiology and Burden of Disease

Overall, it is estimated that 1.7 million new cases of cancer are diagnosed yearly and cancerous conditions are responsible for over half a million deaths per year.[8] Using incidence and survival data from the Surveillance, Epidemiology, and End Result (SEER) Program and population projections from the U.S Census Bureau, the National Cancer Institute (NCI) projects the total cost of cancer care in the United States in 2020 to be $174 billion.[1] Treatment options for cancerous and noncancerous conditions vary depending on the type, location and stage of the condition and can include radiation therapy, chemotherapy, targeted therapy (e.g. inhibitor drugs), immunotherapy (including monoclonal antibodies) and surgery, or combinations of these treatments. Radiation may be delivered systemically via radioactive drugs, however, the two most common forms of radiation therapy are external beam radiation therapy (EBRT) and brachytherapy (internal radiation therapy). The focus of this review will be to evaluate the safety and effectiveness of Proton Beam Therapy (PBT), a form of external beam radiation therapy, to treat both malignant and benign tumors compared with other forms of cancer treatment.

2.2 Overview of Radiation Therapy

Radiation therapy (RT) has evolved to become one of the most powerful and commonly employed modalities for the treatment of a variety of malignancies. Today, approximately 50% of all cancer patients benefit from radiation therapy in the management of their disease and it may be the sole therapy used.[82] High-energy radiation from gamma rays, electron beams, photon beams or proton beams breaks the DNA of cancer cells, inhibiting their ability to proliferate. The radiation may also affect surrounding healthy tissues. Tumor types (and healthy tissues) vary with regard to their sensitivity to radiation. A goal of treatment planning is to damage cancer cells while minimizing damage to surrounding healthy cells including sensitive structures and organs at risk (OARs). Most often radiation is delivered using external beam radiation therapy (EBRT), a method of externally delivering radiation using a machine to aim high-energy beams directly at the tumor from outside the body. Classification of RT may be by the type of beam or particle used (i.e. electron, photon or proton) with photon RT being the most widely available and commonly used.[54] RT may be used for a variety of reasons including to cure a radiosensitive tumor, to shrink a tumor pre-operatively, to prevent recurrence or spread post-operatively (adjuvant treatment), to treat a recurrent tumor or as a palliative treatment. It may be combined with other treatments such as chemotherapy. Radiosensitive tumors for which RT may be curative include, but are not limited to, prostate cancers, head and neck cancers, and non-small cell lung cancer. RT in combination with other treatment regimens is commonly used to combat breast cancer, colon cancer, lung cancers, seminomas, and some cancers of the central nervous system, among others.

2.2.1 Potential Harms from Any Form of Radiation Therapy

Side effects of radiation therapy occur when healthy tissues in the path of the radiation beam are damaged; the effects vary from person to person. A variety of factors impact the location, type, timing and severity of side effects including the type/method of delivery and dose of radiation, the area of the body that is exposed to radiation and a person's overall health. General short-term side effects of radiation therapy may include fatigue and skin irritation (radiation dermatitis) at the radiation site. These usually subside after treatment completion. Other side effects (short and longer term) depend on

the site that was irradiated and the sensitivity of tissues surrounding the tumor. For example, short-term side effects of RT to the head or neck may include difficulty swallowing and dry mouth and later, tooth decay. Radiation to some structures may rarely cause long-term damage. For example, RT for breast cancer treatment may affect the heart. One population-based case-control study of 2168 women who underwent radiation therapy for breast cancer between 1958 and 2001 assessed the risk of major coronary events after therapy. The study found that rates of major coronary events increased by 7.4% for each increase of 1 Gray in the mean radiation dose delivered to the heart (95% CI, 2.9 to 14.5).[61] (It should be noted that the radiation therapy techniques have evolved over this time period and methods of radiation therapy delivery were not specified). Radiation is a carcinogen and rarely, secondary cancers may occur in long-term cancer survivors who have had radiation therapy; this is of particular concern in patients receiving radiation at younger ages. In addition to concerns related to the potential development of secondary cancer in those receiving radiation as children, even lower-dose irradiation of normal tissue in pediatric patients can result in pronounced acute and long-term toxicity.[279] The effects of radiation damage may be more nuanced in children, such as effects on neurocognitive development, especially when administered to children under 3 years of age.[319] The majority of white matter development takes place during early childhood and radiation is thought to have a disproportionate effect on white matter. A decrease in white matter production due to radiation exposure early in life could lead to greater risk for negative cognitive abnormalities, resulting in difficulties developing necessary skills at age appropriate rates.[228] Thus, the opportunity to limit radiation exposure to normal and developing tissues is important and is part of radiation planning.

2.2.2 Radiation Therapy Planning

In its earliest applications, RT planning used X-ray technology to take two-dimensional scans of the tumor location which were then used to determine how best to position the radiation beams in order to effectively treat the tumor. Treatments planned this way are referred to as Two-dimensional RT (2DRT) or Conventional RT (CRT). Major technological developments in computer and imaging technologies further improved upon the ability to deliver a consistent radiation dose to irregularly shaped tumors in difficult anatomic locations, while simultaneously sparing normal tissues from unnecessary radiation. Thus, 2DRT/CRT has largely been replaced by Three-dimensional Conformal Radiation Therapy (3DCRT), which uses three-dimensional imaging, such as Computed Tomography (CT) scans and Magnetic Resonance Imaging (MRI), to very accurately map the location and size of the tumor in three dimensions, as well as identify any critical OARs. Using these 3D images, beams are then matched very precisely to the shape of the tumor and delivered from all directions.[5,26]

2.2.3 Radiation Therapy Delivery

Advancements in EBRT delivery techniques have also been achieved. Initially, EBRT was limited technologically by devices that only produced low X-ray energies which therefore restricting the depth at which photons could penetrate into biological tissues. By the 1930s linear accelerators (LINACs), for delivering photons and electrons, and cyclotrons, for delivering protons and other heavy charged particles, had been developed which made possible the delivery of high-frequency accelerated particles directly to the tumor volume. This meant radiation could penetrate deeper into tissues and more precisely treat the tumor. By 1988, LINACs and cyclotrons were cleared for use by the Food and Drug Association (FDA) based on the FDA 510(k) process.[4] The first hospital-based clinic offering PBT opened two years later in 1990. There are now 27 operating PBT centers in the United States, including one in Seattle that opened in March 2013, and 5 additional centers under construction or in development.[12] It is important to note that PBT centers very expensive to construct. The cost of building cyclotrons, the

heart of proton beam facilities, ranges from $25 million to over $200 million for a multiple gantry facility.[268] Several approaches to reduce the costs of delivering PBT are being explored. One approach is the construction of compact, single-gantry proton facilities that have been estimated to reduce costs to the range of $15-$25 million. Some commentators believe that lower construction costs will mitigate the debt incurred by medical institutions and therefore lead to the ability to reduce the price charged to payers for each treatment course.[263]

Since the initial discovery of LINACs and cyclotrons, several EBRT delivery techniques (utilizing various particles) have become available today, with two of the most common being Intensity Modulated Radiation Therapy (IMRT) and Stereotactic Radiosurgery or Stereotactic Body Radiation Therapy (SBRT). IMRT is a further development of 3DCRT; it employs the same image planning and distribution techniques above but goes a step further by altering the intensity (strength) of the beams being delivered, usually lessening the intensity of the beam when near OARs. This allows for more control of the level of radiation exposure to surrounding healthy tissues while delivering a high dose to the tumor volume.[82] Initially, this technique had only been applied to photon RT but more recently similar methods have been applied to PBT as well, which is often referred to as Intensity Modulated Proton Therapy (IMPT). In this review, IMPT was a common intervention for the treatment of head and neck cancers in adults and IMRT (with photons) was the most common comparator to PBT for the treatment of brain tumors, esophageal cancers, head and neck tumors, lung cancer, and prostate cancer.

Stereotactic Radiosurgery and SBRT are similar to IMRT, however, the beams are delivered in fewer fractions (treatments) and at much higher doses than with IMRT. In addition to dose per fraction, the planning target volume margins are smaller with SBRT, requiring more rigid immobilization. Stereotactic radiosurgery, typically reserved for tumors in the brain and spine, is usually completed in a single session. (It is important to note that although the word “surgery” is utilized, no actual incisions are made during this treatment). SBRT is completed in 3 to 5 sessions and is normally used to treat larger tumors in areas of the body other than the brain.[11,82,254] These techniques are advantageous for patients who cannot tolerate surgery or have tumors in locations that are difficult to remove. Stereotactic Radiosurgery and SBRT can be delivered using photons (Brand Names: Cyber Knife, Novalis Tx, XKnife, Axesse), gamma rays (Brand Name: Gamma Knife), or protons. In the United States, these techniques are most commonly used with photons and gamma rays. More recently, the use of these techniques with protons has emerged but is only offered at a few research centers in the United States. In this review, one study compared the use of Gamma Knife to PBT for the treatment of ocular (uveal) melanoma.

In addition to IMPT and SBRT, described above, there are three additional delivery techniques specific to PBT considered during treatment planning: passive scattering, uniform scanning, and pencil beam scanning (PBS). Historically, PBT has been most commonly delivered using passive scattering techniques. This is also true for the majority of the studies included in our report.[82] Passively scattered PBT involves the use of metal apertures called collimators and wax or acrylic compensators specifically designed for each patient’s tumor. These openings are used to shape the lateral and distal aspects of individual proton beams, allowing for the beam to contour to the shape of the tumor. The second technique, uniform scanning beam, also utilizes collimators and compensators to target the tumor but goes a step further using magnets to scan a broad beam across the treatment field. Because collimators and compensators are designed specifically for an individual tumor, these radioactive apparatuses must be disposed of at the conclusion of treatment, a disadvantage to these two PBT techniques. The third and most recent technique, PBS, uses magnets to steer a small original proton beam across the lateral aspects of the tumor, thus eliminating the need for collimators and compensators. Essentially, PBS

"traces" the dimensions of the tumor allowing for the dynamic position of the beam throughout the target volume. Although more precise, treatment using PBS takes longer and there is more concern over organ motion that occurs during the treatment period which may affect where the radiation treatment occurs (i.e. healthy tissue could be targeted).[208,284] In this review, no study used PBS as the sole intervention. In two comparative studies, one in pediatric brain tumors and one in lung cancer in adults, 10% and 19% of patients, respectively, received PBS; additionally, several case-series, primarily in pediatric patients, employed PBS in a subset of their populations.

When radiation is delivered internally, as opposed to externally as with EBRT, it is called Brachytherapy. Brachytherapy treatment delivery involves a physician placing small seeds of radioactive material directly into or very close to the tumor. Brachytherapy is a common treatment for eye tumors and patients are considered good candidates if they have large, medium, or small tumors with documented growth (although visual outcomes may be compromised in patients with large tumors). Patients presenting with extensive circumpapillary/papillary extension, bulky extrascleral extension, ring melanoma, tumor involvement of more than half the ciliary body, a very large tumor, or blind and painful eyes, are considered poor candidates for Brachytherapy.[36] Two studies in this report compare the use of PBT to Brachytherapy for the treatment of Uveal Melanomas (in conjunction with trans-scleral resection in one study).

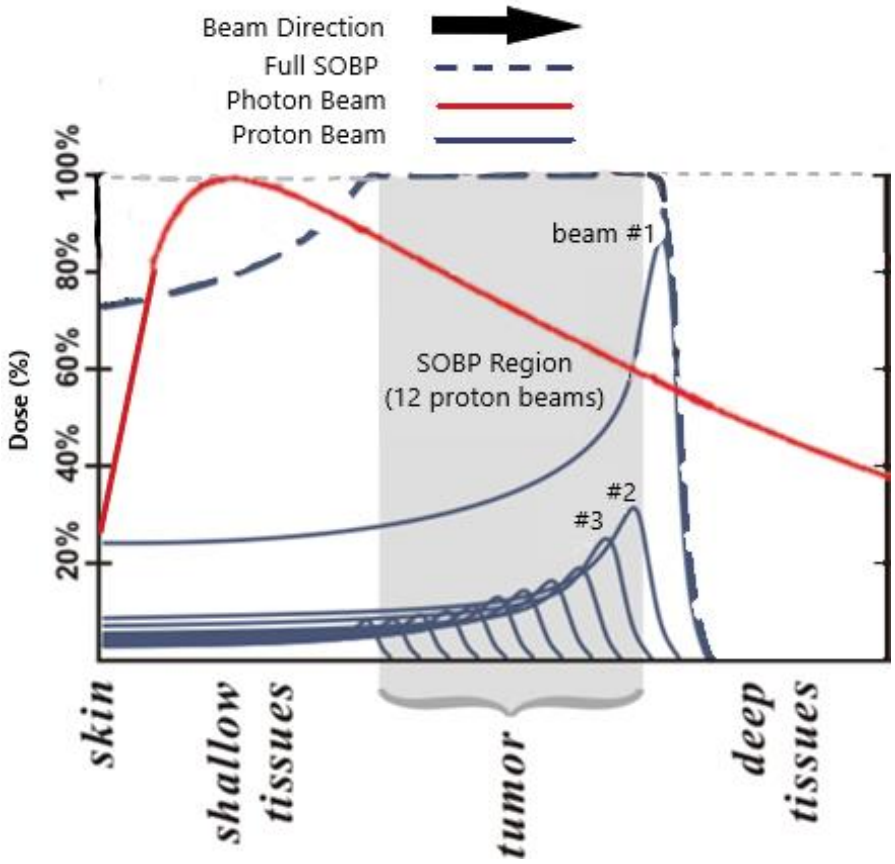
2.3 Physical Properties of Radiation Particles

With treatment planning and delivery techniques evolving similarly between varying types of EBRT, the real difference between modalities lies within the physical properties of each particle and how each reacts with tissue inside the body. Particles have different physical properties and thus their damaging effect on tissue varies.

Photons are uncharged and massless particles that reside within atoms and are characterized by a high deposit of energy near to the body surface with an exponential decrease of energy release as a function of depth.[82] As Figure 2 demonstrates, this has been a challenge for conventional photon therapy due to the amount of radiation deposited both before and after the target is reached. While the amount of photon radiation at entry into the body is much higher than at exit, photon beams typically "hit" normal tissues after leaving the target. In other words, photon beams contain an "exit dose" meaning that healthy tissue downstream from the tumor could be at an increased risk of exposure to unnecessary radiation.

This so-called "exit dose" is absent for protons, as tissue beyond the point of peak energy deposition receives little to no radiation.[146] Protons, heavy positively charged particles, can effectively treat cancerous cells at the end of their path while simultaneously lessening the damage to surrounding healthy tissues, possibly allowing for a greater dose of radiation to be delivered to the target neoplasm(s).[160] This phenomenon is referred to as the "spread out Bragg peak" (SOBP) region, which is created by varying the energy of a proton beam, creating a range of energies. For example, a shallower beam will have lower energy compared to a deeper beam (Figure 2). The large mass and acceleration applied to the protons provide each proton with a specific momentum that is mostly dispelled after traveling a defined distance. Protons are slowed down by interactions with their target which results in a sharp burst of energy deposited at the end of its path, followed by no further dose delivery ("exit dose").[280] This physical characteristic distinguishes PBT from other EBRT modalities such as photon RT. In theory, PBT offers physical advantages, though the technology is still new and more prospective clinical comparative evaluations still need to be completed.

Figure 2. Adapted from Levin WP, Kooy H, Loeffler, DeLaney TF. Proton Beam Therapy. BR J Cancer. 2005; 93(8):849-854.



Additionally, it has been previously assumed that the biological effects of protons are equivalent to that of photons, but recent studies have shown that the Relative Biological Effectiveness (RBE) of protons in relation to photons are not known with absolute certainty for all types of tissues and fractionation schemes.[215] The unit Gy is what can be measured with instrumentation, but the RBE allows clinicians to understand what is happening to the tissues on a cellular level (i.e., the biological damage occurring to cells). In dosimetry calculations, Linear Energy Transfer (LET), the average energy deposited by an ionizing particle in each unit of length, is used to calculate RBE.[85,321] For PBT treatment planning, an RBE of 1.1 is usually assumed. However, RBE is dependent on several factors such as dose per fraction, LET, tissue radio-sensitivity, particle speed, tissue type, and local microenvironments such as oxygen level.[85] One study identified situations in which RBE was found to be both larger and smaller than 1.1 and another found that ignoring possible variations in RBE could lead to suboptimal PBT treatment plans. The concern with assuming a 1.1 RBE for all tumor types treated with PBT is that it may result in treatment plans that deliver a lower biological dose to the target and a higher biological dose to the normal tissue.[87]

Further, the dose range is relatively certain for tumors that are close to the skin, but there is more uncertainty around the end of the dose range when deep-seated tumors such as prostate cancer are considered.[88] Protons are also very sensitive to tissue heterogeneity, and the precision of the beam may

be disturbed as it passes through different types of tissue.[283] Another concern is the effects of neutrons, which are produced by passively-scattered proton beams and result in additional radiation dose to the patient. The location of neutron production in a PBT patient and its biologic significance is currently a topic of significant debate.[100,127]

2.4 Comparator: Transarterial Chemoembolization

Transarterial Chemoembolization (TACE) is a common treatment for liver cancers in which the blood supply to a tumor is blocked after anticancer drugs are given in blood vessels near the tumors. To reduce side effects, the anticancer drugs are sometimes attached to small beads that are injected into an artery that feeds the tumor. This blocks the vein and reduces the level of drugs that are delivered to other parts of the body. TACE is typically indicated for patients with intermediate stage liver disease and is most often used as a means to buy time until a patient can receive a liver transplant. Patients with advanced stage liver disease typically respond worse to TACE, but TACE can be administered to these patients as a form of palliative care.[257] Serious complications from TACE occur after about one in 20 procedures. Most major complications involve either infection in the liver or damage to the liver. Approximately 1% of procedures result in death, usually due to liver failure.[3] In this review, one RCT compared PBT with TACE for the treatment of unresectable hepatocellular carcinoma.

2.5 Clinical Guidelines, Consensus Statements, & Appropriateness Criteria

The table below summarizes information from across guidelines and appropriateness criteria as well Centers for Medicare & Medicaid Services (CMS) and payer policies related to the use of proton beam therapy for a range of cancerous and non-cancerous conditions. For CMS and payer policies, specific detail was only provided for conditions that were or were not considered medically necessary. For a broader list of conditions considered “investigational” by CMS or representative bellwether payer policies, please refer to Table 6.

Table 2. Summary of proton beam therapy recommendations by cancer type across guidelines, appropriateness criteria, CMS coverage, and payer policies

Guideline & Appropriateness Criteria				CMS and Payer Policies
Condition	**Recommendation**	**Strength of Recommendation**	**Evidence Quality**	**Coverage**
Bone Cancer [202,229]	NCCN: M ACR*: N	NCCN: Moderate ACR*: NR	NCCN: 2A ACR*: NR	Investigational or NR
Brain, Spinal, Paraspinal Cancer[74,105,202]	NCCN: M (CNS cancers) NICE: Y AIM: Y (CNS tumors, chordomas, chondrosarcoma)	NCCN: Moderate NICE: NR AIM: NR	NCCN: 2A NICE: NR AIM: NR	LCDs† CMS[7,9,10]: Y (unresectable, pituitary, chordomas, chondrosarcomas) Payer Policies Aetna: Y (chordomas/chondrosarcomas of skull, cervical spine; pituitary, Intracranial arteriovenous malformation ; CNS)
Breast Cancer[105]	AIM: N	AIM: NR	AIM: NR	Investigational or NR
Esophageal Cancer[105]	AIM: N	AIM: NR	AIM: NR	Investigational or NR
Gastrointestinal Cancer[105]	AIM: N AIM: N (pancreatic)	AIM: NR	AIM: NR	Investigational or NR
Gynecologic Cancer [105,229]	AIM: N ACR*: N	AIM: NR ACR*: NR	AIM: NR ACR*: NR	Investigational or NR
Head & Neck Cancer [105,202,229]	NCCN: M AIM: N ACR*: Y	NCCN: Moderate AIM: NR ACR*: NR	NCCN: 2A AIM: NR ACR*: NR	LCDs† CMS[7,9,10]: Y (advanced/unresectable; paranasal/sinus)
Liver Cancer [105,202]	NCCN: M AIM: N	NCCN: Moderate AIM: NR	NCCN: 2A AIM: NR	Investigational or NR
Lung Cancer [74,105,144,202]	ASCO: Y (pleural mesothelioma) NCCN: M (pleural mesothelioma & NSCLC) AIM: N ACR*: N	ASCO: Strong NCCN: Moderate AIM: NR ACR*: NR	ASCO: Intermediate NCCN: 2A AIM: NR ACR*: NR	Investigational or NR
Lymphomas [105,202,229]	NCCN: M AIM: N ACR: M	NCCN: Moderate AIM: NR	NCCN: 2A AIM: NR	Investigational or NR

Guideline & Appropriateness Criteria				CMS and Payer Policies
Condition	**Recommendation**	**Strength of Recommendation**	**Evidence Quality**	**Coverage**
Ocular Cancers [105,202]	NCCN: M (uveal melanoma) AIM: Y	NCCN: Moderate AIM: NR	NCCN: 2A AIM: NR	LCDs† CMS[7,9,10]: Y Payer Policies Aetna: Y (uveal) Anthem: Y (uveal) Anthem: N (choroidal neovascularization secondary to age-related macular degeneration)
Pediatric Cancers [74,229]	NICE: Y AIM: Y	NICE: NR AIM: NR	NICE: Not sufficient	LCDs† CMS[7,9,10]: Y Payer Policies Aetna: Y Anthem: Y
Prostate Cancer [74,105,202,211,229]	ASTRO: N NCCN: N NICE: N AIM: N ACR*: M	ASTRO: Moderate NCCN: Moderate AIM: NR ACR: NR	ASTRO: Grade C NCCN: 2A AIM: NR ACR*: NR	Aetna: N
Sarcomas [202]	NCCN: M	NCCN: Moderate	NCCN: 2A	LCDs† CMS: Y (unresectable retroperitoneal sarcoma)
Seminomas	NR	NR	NR	Investigational or NR
Thymomas [202]	NCCN: M	NCCN: Moderate	NCCN: 2A	Investigational or NR

ACR = American College of Radiology; AIM = American Imaging Management; ASTRO = American Society for Radiation Oncology; CMS = Centers for Medicare and Medicaid Services; CNS = central nervous system; LCD = local coverage determination; NCCN = National Cancer Care Network; NICE = The National Institute for Health and Care Excellence; NR = not reported; Y = Yes.

*ACR ratings are associated with N, M, and Y ratings based on their 1-9 rating system; in this table N = 1, 2, 3 (usually not appropriate); M = 4, 5, 6 (may be appropriate); and Y = 7, 8, 9 (usually appropriate). For more information on their rating system see Appendix Table L2.

†At the time of this report the only CMS policy related to proton beam therapy and applied to Washington State had been retired as of Sept. 2017; two LCDs active in twelve states (not including Washington State) are active however, with only minor differences in coverage determinations. Information on the coverage decisions are reported here for reference, more detail is available in section 2.7, Table 6.

2.6 Previous Health Technology Assessments & Systematic Reviews

2.6.1 Summary of Previous HTAs of Proton Beam Therapy in Adults and Pediatrics

A total of six Health Technology Assessments (HTAs) [43,66,106,141,166,302] were identified regarding the comparative effectiveness, safety, and/or economic value of PBT for the treatment of tumors compared to other various types of treatments. One identified HTA[302] did not include any new SRs or studies relevant to PBT that were published subsequent to the search dates of the previous report and another HTA[106] only cited the 2014 WA State HTA. For the reasons stated, these HTAs have not been summarized in the table below. Across the 4 summarized HTAs containing newly published studies and/or systematic reviews, 2 HTAs [43,141] included analysis in adults and pediatrics and 2 HTAs[66,166] included analysis in adults only.

Table 3. Previous Health Technology Assessments of PBT in adult and pediatric populations

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
CDATH (2016)[43] Rapid Response Report *Proton Beam Therapy versus Photon Radiotherapy for Adult and Pediatric Patients: A Review of the Clinical and Cost-Effectiveness*	**Cancer Type(s):** Adult and pediatric oncology patients requiring radiation therapy for various cancers **Treatments Evaluated**: PBT vs. photon radiotherapy **Key Qs:** 1. What is the clinical effectiveness of proton beam therapy for the treatment of cancer patients? 2. What is the cost-effectiveness of proton beam therapy for the treatment of cancer patients?	**Evidence Base Available** SRs of clinical evidence: n=2 (2 new) SRs of economic evidence: n=2 (2 new) Primary Economic Evaluation: n=1 (1 new) Critical Appraisal (RoB) = AMSTAR (SRs), Drummond strengths and limitations tool (economic evaluations) SOE: Yes **Search Dates:** January 1, 2013 to April 22, 2016	**Clinical and Safety outcomes** **Adults and Pediatrics – General Conclusions** Across Various Cancers (craniopharyngioma or retinoblastoma in children; breast cancer, head and neck cancer, uveal hemangioma, NSCLC, meningioma in adults) 2 SRs with 20 comparative studies - Most evidence was deemed low-strength or insufficient, meaning it is unlikely to allow for any definitive conclusions. **Adults – Detailed Conclusions by Cancer Type and Treatment** Breast 1 non-randomized study - PBT vs. 3DCRT: comparable 7-year cumulative recurrence rates, incidence of fat necrosis, moderate to severe fibrosis, 7-year moderate to severe breast pain, and cosmetic outcomes; PBT had higher rates of 7-year skin toxicities Medulloblastoma 1 "low-quality" retrospective non-randomized study - PBT vs. IMRT: Comparable OS and PFS; Lower acute toxicity and AEs rates in PBT group Spinal Cord Glioma 1 retrospective non-randomized studies - PBT vs. IMRT: No difference in the local recurrence or death rate at 1-year; PBT patients more likely to progress to death within 5-years

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			Esophageal 2 non-randomized studies - PBT vs IMRT: No difference in odds of post-operative pulmonary complications - PBT vs. 3DCRT: Reduced odds of post-operative pulmonary complications in 3DCRT group - Increased rates of acute pneumonitis in the PBT group vs. either IMRT or 3D-CRT - No difference in odds of GI complications between either PBT, IMRT, 3DCRT Meningioma 1 small "poor-quality" retrospective non-randomized study - Results deemed unreliable Uveal Hemangioma 1 "poor-quality" retrospective non-randomized study - PBT vs. Photon RT: No difference in stabilization of visual acuity, optic disc or nerve atrophy, retinopathy or grade 3 or 4 side effects but, potential confounding due to baseline imbalances and the limited scope of the evidence Head and Neck 1 small (n=6) "poor-quality" retrospective non-randomized study - Insufficient evidence to support use of PBT in patients with head and neck cancers Lung 2 historically controlled non-randomized studies - No difference in OS between PBT and IMRT or 3DCRT - PBT superior to IMRT in terms of acute esophagitis at 6 months and grade 3 esophagitis at 15 to 17 months - Similar rates of grade 3 pneumonitis, grade 3 dermatitis and grade 3 fatigue between PBT and IMRT - PBT superior to 3DCRT for rates of grade 3 esophagitis at 15 to 17 months, and grade 3 pneumonitis - PBT higher rates of grade 3 dermatitis compared to 3DCRT - PBT vs. 3DCRT: Similar rates of acute esophagitis at 6 months and grade 3 fatigue

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			Prostate 1 "poor-quality case series, several historically controlled and retrospective studies (n=unclear) - Role of PBT in the context of other available therapies for prostate cancer remains unclear Various Cancers – Risk of Secondary Malignancies 1 retrospective study - Findings were deemed to be unreliable Recurrent Cancers 2 retrospective non-randomized studies - Insufficient evidence to allow conclusions regarding the comparative effectiveness of PBT and CRT, but studies reported similar outcomes for both groups **Pediatrics – Detailed Conclusions by Cancer Type and Treatment** 2 retrospective non-randomized studies from 1 SR Craniopharyngioma - PBT vs. IMRT: No significant difference in the risk of secondary malignancy, but significantly lower rates of radiation therapy induced or in-field secondary malignancies in the PBT group Retinoblastoma - PBT vs. CRT: No significant difference in risk of secondary malignancy; PBT group had significantly lower rates of radiation therapy induced or in-field secondary malignancies compared to CRT group **Economic Outcomes** Evaluations are limited by the absence of high quality, long-term clinical evidence; evaluations largely rely on modeled outcomes from case series and effectiveness of PBT may be over-estimated making it appear more cost effective than it is. **Adults – Detailed Conclusions by Cancer Type and Treatment** Prostate 4 evaluations - PBT was not cost-effective vs. photon modalities in older men with prostate cancer

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			Breast 3 evaluations - PBT was not cost-effective vs. whole breast irradiation and CRT in breast cancer patients of various ages without cardiac risk factors - PBT more likely to be cost-effective in women with cardiac risk factors and younger patients (aged 40 or 50, vs. 60) Lung 1 evaluation - PBT not considered cost-effective for NSCLC Head and Neck 2 evaluations - PBT not considered cost-effective compared to CRT or IMRT **Pediatrics – Detailed Conclusions by Cancer Type and Treatment** Medulloblastoma 4 evaluations • - Across all 4 studies, PBT was determined to be cost-effective, but there were methodological issues with some evaluations Brain 1 evaluation - PBT considered to be cost-effective over a broad range of costs
CDATH (2017) [141] *Proton Beam Therapy for the Treatment of Cancer in Children and Adults: A Health Technology Assessment*	**Cancer Type(s):** Adults and children with various cancers **Treatments Evaluated:** PBT vs. 3DCRT PBT vs. IMRT PBT vs. photon RT PBT vs. SRT PBT vs. carbon ion RT PBT vs. helium ion RT PBT plus photon RT vs. photon RT	**Evidence Base Available:** Clinical SRs: N=9 SRs, 11 publications (6 new SRs with 4 new studies across 5 publications) Critical Appraisal (RoB) = ROBIS tool (SRs), ROBINS-I tool (non-randomized studies), CASP (economic studies)	9 SRs with 34 unique primary studies (mostly low-quality evidence from poor-quality primary studies) Across all four domains of ROBIS, 3 SRs = low-level of concern and 6 SRs = mixed-level of concern (only 2 SRs received a high or moderate level of concern in more than one of the four domains) All but two of the SRs included in the report assessed the quality or risk of bias of their included primary studies. **Clinical Outcomes** **Adults – General Conclusions** - Clinical effectiveness of PBT, alone or in combination with photon RT, was similar to other types of RT in giant-cell bone

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
	PBT vs. brachytherapy PBT plus photon RT vs. brachytherapy **Key Qs** <u>Clinical</u> 1. What are the clinical benefits of PBT compared with other types of radiotherapy for the treatment of cancer in children and adults? 2. What are the clinical harms of PBT compared with other types of radiotherapy for the treatment of cancer in children and adults? <u>Economic Analysis</u> 3. What is the budget impact of installing a Canadian-based PBT facility as an alternative to out-of-country referrals for PBT for the treatment of cancer in children and adults? <u>Patient Experiences and Perspectives</u> 4. What are the experiences and perspectives of adults and children diagnosed with cancer and their family members and caregivers related to travelling to receive cancer treatment? <u>Ethical Issues</u> 5. What ethical issues are known in cancer treatment, and how might the availability of PBT influence these issues? 6. What new ethical issues are raised by the use of PBT in cancer treatment? In particular, what issues are raised by the need to travel out of country for treatment?	SOE: Yes **Search Dates:** January 1, 2007 to June 30, 2017	tumors, breast cancer, medulloblastoma, esophageal cancer, liver cancer, lung cancer, and most prostate cancer - PBT, alone or in combination with photon RT, was associated with greater benefits in meningioma and both greater benefits as well as lower benefits in eye cancer - Evidence from subgroup analyses suggests that the effect of PBT may be greater in malignant meningioma and poorly-differentiated tumors of prostate cancer, although it is unclear whether these findings from subgroup analyses are clinically meaningful **Pediatrics – General Conclusions** - Clinical effectiveness and safety of PBT were similar to IMRT in craniopharyngioma (i.e., 3-year OS or DFS). PBT, compared with photon RT, was associated with lower harms in retinoblastoma (i.e., lower 10-year RT-induced or in-field secondary malignancy). **Adults and Pediatrics – Detailed Conclusions by Cancer Type and Treatment** <u>Bone Cancer</u> 2 SRs with 1 unique "poor-quality" study - PBT plus photon RT vs. photon RT alone and with or without partial tumor resection: No significant differences in distant metastases or PFS after 20-year follow-up or 9-year median follow-up <u>Breast</u> 1 SR with 1 unique "fair-quality" and "low SOE" study - PBT vs. 3DCRT: No significant difference in 7-year cumulative local recurrences <u>Central Nervous System Cancers</u> 3 SRs with 6 unique primary studies (most of which were "low-quality") - PBT vs. IMRT: Very low-quality evidence indicated no statistically significant differences in 3-year OS or DFS

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
	7. If PBT therapy is installed and implemented more widely in Canada, how should it be provided to best address the identified issues? <u>Implementation Issues</u> 8. What are the main challenges and enablers to implementing PBT in Canada?		- PBT vs. photon RT: Low-strength evidence indicated no significant differences in 2-year or 5-year locoregional failure, and 2-year OS or PFS - PBT plus photon RT vs. photon RT alone: PBT plus photon RT was associated with statistically significantly higher 5-year local control (however, insufficient evidence to make a definitive conclusion about the benefits of PBT) - PBT vs. IMRT: low-strength evidence indicated no statistically significant differences in local recurrences or metastases after 24 months of median follow-up; PBT was associated with statistically significantly lower chances of 5-year survival or higher mortality rates after 24 months of median follow-up in multivariate analyses controlling for age, tumor pathology, and treatment modality <u>Esophageal Cancer</u> 1 SR with 2 unique primary studies - PBT vs. IMRT or 3DCRT: No significant differences in 90-day mortality, 3-year OS (over an unknown duration), or disease-specific survival over an unknown duration. <u>Eye Cancer</u> 2 SRs with 2 unique primary studies - PBT vs. brachytherapy: PBT had a statistically significantly lower rate of local recurrence but a higher mortality rate - No 5-year or 15-year late recurrences after PBT vs. helium ion RT, but some late recurrences with iodine-125 brachytherapy. <u>Liver Cancer</u> 3 SRs with 3 unique primary studies of low-, unknown-, and poor-quality - PBT vs. carbon ion RT: Similar 5-year local control and 5-year OS - PBT vs. photon RT: Similar rate of tumor recurrences of unknown duration and mortality after 1.5 years of follow-up, but did not always provide statistical testing results. (Based on a small (n=8) poor-quality study) <u>Lung Cancer</u>

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			2 SRs with 2 unique primary studies - PBT vs. 3DCRT vs. IMRT: No significant differences in median survival times - PBT vs. carbon ion RT: no statistically significant differences in three-year local control, overall survival, or progression-free survival <u>Prostate Cancer</u> 4 SRs with 8 unique primary studies - PBT vs. 3DCRT: Clinically and statistically significant decreases in bowel, but not urinary, QoL from baseline to 24 months of follow-up for both treatments - PBT vs. IMRT: no statistically significant differences in 2-year bowel, urinary, or sexual QoL or 4-year QoL associated with urinary incontinence or erectile dysfunction diagnoses - PBT vs. photon RT: no significant differences in 18-month QoL - PBT plus photon RT vs. brachytherapy: no statistically significant differences in any of the examined outcomes after 5 or 8 years of follow-up; no significant differences in 8-year distant metastases based on tumor stages (subgroup analysis) - PBT plus photon RT vs. photon RT alone: No significant differences in QoL after 18 months follow-up or tumor control, cancer control, or survival after 5 or 8 years follow-up; statistically significantly greater 8-year local control in poorly-differentiated tumors with PBT plus photon RT (subgroup analysis) **<u>Safety Outcomes</u>** **Adults – General Conclusions** - PBT, alone or in combination with photon RT, was… • similar to other types of RT in most breast cancer, meningioma, some esophageal cancer, choroidal or uveal hemangioma, head and neck cancer, some lung cancer, and some prostate cancer • associated with greater harms in some breast cancer and some prostate cancer

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			• associated with lower harms in medulloblastoma • associated with greater harms as well as lower harms in in optic nerve sheath meningioma, some lung cancer, and some esophageal cancer **Pediatrics – General Conclusions** - Clinical effectiveness and safety of PBT were similar to IMRT in craniopharyngioma (i.e., 3-year OS or DFS). PBT, compared with photon RT, was associated with lower harms in retinoblastoma (i.e., lower 10-year RT-induced or in-field secondary malignancy). **Adults and Pediatrics – Detailed Conclusions by Cancer Type and Treatment** <u>Breast</u> 1 SR with 1 unique "fair-quality" and "low SOE" study - PBT vs. 3DCRT: Statistically significant higher risk of 7-year skin toxicity associated with PBT - PBT vs. 3DCRT: No statistically significant differences in the occurrences of fat necrosis over an unknown duration, moderate/severe fibrosis over an unknown duration, 7-year moderate or severe breast pain, or 5-year rib fracture <u>Central Nervous System Cancers</u> 3 SRs with 6 unique primary studies (most of which were "low-quality") - PBT vs. IMRT: Very low-quality evidence indicated no statistically significant differences in vascular injury, visual dysfunction, hypothalamic obesity, panhypopituitarism, and other endocrinopathies after nine months to 185 months of follow-up - PBT vs. photon RT: Low-strength evidence indicated that PBT was associated with statistically significantly lower risk of 1-month acute toxicity, including weight loss, esophagitis, and nausea or vomiting - PBT vs. photon RT: No significant differences in side effects after 12 months to 42 months of follow-up (authors deemed evidence as insufficient)

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			- PBT vs. IMRT: low-strength evidence indicated that neither treatment was associated with any long-term toxicity or myelopathy <u>Esophageal Cancer</u> 3 SRs with 4 unique primary studies - PBT vs. IMRT: No statistically significant differences in 30-day gastrointestinal and pulmonary post-operative complications or esophagitis, pneumonitis, and dermatitis rates - PBT vs. 3DCRT: PBT was associated with statistically significantly lower risk of 30-day pulmonary post-operative complications - PBT vs. 3DCRT or IMRT (analyzed together): PBT was associated with statistically significantly higher risk of acute pneumonitis <u>Eye Cancer</u> 3 SRs with 3 unique primary studies - PBT vs. RT: Very low-quality evidence indicated that PBT was associated with statistically significantly lower 10-year RT-induced or in-field secondary malignancy - PBT vs. photon therapy: Differences in grade 1 or 2 side effects in optic or disc nerve atrophy, ocular pressure, effect on retina, and retinopathy after 28 months of median follow-up. After adjusting for between-group differences, no statistically significant effects on optic or disc nerve atrophy or retinopathy were observed. No statistically significant differences in grade 3, 4, or 5 side effects in lacrimation, lens, or retinopathy after 28 months of median follow-up. (Evidence for comparative effectiveness of PBT compared to photon therapy deemed insufficient) - PBT alone or PBT plus photon RT vs. photon RT alone: PBT alone or PBT plus photon RT was associated with lower rates of acute orbital pain or headache but higher rates of late asymptomatic retinopathy. (Based on a “poor-quality” study) <u>Head and Neck Cancers</u> 1 SR with 1 unique primary study

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			- PBT vs. carbon ion RT: Similar in both unadjusted and adjusted rates of vision loss over unknown duration. However, statistical testing results were not always provided <u>Liver Cancer</u> 1 SR with 2 unique primary studies - PBT vs. carbon ion RT: Similar rates of grade 2 dermatitis, increased transaminase, rib fracture, nausea, anorexia, pain, or ascites, grade ≥ 3 late toxicity, and deaths related to treatment-related toxicity after 31 months of median follow-up, but did not provide statistical testing results - PBT vs. photon RT: No occurrences of bone marrow depression or gastrointestinal complications over unknown duration with PBT or photon RT, but concluded there was insufficient evidence to make a definitive conclusion about the benefits of PBT, compared with X-rays <u>Lung Cancer</u> 2 SRs with 3 unique primary studies - PBT vs. 3DCRT vs. IMRT: No statistically significant differences in hematologic toxicity after 1.8 months to 76.1 months of follow-up - Risk of six-month acute severe esophagitis was similar between PBT and 3DCRT but statistically significantly lower with PBT compared with IMRT - occurrences of 15- to 17-month esophagitis and grade ≥ 3 pneumonitis after 1.8 months to 76.1 months of follow-up were the lowest with PBT compared with 3DCRT or IMRT; grade ≥ 3 dermatitis after 1.8 months to 76.1 months of follow-up was the highest with PBT compared with 3DCRT or IMRT; and grade ≥ 3 fatigue after 1.8 months to 76.1 months of follow-up was similar among the three modalities - PBT vs. carbon ion RT: no statistically significant differences in the rates of dermatitis, pneumonitis, and rib fracture after 3.5 years of median follow-up <u>Prostate Cancer</u> 5 SRs with 7 unique primary studies

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			- PBT vs. 3DCRT: 1-year adjusted gastrointestinal toxicity rate was significantly higher with PBT (low-strength evidence) - PBT vs. IMRT: 4 SRs that reported on toxicity were in disagreement regarding the toxicity of PBT, some reporting a statistically significant difference in gastrointestinal toxicities and some reporting no statistically significant difference - PBT plus photon RT vs. photon RT: No significant differences in 18-month gastrointestinal, sexual, or urinary toxicity, 5-year rectal or urinary toxicity, or eight-year gross hematuria between the two interventions; 8-year rates of rectal bleeding and urethral stricture statistically significantly higher with PBT plus photon RT - PBT vs. brachytherapy: higher rates of gastrointestinal toxicity, including bleeding, over unknown duration with PBT (no between group comparison or statistical testing)
VA (2015)[66] *Comparative Effectiveness of Proton Irradiation Treatment*	**Cancer Type(s):** Adults only with various cancers **Treatments Evaluated:** Single field PBT vs. 3DCRT PBT vs. IMRT PBT vs. photon RT PBT vs. IMRT PBT vs. IMRT or 3DCRT PBT plus photon RT vs. photon RT PBT plus photon RT vs. brachytherapy PBT plus photon RT vs. various photon RT modalities PBT vs. various photon RT modalities **Key Qs:** 1. What is the effectiveness of proton beam irradiation compared to conventional X-ray-based external beam modalities?	**Evidence Base Available** N=51 studies n=25 comparative studies (2 new), 6 SRs (1 new), 20 non-comparative Critical Appraisal (RoB) = Cochrane Collaboration's Risk of Bias Tool (RCTs), methods from the Drug Effectiveness Review Project (observational studies), AMSTAR (SRs) SOE: Yes **Search Dates:** NR to December 10, 2014	**<u>Clinical and Safety Outcomes</u>** **Adults – General Conclusions** - Comparative studies have not demonstrated any common clinical situations in which proton beam therapy has an important clinical advantage over photon radiotherapy modalities on meaningful *long-term* health outcomes - Low-strength evidence of the potential of PBT for increased late toxicity compared with IMRT and 3D-CRT for breast, esophageal, prostate, and spinal cord glioma cancers was uncovered **Adults – Detailed Conclusions by Cancer Type** 6 SRs and 25 primary comparative studies <u>Breast</u> 1 fair-quality prospective study (Low SOE) - PBT vs. photon-based 3D conformal accelerated partial-breast irradiation: Comparable 7-year cumulative local recurrence; 7-year skin toxicities were more common in PBT; No difference in patients' ratings of good or excellent for 7-year overall cosmetic outcomes or in local failure rates <u>Esophageal</u>

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
	2. What is the effectiveness of proton beam irradiation compared to state-of-the-art therapies? 3. In patients with local recurrences after irradiation, what is the effectiveness of proton beam irradiation compared to conventional X-ray-based external beam modalities and state-of-the-art therapies? 4. What are the short- and long-term harms of proton beam irradiation compared to conventional X-ray-based external beam modalities and state-of-the-art therapies? 4A. What are the harms of proton beam irradiation compared to photon-based therapies in treating mobile targets that may move during treatment?		(Low SOE) - PBT vs. IMRT: Comparable risk of postoperative pulmonary complications and GI complications - PBT vs. 3DCRT: Comparable risk of GI complications; 3DCRT had a higher risk of pulmonary complications - PBT vs. IMRT/3DCRT: PBT is associated with a higher risk of acute pneumonitis <u>Medulloblastoma</u> 1 retrospective cohort study (Low SOE) - PBT vs. photon therapy: Comparable 2-year OS and PFS, proportion of patients with treatment breaks, and locoregional failure; Some 1-month toxicities were less common in PBT, including medical management of esophagitis, weight loss, and Grade ≥ 2 nausea/vomiting <u>NSCLC</u> 1 comparative cohort study (Low SOE) - PBT vs. 3DCRT vs. IMRT: Similar acute risk of severe esophagitis for 3DCRT and PBT, but lower than IMRT - Insufficient evidence to draw conclusions about proton-based stereotactic ablative therapy for early-stage lung cancer compared with photon-based stereotactic ablative therapy <u>Prostate</u> 1 RCT, 9 cohort studies (all low SOE unless otherwise specified) - PBT vs. IMRT: Similar QoL; Transiently lower Genitourinary toxicity at 0-6 months for PBT, similar GI and Genitourinary toxicity at 12-24 months (low to moderate SOE), increased GI toxicity with PBT at 4- 5 years - PBT vs. 3DCRT: Similar QoL (insufficient SOE); Increased acute GI toxicity with PBT - PBT vs. brachytherapy: Similar 8-year survival and distant metastasis - PBT plus photon RT vs. photon RT alone: Similar overall 5 to 8 year survival and QoL; Increased 8-year rectal bleeding and urethral stricture

Assessment (year) *Title*	Cancer Types, Treatments Evaluated, and Key Questions	Evidence Base Available and Search Dates	Primary Conclusions
			Spinal Cord Glioma 1 retrospective cohort (Low SOE) - PBT vs. Photon RT: PBT demonstrated a reduced chance of 5-year OS Mixed cancer types – secondary malignancy - There is insufficient evidence to draw conclusions about how PBT compares to other radiation modalities in the risk of secondary malignancy
Ludwig Boltzmann Institute (2018)[166] *Stereotactic radiotherapy, proton therapy and irreversible electroporation for the treatment of localized prostate cancer*	**Cancer Types:** Prostate Cancer **Treatments Evaluated:** PBT, Irreversible electroporation, SBRT **Key Qs:** 1. Are irreversible electroporation, stereotactic radiotherapy and proton therapy more effective and safer for the treatment of localized prostate cancer – in terms of predefined outcome parameters – in comparison with other treatment options for prostate cancer?	**Evidence Base Available:** (PBT only) 5 RCTs, 12 prospective studies Critical Appraisal (RoB) = GRADE SOE: Yes **Search Dates:** Unclear	**Clinical and Safety Outcomes** **Adults – General Conclusions** Prostate SOE considered to be moderate to low for clinical outcomes and low to very low for safety outcomes - There is generally no evidence to suggest that PBT confers any advantage regarding QoL, in terms of urinary and gastrointestinal symptoms - Hypofractionated PBT vs. Standard PBT: Hypofractionated PBT resulted in a statistically significantly worse result for sexual functioning - PBT vs. SBRT: Similar frequencies of toxicities; late GI toxicity after PBT occurred frequently

3DCRT: Three-dimensional conformal radiotherapy; AE: Adverse event; AMSTAR: A Measurement Tool to Assess Systematic Reviews; CASP: Critical Appraisal Skills Program; CRT: Conformal radiotherapy; DFS: Disease Free Survival; GI: Gastrointestinal; GRADE: Grading of Recommendations Assessment, Development, and Evaluation; IMRT: Intensity-modulated radiation therapy; NSCLC: Non-small cell lung cancer; OS: Overall survival; PBT: Proton Beam Therapy; PFS: Progression free survival; QoL: Quality of Life; Qs: Questions; RoB: Risk of Bias; ROBINS-I: Risk of Bias in Non-randomized studies of Interventions; ROBIS: Risk of Bias in Systematic Reviews; RT: Radiotherapy; SOE: Summary of Evidence; SR: Systematic Review

2.6.2 Previous Systematic Reviews of Proton Beam Therapy in Adult Populations

Eleven systematic reviews (SRs) evaluating the effectiveness and/or safety of PBT in adult populations that included studies published subsequent to the prior report's search dates were identified via the search strategy and hand searching (Table 4). Studies contained in these reviews that met inclusion criteria for this HTA were included (excluding those published during the span of the search dates reviewed and considered by the previous report). 1 SR in Head and Neck Cancers(including skull-base), [220] 1 SR in Lung Cancer, [47] 2 SRs in Breast Cancer, [133,296] 2 SRs in Brain, Spinal, and Paraspinal Cancers, [222,328] 1 SR in Ocular Tumors, [294] and 4 SRs in Mixed Cancer and Tumor Types [68,293,295,297] were identified.

Table 4. Summary of Previous Systematic Reviews of PBT in Adult Populations

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available	Primary Conclusions
Head and Neck Cancer						
Patel (2014a)[220]	Paranasal sinus and nasal cavity malignant diseases	PBT alone, PBT plus photon RT	No Meta-Analysis of Case Series	Date of inception of every database to April 2014	10 PBT studies (3 new) (2 comparative, 8 case series)	*Indirect comparison of case series of PBT vs. case series of photon was done. Only pooled data for PBT is presented.* - OS at longest follow-up (n=8 studies, 191 patients): 63% (95% CI, 53% to 76%) - 5-year OS (n=5 studies, 124 patients): 66% (95% CI, 52% to 85%) - DFS at longest follow-up (n=2 studies, 56 patients): 49% (95% CI, 21% to 116%) - 5-year DFS (n= 1 study, 36 patients): 72% (95% CI, 59% to 89%) - Locoregional control at longest follow-up (n=7 studies, 147 patients): 81% (95% CI, 71% to 92%) - 5-year locoregional control (n=2 studies, 36 patients): 43% (95% CI, 9% to 210%)

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available	Primary Conclusions
Lung Cancer						
Chi (2017)[47]	NSCLC	PBT	No Meta-Analysis of Case Series	January 2000 to June 2016	9 PBT case series (3 new)	*Indirect comparison of case series of PBT vs. case series of SBRT was done. Only pooled data for PBT is presented.* - 1-year OS (95% CI): 91.7% (82% to 100%) - 3-year OS (95% CI): 69.5% (39% to 100%) - 5-year OS (95% CI): 60% (23% to 97%) - 1-year PFS (95% CI): 85.3% (76% to 95%) - 3-year PFS (95% CI): 63.5% (37% to 90%) - 5-year PFS (95% CI): 57.2% (19% to 95%) - 1-year LC (95% CI): 96.3% (90% to 100%) - 3-year LC (95% CI): 87.4% (73% to 100%) - 5-year LC (95% CI): 87.2% (73% to 100%) Toxicity (N=614 patients) - Grade 3-5 toxicity (95% CI): 4.8% (3.4% to 6.7%) - Radiation Pneumonitis ≥ grade 3 (95% CI): 0.9% (0.4% to 1.9%) - Chest Wall Toxicity ≥ grade 3 (95% CI): 1.9% (1.1% to 3.3%) - Rib fractures (95% CI): 13% (11% to 16%)
Breast Cancer						
Verma (2016a) [296]	Breast Cancer	PBT vs. various photon therapy treatments	No	Date of inception of database to July 1, 2015	9 PBT case series (4 new that are not abstracts from conferences)	- Conventionally fractionated breast/chest wall PBT produces grade 1 dermatitis rates of approximately 25% and grade 2 dermatitis in 71% to 75%. - The incidence of esophagitis was decreased if the target coverage was compromised in the medial supraclavicular volume.

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available	Primary Conclusions
						- PBT offers excellent potential to minimize the risk of cardiac events, keeping the mean heart dose at <1 Gy.
Kammerer (2018) [133]	Breast Cancer	Passive Scatter PBT vs. Pencil Beam PBT vs. photon therapy (various)	No	NR	13 PBT case series (7 new)	- PBT vs. photon therapy: PBT appears to be promising in left breast cancer adjuvant radiotherapy including nodal areas. Dosimetric gains seem to be consistent, particularly in case of post-mastectomy irradiation, or irradiation of CMI. - It remains to be evaluated whether PBT actually brings a reduction in late cardiac toxicity. - Skin toxicity remains a concern but may be reduced with new techniques of PBT such as PBS.
Brain, Spinal, and Paraspinal (Including Skull-base)						
Pennicooke (2016) [222]	Chordoma of the Spine and Sacrum	PBT, PBT + photon RT	No Meta-Analysis of Case Series	1974 to March 2016	17 PBT case series (6 new)	*Indirect comparison of case series of PBT vs. case series of photon RT vs. case series of Carbon Ion therapy was done. Only pooled data for PBT is presented.* - The PBT studies shows a clear trend towards optimal LC rates with primary RT for de novo chordoma only when the dose deliver is >70 Gy(RBE) in 16 fractions. However, such a treatment modality is also associated with higher toxicity rates and adverse effects
Zhou (2018) [328]	Chordoma	PBT	No	Database inception to May 2017	9 PBT case series (1 new)	*Indirect comparison of case series of PBT vs. case series of photon RT vs. case series of Carbon Ion therapy was done. Only pooled data for PBT is presented.*

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available	Primary Conclusions
			Meta-analysis of Case Series			- 3-year OS (95% CI): 89% (85%-93% - 5-year OS (95% CI): 78% (23%-84%) - 10-year OS (95% CI): 60% (43%-77%)
Ocular Tumors						
Verma & Mehta (2016) [294]	Uveal Melanoma	PBT	No	January 1, 2000 to June 2015	14 case series (3 new)	- 5-year LC rates exceed 90%, which persisted at 10 and 15 years. - 5-year OS ranged from 70% to 85% - 5-year metastasis-free survival and disease-specific survival rates ranged from 75% to 90%, with more recent series reporting higher values. - With the removal of smaller studies, 5 year enucleation rates were consistently between 7% and 10%. - Many patients (60%-70%) showed a post-PBT visual acuity decrease, but still retained purposeful vision (>20/200); more recent, higher-volume series reported superior numbers. - Complication rates were quite variable but showed improvements on historical plaque brachytherapy data.
General/Mixed						
Verma (2016) [293]	Stomach (n=2 studies); Esophageal (n=13 studies); Pancreas (n=6 studies); Hepatobiliary (n=14 studies); Liver (n=4 studies); Retroperitoneal (n=2 studies)	PBT	No	Date of inception of database to October 15, 2015	39 cohorts from 41 publications (9 new that aren't abstracts from conferences; n=2 Esophageal; n=6 Hepatobiliary; n=1 Retroperitoneal)	- Limited quality (and quantity) of data hamper direct comparisons and conclusions. However, the available data, despite the inherent caveats and limitations, suggest that PBT offers the potential to achieve significant reduction in treatment-related toxicities without

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available	Primary Conclusions
						compromising survival or LC for multiple malignancies. **Esophageal** - PBT is associated with reduced toxicities, postoperative complications, and hospital stay as compared to photon radiation, while achieving comparable local control (LC) and overall survival (OS). **Pancreatic** - Numerical survival for resected/unresected cases is similar to existing photon data, whereas grade ≥3 nausea/emesis and post-operative complications are numerically lower than those reported with photon RT. **Hepatocellular Carcinoma** - The strongest data in support of PBT for HCC comes from phase II trials demonstrating very low toxicities, and a phase III trial of PBT versus transarterial chemoembolization demonstrating trends towards improved LC and PFS with PBT, along with fewer post-treatment hospitalizations. **Other** - Survival and toxicity data for cholangiocarcinoma, liver metastases, and retroperitoneal sarcoma are also roughly equivalent to historical photon controls.
Verma (2017) [295]	Ocular (n=1 study); Brain, Spinal, and Paraspinal (n=5 studies); Head and Neck (n=4 studies); Lung (n=2 studies); GI (n=4 studies)	PBT for reRT	No	Date of inception of database to June 2017	16 studies (13 new; n=2 Adult Brain; n=2 Pediatric Brain; n=4 Head and Neck; n=1 Lung; n=4 GI)	**Ocular** - PBT for recurrent uveal melanoma achieved 5-year eye retention of 55% **Brain - Adult**

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available	Primary Conclusions
						- For chordomas, reRT afforded a 2-year local control and OS of 85% and 80%, respectively. - Multiple PBT reRT studies for adult gliomas illustrate no grade ≥3 toxicities. **Brain - Pediatric** - Two pediatric CNS tumor studies demonstrated the safety and efficacy of reRT, with one total grade 3 toxicity and achievement of longer-term OS. **Head and Neck** - PBT for Head and Neck malignancies shows appropriate local/locoregional control and favorable toxicity profiles versus historical photon-based methods, including low (9–10%) rates of feeding tube placement. **Lung** - PBT for recurrent lung cancer can achieve favorable survival with expected toxicities/complications of reRT, especially with concurrent chemotherapy and centrally located recurrences. **GI** - PBT reRT in gastrointestinal malignancies induced very few high-grade complications.
Verma (2017) [297]	Head/Neck/Thoracic (n=6 studies); Prostate (n=8 studies); Pediatric (n=3 studies)	PBT	No	Date of inception of database to March 2017	17 studies (13 new; Head/Neck/Thoracic, n=5; Prostate, n=6; Pediatric, n=2)	- Based on limited data, PBT provides favorable QOL/PRO profiles for select brain, head/neck, lung, and pediatric cancers; measures for prostate and breast cancers were more modest. These results have implications for cost-effective cancer care

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available	Primary Conclusions
						and prudently designed QOL evaluation in ongoing trials.
Doyen (2016) [68]	Lung (n=4 studies); Breast (n=2 studies); Esophageal (n=3 studies); Head and Neck (n=3 studies); Pancreas (n=1 study); Liver (n=6 studies); Lymphoma (n=1 study); Soft Tissue Sarcoma (n=1 study); Prostate (n=2 studies); Gynecological (n=1 study)	PBT	No	NR	24 studies with clinical outcomes (12 new; Lung, n=2; Breast, n=1; Esophageal, n=2; Head/Neck, n=2; Liver, n=3; Lymphoma, n=1; Soft Tissue Sarcoma, n=1)	- Use of PBT in cancers should be prioritized for patients with high survival rates and/or young patients, for example Hodgkin lymphoma or breast cancer: in these locations PT could yield less cardiac toxicities and radiation-induced cancers compared to photon therapy as demonstrated in the literature.

CNS: Central Nervous System; DFS: Disease Free Survival; GI: Gastrointestinal; HCC: Hepatocellular Carcinoma; IMRT: Intensity Modulated Radiation Therapy; LC: Local Control; NSCLC: Non-small Cell Lung Cancer; OS: Overall Survival; PBS: Pencil Beam Scanning; PBT: Proton Beam Therapy; PFS: Progression Free Survival; PRO: Patient Reported Outcomes; QOL: Quality of Life; reRT: Re-irradiation Radiotherapy; SBRT: Stereotactic Body Radiotherapy

2.6.3 Previous Systematic Reviews of Proton Beam Therapy in Pediatric Populations

A total of 2 Systematic Reviews, one in brain tumors[21] and one in multiple cancer types[158], with new evidence since the prior report's search dates were identified by the search evaluating clinical and/or safety outcomes of PBT in pediatric populations.

Table 5. Summary of Previous Systematic Reviews of PBT in Pediatric Populations

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available (Number of included studies published since last report)	Primary Conclusions
Brain, Spinal, and Paraspinal						
Armoogum (2015) [21]	Medulloblastoma, carniopharyngioma, low grade gliomas, endymoma	PBT vs. IMRT	No	Date of inception of database to October 2014 with an update in November 2014	2 studies evaluating clinical/safety outcomes (2 new for clinical outcomes)	- The authors state that the data they reviewed demonstrated superiority of protons over photons for CNS tumors in children in terms of late neurocognitive, behavioral, vascular effects, health-related quality of life scores, endocrine abnormalities and second tumors.
General/Mixed						
Leroy (2016) [158]	Skull-base chondrosarcoma, Paraspinal and skull-based chordoma, craniopharyngioma, ependymoma, Esthesioneuroblastoma, Ewing sarcoma, CNS germinoma, Low-grade glioma, medulloblastoma, nonresectable osteosarcoma, retinoblastoma, Rhabdomyosarcoma	PBT PBT vs. photon RT	No	January 1, 2007 to March 21, 2014 with an update on June 25, 2015	23 studies (9 New: 5 Brain; 2 Ocular; 2 Soft Tissue Sarcoma)	- For retinoblastoma, very low-level evidence was found that PBT might decrease the incidence of second malignancies. - For chondrosarcoma, chordoma, craniopharyngioma, ependymoma, esthesioneuroblastoma, Ewing sarcoma, central nervous system germinoma, glioma, medulloblastoma, osteosarcoma, and rhabdomyosarcoma, there was insufficient evidence to either support or refute PBT in children. - For pelvic sarcoma (i.e., nonrhabdomyosarcoma and non-Ewing sarcoma), pineal parenchymal tumor, primitive neuroectodermal tumor, and

Assessment (year)	Specific Diagnosis	Treatments Evaluated	Network Meta-analysis or Indirect Analysis?	Search Dates	Evidence Base Available (Number of included studies published since last report)	Primary Conclusions
						"adult-type" soft tissue sarcoma, no studies were identified that fulfilled the inclusion criteria. - Although there is no doubt that PBT reduces the radiation dose to normal tissues and organs, to date the critical clinical data on the long-term effectiveness and harm associated with the use of PBT in the 15 pediatric cancers under investigation are lacking.

CNS: Central Nervous System; IMRT: Intensity Modulated Proton Therapy; PBT: Proton Beam Therapy; RT: radiotherapy

2.7 Medicare and Representative Private Insurer Coverage Policies

For the purposes of this report we obtained and summarized payer policies from two bellwether payers and any relevant information on NCDs and/or LCDs from the Centers for Medicare and Medicaid Services. Additionally, we received model policies from several proton beam therapy organizations which have been addressed in public comment summary documents available through the Washington Health Technology Assessment Program's website. An overview of CMS and payer policies decisions is available in section 2.3. Policies are summarized below:

- **Centers for Medicare and Medicaid Services (CMS) LCD (2018) [7,9,10]**
  In 2015, CMS released a Local Coverage Determination by the Wisconsin Physicians Service Insurance Corporation with jurisdiction applicable on a state-by-state basis (including Washington State).[7] Until this time, although proton beam therapy had been considered for selection, no National Coverage Determination had been reached. Later, as of September 2017, the LCD was retired and no determination applicable to Washington State has replaced it. However, two LCDs applicable to twelve states not including Washington are currently active. [9,10] Details of coverage in these LCD are provided below.

- **Aetna (2018)[14]**
  Aetna considers proton beam therapy as medically necessary for skull-base chordomas or chondrosarcomas, pediatric malignancies in children (21 years of age and younger) and uveal melanomas confined to the globe (i.e., not distant metastases). Other conditions are considered either not medically necessary (such as localized prostate cancer), or investigational/experimental (all other conditions, see full list below).

- **Anthem (2018)[18]**
  Anthem considers proton beam radiation therapy as medically necessary for primary therapy use in non-metastatic uveal melanoma of the uveal tract, post-operative use in residual, non-metastatic chordoma or low-grade (I or II) chondrosarcoma of the basisphenoid region, as an alternative when conventional radiation is not available for pituitary adenoma, for non-operative intracranial arteriovenous malformation (AVM), for malignancies adjacent to optical nerve, brain stem or spinal cord, and for pediatric malignancies. Choroidal neovascularization secondary to age-related macular degeneration (AMD) is considered not medically necessary, whereas proton beam irradiation is considered investigational for all other conditions.

Table 6. Overview of Medicare and Payer Policies

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
Centers for Medicare and Medicaid Services [7,9,10]	71 references, evidence not characterized	At present, there is no NCD for proton beam therapy; additionally, the only published LCD (L34634) on PBT that covered all states (including Washington) and was used in the prior report was retired as of Sept. 1st 2017 (https://www.cms.gov/medicare-coverage-database/details/lcd-details.aspx?LCDId=34634&ver=15&Date=&DocID=L34634), however, two LCDs (L35075 and L36658) applying to twelve states (not including Washington) are active with similar coverage conditions as the retired LCD. Conditions of the active and retired LCDs are provided below with additions from the active LCDs highlighted in bold: Conditions for Medical Necessity CMS considers PBT reasonable when sparing the surrounding normal tissue cannot be adequately achieved with photon-based radiotherapy and is of added clinical benefit to the patient. Examples of treatment advantage may include: 1. The target VOLUME is in close proximity to one or more critical structures and a steep dose gradient outside the target must be achieved to avoid exceeding the tolerance dose to the critical structure(s) 2. A decrease in the amount of dose inhomogeneity in a large treatment VOLUME is required to avoid an excessive dose "hotspot" within the treated VOLUME to lessen the risk of excessive early or late normal tissue toxicity. 3. A photon-based technique would increase the probability of clinically meaningful normal tissue toxicity by exceeding an integral dose-based metric associated with toxicity. 4. The same or an immediately adjacent area has been previously irradiated, and the dose distribution within the patient must be sculpted to avoid exceeding the cumulative tolerance dose of nearby normal tissue. Conditions considered frequently supported by the above requirements (Group 1) include: • Ocular Tumors, including intraocular melanomas • Skull-base tumors including but not limited to: ○ Chordomas ○ Chondrosarcomas ○ Primary or metastatic tumors of the spine where spinal cord tolerance may be exceeded with conventional treatment or where the spinal cord has previously been irradiated	Rationale: NR

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
		• Unresectable benign or malignant tumors of the CNS, including but not limited to: ○ Astrocytoma, glioblastoma, medulloblastoma, acoustic neuroma, craniopharyngioma, benign and atypical meningioma, pineal gland tumors, and arteriovenous malformations • Primary hepatocellular cancer treated in a hypofractionated regimen • Pediatric Primary or benign solid tumors in children treated with curative intent and occasional palliative treatment of childhood tumors when at least one of the four criteria noted above apply • Pituitary neoplasm • Advanced staged and/or unresectable malignant lesions of the head and neck • Malignant tumors of the paranasal and other accessory sinuses • Unresectable retroperitoneal sarcoma • **Patients with genetic syndromes making total volume of radiation minimization crucial such as but not limited to NF-1 patients and retinoblastoma patients** Coverage is considered <u>investigational</u> and limited to providers who have demonstrated experience in data collection and analysis with a history of publication in the peer-reviewed medical literature for the following conditions (group 2): • Unresectable lung cancers, upper abdominal cancers, and left breast tumors • Advanced, unresectable pelvic tumors, pancreatic and adrenal tumors • Skin cancer with nerve innervation of the skull base • Unresectable lesions of the liver, biliary tract, anal canal and rectum • Non-metastatic prostate cancer, with documented clinical staging and demonstration of clinical necessity of PBT • **Hodgkin or Non-Hodgkin Lymphoma involving the mediastinum or in non-mediastinal sites where PBT has the potential to reduce the risk of pneumonitis or late effects of radiation therapy**	
Bellwether Policies			
Aetna (2018) [14]	Literature Review (166 references) including:	Aetna considers proton beam radiotherapy (PBRT) **medically necessary** in any of the following radiosensitive tumors:	<u>Rationale: NR</u>

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
	1 CER (VHA 2015), 2 CADTH assessments, 1 assessment of economic evaluation (VATAP, Flynn 2010), 1 AHRQ assessment (Trikalinos 2009),4 HTAs (Wild 2013,RIHTA 2011, ICER 2008, Washington HTA 2014), guidelines from ASTRO NCCN, ACR, and Alberta Health Services; 7 SRs (Lodge 2007; Lance, 2010; Brada et al, 2009; Efstathiou et al, 2009; ICER, 2008; Wilt et al, 2008; Brada et al, 2007; Olsen et al, 2007), various studies	a. Chordomas or chondrosarcomas arising at the base of the skull or cervical spine without distant metastases; or b. Malignancies in children (21 years of age and younger); or c. Uveal melanomas confined to the globe (i.e., not distant metastases) (the uvea is comprised of the iris, ciliary body, and choroid [the vascular middle coat of the eye]). Aetna considers proton beam radiotherapy for treatment of prostate cancer **not medically necessary** for individuals with localized prostate cancer because it has not been proven to be more effective than other radiotherapy modalities for this indication. Proton beam therapy for metastatic prostate cancer is considered experimental and investigational. Aetna considers proton beam radiotherapy **experimental and investigational** for all other indications, including the following indications in adults (over age 21) (not an all-inclusive list) because its effectiveness for these indications has not been established: • Adenoid cystic carcinoma • Age-related macular degeneration (AMD) • Angiosarcoma • Atypical meningioma • Bladder cancer • Brain tumors • Breast cancer • Cardiac intimal sarcoma • Carotid body tumor • Cavernous hemangioma • Cervical cancer • Cholangiocarcinoma • Choroidal hemangioma • Dermatofibrosarcoma protuberans • Desmoid fibromatosis • Desmoid tumor (aggressive fibromatosis) • Ependymoma • Esophageal cancer • Ewing's sarcoma • Fibrosarcoma of the extremities • Gangliomas • Glioma • Head and neck cancer (including nasopharyngeal carcinoma) • Hemangioblastoma • Hemangioendothelioma	

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
		• Hepatocellular carcinoma • Lymphomas (Large cell lymphoma, Hodgkin's lymphoma, Non-Hodgkin lymphoma) • Intracranial arterio-venous malformations • Leiomyosarcoma of the extremities • Liposarcoma • Liver metastases • Lung cancer (including non-small-cell lung carcinoma) • Maxillary sinus tumor • Mesothelioma • Multiple myeloma • Nasopharyngeal tumor • Non-uveal melanoma • Oligodendroglioma • Optic nerve schwannoma • Optic nerve sheath meningioma • Pancreatic cancer • Parotid gland tumor • Pineal tumor • Pituitary neoplasms • Rectal cancer • Retroperitoneal/pelvic sarcoma • Rhabdomyoma • Sacral chordoma • Salivary gland tumors (e.g., sublingual gland tumor, submandibular gland tumor) • Seminoma • Sino-nasal carcinoma • Small bowel adenocarcinoma • Soft tissue sarcoma • Squamous cell carcinoma of the eyelid, tongue/glottis • Thymic tumor • Thymoma • Tonsillar cancer • Uterine cancer • Vestibular schwannoma • Yolk cell tumor	
Anthem (2018) [18]	Literature review (149 references) including: Guidelines from ASTRO, ACR, AAO, NCCN; 1 BCBS technology	Updated 02/2018 Anthem considers proton beam radiation therapy, with or without stereotactic techniques, as **medically necessary** for any of the following conditions: a. As primary therapy for melanoma of the uveal tract (iris, choroid, or ciliary body) involving tumors of up to 24 mm in largest diameter and 14 mm in height,	Rationale: NR

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
	assessment, 2 ongoing trials; 4 AHRQ reviews	and with no evidence of metastasis or extrascleral extension; or b. As postoperative therapy for individuals who have undergone biopsy or partial resection of a chordoma or low-grade (I or II) chondrosarcoma of the basisphenoid region (for example, skull-base chordoma or chondrosarcoma) or cervical spine and have residual, localized tumor without evidence of metastasis; or c. Pituitary adenoma when conventional stereotactic radiation is not an available option; or d. Intracranial arteriovenous malformation (AVM) not amenable to surgical excision or other conventional forms of treatment; or e. Central nervous system (CNS) lesions including but not limited to, primary or metastatic CNS malignancies or AVM, adjacent to critical structures such as the optic nerve, brain stem or spinal cord; or f. Primary or benign solid tumors in children treated with curative intent. Proton beam radiation therapy is considered **not medically necessary** for the following condition: Choroidal neovascularization secondary to age-related macular degeneration (AMD). Proton beam radiation therapy is considered **investigational and not medically necessary** when criteria are not met and for all other indications, including, but not limited to, the treatment of: Localized prostate cancer.	

CMS = Centers for Medicare and Medicaid Service; HTA = Health Technology Assessment; NCDs = National coverage determination; NR = not reported; SR = systematic review

3 The Evidence

3.1 Methods of the Systematic Literature Review

3.1.1 Objectives

The aim of this report is to update the 2014 HTA on proton beam therapy (PBT) by systematically reviewing, critically appraising and analyzing new research evidence on the safety and efficacy of PBT, both as a primary or as a salvage therapy (i.e., for recurrent disease or failure of initial therapy), for the treatment of multiple types of cancer as well as selected noncancerous conditions in adults and children.

3.1.2 Key Questions

1. What is the comparative impact of proton beam therapy treatment with curative intent on survival, disease progression, health-related quality of life, and other patient outcomes versus radiation therapy alternatives and other cancer-specific treatment options (e.g., surgery, chemotherapy) for the following conditions:
    a. Cancers
        i. Bone tumors
        ii. Brain, spinal, and paraspinal tumors
        iii. Breast cancer
        iv. Esophageal cancer
        v. Gastrointestinal cancers
        vi. Gynecologic cancers
        vii. Head and neck cancers (including skull base tumors)
        viii. Liver cancer
        ix. Lung cancer
        x. Lymphomas
        xi. Ocular tumors
        xii. Pediatric cancers (e.g., medulloblastoma, retinoblastoma, Ewing's sarcoma)
        xiii. Prostate cancer
        xiv. Soft tissue sarcomas
        xv. Seminoma
        xvi. Thymoma
        xvii. Other cancers
    b. Noncancerous Conditions
        vii. Arteriovenous malformations
        viii. Hemangiomas
        ix. Other benign tumors (e.g., acoustic neuromas, pituitary adenomas)
2. What is the comparative impact of salvage treatment (including treatment for recurrent disease) with proton beam therapy versus major alternatives on survival, disease progression, health-related quality of life, and other patient outcomes versus radiation therapy alternatives and other cancer-specific treatment options (e.g., surgery, chemotherapy) for the condition types listed in key question 1?
3. What are the comparative harms associated with the use of proton beam therapy relative to its major alternatives, including acute (i.e., within the first 90 days after treatment) and late (>90

days) toxicities, systemic effects such as fatigue and erythema, toxicities specific to each cancer type (e.g., bladder/bowel incontinence in prostate cancer, pneumonitis in lung or breast cancer), risks of secondary malignancy, and radiation dose?

4. What is the differential effectiveness and safety of proton beam therapy according to factors such as age, sex, race/ethnicity, disability, presence of comorbidities, tumor characteristics (e.g., tumor volume and location, proliferative status, genetic variation) and treatment protocol (e.g., dose, duration, timing of intervention, use of concomitant therapy)?
5. What is the comparative cost-effectiveness of proton beam therapy in the short- and long-term relative to other types of radiation therapy, radiation therapy alternatives or other cancer-specific treatment options (e.g., surgery, chemotherapy)?

3.1.3 Inclusion/Exclusion Criteria

Table 7Table 7. Briefly, included studies met the following requirements with respect to participants, intervention, comparators, outcomes, and study design:

- **Population**: Adults and children undergoing treatment of primary or recurrent disease, to include cancer types (bone cancer, brain, spinal, and paraspinal tumors, breast cancer, esophageal cancer, gastrointestinal cancer, gynecologic cancer, head and neck cancer, liver cancer, lung cancer, lymphomas, ocular tumors, pediatric cancers, prostate cancer, sarcomas, seminoma, thymoma, other cancers) and noncancerous conditions (arteriovenous malformations, hemangiomas, other benign tumors).

- **Interventions:** Proton beam therapy; all approaches were considered including monotherapy, use as a "boost" mechanism to conventional radiation, and combination therapy with other treatment modalities (e.g., chemotherapy, surgery).

- **Comparators:** Primary comparators include other radiation alternatives (e.g., intensity-modulated radiation therapy (IMRT), stereotactic radiation techniques and other external beam therapies, and brachytherapy). Other treatment alternatives specific to each condition type treated, and may include chemotherapy, immunotherapy, surgical procedures, and other devices (e.g., laser therapy for ocular tumors).

- **Outcomes:**

Primary Clinical outcomes:

- Overall survival/disease-free survival
- All-cause and/or disease-related mortality
- Direct measures of tumor regression, control or recurrence
- Incidence of metastases

Secondary or indirect (intermediate) outcomes

- Patient reported outcomes including health-related quality of life (HrQoL) using validated instruments
- Requirements for subsequent therapy
- Other outcomes specific to particular conditions (e.g., visual acuity for ocular tumors, shunt requirements for arteriovenous malformations)
- Intermediate measures of tumor recurrence such as biochemical measures

Safety outcomes:

- Treatment-related harms, to include generalized effects (e.g., fatigue, erythema) and localized toxicities specific to each condition (e.g., urinary incontinence in prostate cancer, pulmonary toxicity in lung or breast cancer); the primary focus is on adverse effects requiring medical attention
- Secondary malignancy risk due to radiation exposure

Economic outcomes:

- Long term and short term comparative cost-effectiveness measures

- **Studies:** The focus will be on high quality (low risk of bias) comparative studies (e.g., randomized controlled trials, comparative cohort studies with concurrent controls) will be considered for Key Questions 1-4. Comparative observational studies with long term clinical outcomes or safety will be considered for Key Questions 1-4. Case series will be considered but will not be the primary focus of evaluation for each key question. Dosimetry and planning studies will be included for context; the will be included as evidence if they directly answer the key questions. Full, comparative, formal economic studies (i.e., cost-effectiveness, cost-utility, cost-minimization, and cost-benefit studies) will be sought for Key Question 5; studies using modeling may be used to determine cost-effectiveness.

Table 7. Summary of inclusion and exclusion criteria

Study Component	Inclusion	Exclusion
Population	Adults and children undergoing treatment of primary or recurrent disease to include: • Cancers (bone, brain/spinal/paraspinal, breast, esophageal, gastrointestinal, gynecologic, head and neck, liver, lung, ocular, pediatric, and prostate cancers; lymphomas, sarcomas, seminomas, thymomas, other cancers) • Noncancerous conditions (arteriovenous malformations, hemangiomas, other benign tumors).	• Conditions not amenable to proton-beam therapy or for which proton beam therapy would be contra-indicated.
Interventions	Proton beam therapy (PBT) use as a • Curative therapy • Primary or monotherapy • "Salvage" treatment (e.g. following failure of initial therapy or disease recurrence) • "Boost" mechanism to conventional radiation • Combination therapy with other treatments (e.g., chemotherapy, surgery).	• Devices or therapies that are not FDA approved or cleared
Comparator	• Other radiation therapy alternatives (e.g., intensity-modulated radiation therapy (IMRT), stereotactic radiation techniques, other external beam therapies, and brachytherapy) • Other treatment alternatives specific to each condition type treated; may include	• Technologies or treatments that are not widely available or are no longer routinely used • Devices or therapies that are not FDA approved or cleared

Study Component	Inclusion	Exclusion
	chemotherapy, immunotherapy, surgical procedures, and other devices (e.g., laser therapy for ocular tumors). • Dose/fractionation comparison (will be included for completeness as was done in prior report) but not formally evaluated as evidence	
Outcomes	**Clinical outcomes**: <u>Primary</u> • Overall survival/disease-free survival • All-cause and/or disease-related mortality • Direct measures of tumor regression, control or recurrence • Incidence of metastases <u>Secondary or indirect (intermediate) measures</u> • Patient reported outcomes, including health-related quality of life (HrQoL), based on validated instruments • Requirements for subsequent therapy • Other outcomes specific to particular conditions (e.g., visual acuity for ocular tumors, shunt requirements for arteriovenous malformations) • Intermediate measures of tumor recurrence such as biochemical measures **Safety outcomes**: • Treatment-related harms, with a focus on adverse effects requiring medical attention, to include: ♦ Generalized effects (e.g., fatigue, erythema) ♦ Localized toxicities specific to each condition (e.g., urinary incontinence in prostate cancer, pulmonary toxicity in lung or breast cancer) to include consideration of: ▪ Early (≤90 days post-treatment) ▪ Late (>90 days post-treatment) • Secondary malignancy risk due to radiation exposure **Economic outcomes**: • Long term and short term comparative cost-effectiveness measures (e.g. ICER)	• Non-clinical outcomes
Study Design	• Focus will be on highest quality (lowest risk of bias) comparative studies (e.g., randomized controlled trials, comparative cohort studies with concurrent controls) for questions 1-4.	• Simulation studies • Studies of low quality (high risk of bias) • Comparative studies with fewer than 10 per treatment arm • Case reports

Study Component	Inclusion	Exclusion
	• Case series will be considered but will not be the primary focus of evaluation for each key question. • Case series in children with <10 patients will be considered if no comparative studies are available. • Case series designed specifically to evaluate safety may be included • Dosimetry and planning studies may be included for context. To the extent that they specifically answer the key questions, information will be included as part of the evidence base. • Formal, full economic studies will be sought for question 5. Studies using modeling may be used to determine cost-effectiveness.	• Case series in adults with <30 patients; Case series of ≥ 10 patients may be considered for very rare conditions. • Studies comparing modes of therapy; dose comparisons may be included for completeness/context per previous report
Publication	• Studies published in English in peer reviewed journals, technology assessments or publically available FDA reports • Studies published subsequent to the 2014 report (previous report search date through February 2014) • For question 5, comparative, full formal economic analyses (e.g., cost-effectiveness, cost-utility studies) published in English in a peer reviewed journal	• Abstracts, editorials, letters • Duplicate publications of the same study that do not report different outcomes or follow-up times • Single reports from multicenter trials • White papers • Narrative reviews • Articles identified as preliminary reports when full results are published in later versions • Incomplete economic evaluations such as costing studies

*In the absence of such studies, contextual information on treatments and outcomes in untested patients will be described.

3.1.4 Data Sources and Search Strategy

We searched electronic databases from November 2013 to December 2018 to identify publications assessing the use of Proton Beam Therapy for the treatment of primary or recurrent cancerous and non-cancerous conditions that had been published since the original report. A formal, structured systematic search of the peer-reviewed literature was performed across a number of databases including PubMed, EMBASE, the Cochrane Central Register of Controlled Trials, and the Cochrane Database of Systematic Reviews (see Appendix B for full search strategy) to identify relevant peer reviewed literature as well as other sources (ClinicalTrials.gov, National Guideline Clearinghouse, Center for Reviews and Dissemination Database) to identify pertinent clinical guidelines and previously performed assessments. Additional details on the search strategy conducted for clinical guidelines can be found in Appendix H. We also hand searched the reference lists of relevant studies and the bibliographies of systematic reviews.

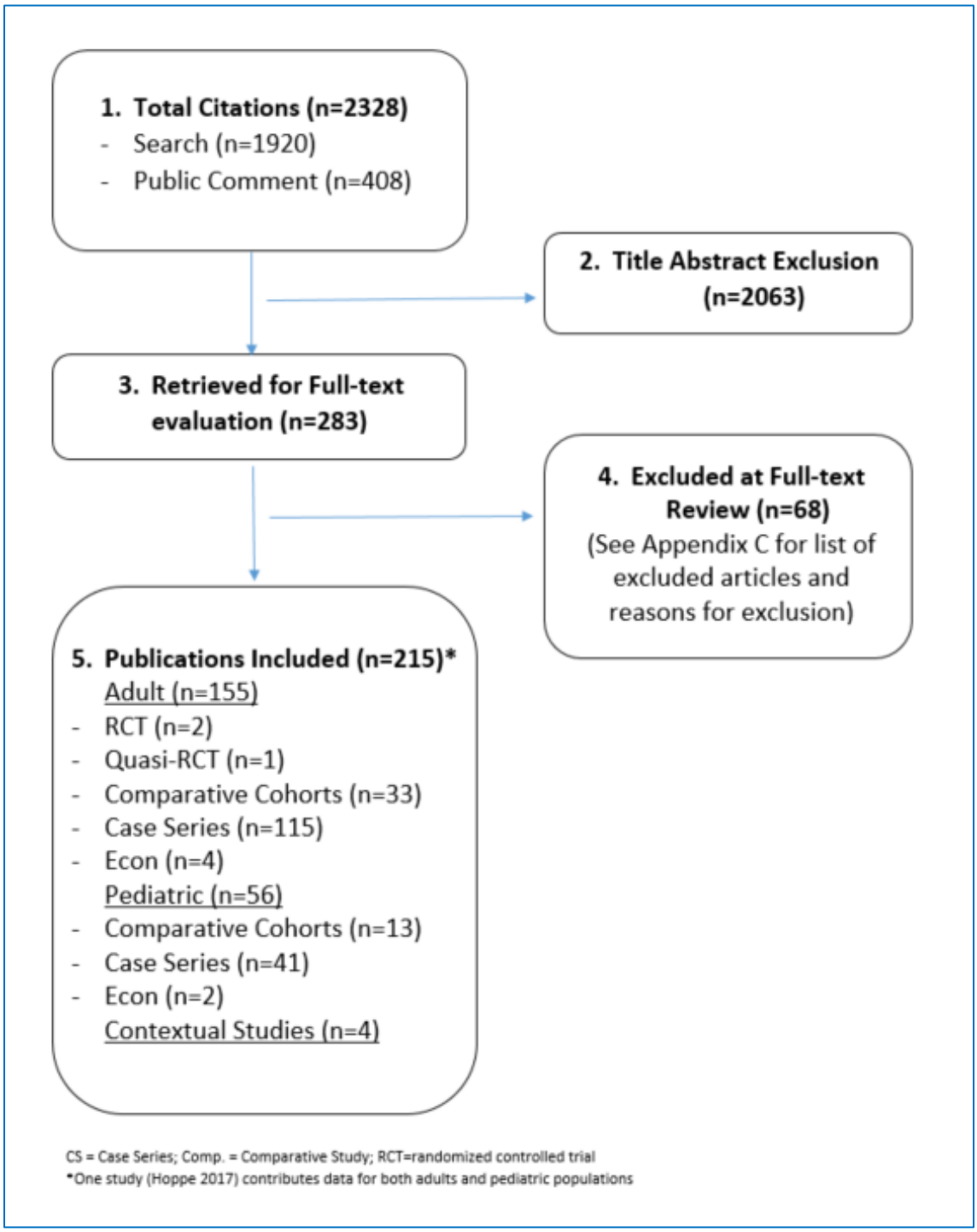
The clinical studies included in this report were identified using the algorithm shown in Appendix A. The search took place in four stages. The first stage of the study selection process consisted of the comprehensive electronic search and bibliography review. We then screened all possible relevant articles using titles and abstracts in stage two. This was done by two individuals independently. Those articles that met a set of *a priori* retrieval criteria were included for full-text review. We excluded conference abstracts, non-English-language articles, duplicate publications that did not report different data or follow-up times, white papers, narrative reviews, preliminary reports, and incomplete economic evaluations. Any disagreement between screeners that were unresolved resulted in the article being included for the next stage. Stage three involved retrieval of the full text articles remaining. The final stage of the study selection algorithm consisted of the review and selection of those studies using a set of *a priori* inclusion criteria, again, by two independent investigators. Discrepancies were resolved through discussion and if necessary adjudicated by a third investigator. A list of excluded articles along with the reason for exclusion is available in Appendix C.

Additionally, a total of 1,426 citations were received from comment received during the Topic Nomination and Draft Key Question public comment phase for this project, of which 390 remained after removal of duplicate citations and elimination of citations published prior to our specified search date range. These 390 studies were reviewed and compared in stage 2 alongside results from the search and included or excluded based on *a priori* criteria outlined above.

Consistent with the 2014 report, we focus on comparative studies. Comparative studies that provide a direct comparison of treatments in the same underlying patient population are considered to provide stronger evidence versus indirect comparisons of case series. Studies which indirectly compared results from separate case series were treated as case series and reported for PBT only.

Consistent with the 2014 report, given uncertainties regarding proton physics and the relative biological effectiveness of PBT in all tissues, particularly in adults, only limited appraisal and abstraction of studies included dosimetry, planning and simulation studies included for context was done and focused on any clinical outcomes reported. Studies that did not report on clinical outcomes were not included.

Figure 3. CONSORT Diagram - Flow of Studies



3.1.5 Data Extraction

Reviewers extracted the following data from the clinical studies: study design, study period, setting, country, sample size, inclusion and exclusion criteria, study population characteristics, follow-up time, study funding and conflicts of interest, proton beam therapy characteristics (e.g., technique, dose information, fractionation), tumor characteristics (e.g. histology, location, metastatic status), study outcomes and adverse events. For economic studies, data related to sources used, economic parameters and perspectives, results, and sensitivity analyses were abstracted. An attempt was made to reconcile conflicting information among multiple reports presenting the same data. Detailed study and patient characteristics and results are available in the Data Abstraction Appendices A-Q.

3.1.6 Quality Assessment: Overall Strength of Evidence, Risk of Bias, & QHES evaluation

The method used by Aggregate Analytics, Inc. (AAI) for assessing the quality of evidence of individual studies as well as the overall strength of evidence (SOE) for each primary outcome from comparative studies are based on criteria and methods established in the Cochrane Handbook for Systematic Reviews of Interventions[107], precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group[2,22,23,96,97], and recommendations made by the Agency for Healthcare Research and Quality (AHRQ). Economic studies were evaluated according to The Quality of Health Economic Studies (QHES) instrument developed by Ofman et al.[209] Based on these quality criteria, each comparative study chosen for inclusion for a Key Question was given a RoB (or QHES) rating; details of each rating are available in Appendix E. For comparative cohort studies, lost to follow-up (including differential loss to follow-up) and control for potential confounding are generally the primary sources of bias. Risk of bias was not assessed for case series (single arm studies); limited RoB assessment was done for dosimetry included for context. All case series were considered to be at high risk of bias. Standardized, pre-defined abstraction guidelines were used to determine the RoB (or QHES) rating for each study included in this assessment. Criteria are detailed in Main Appendix D.

In the absence of high quality RCTs, comparative, nonrandomized observational studies were included. Given that the primary evidence base for this report is from such studies, key criteria areas for potential bias included control for confounding and loss to follow-up, including differential loss to followup. Credit for confounding control was given if studies reported explicit evaluation of confounders and/or control of them using appropriate methods (e.g. multivariate analysis, matching). Studies using propensity scoring should specify and justify choice of covariates used for matching, describe statistical methods for propensity scoring and matched analyses and use processes to assess degree of balance between groups. Such studies were considered to have controlled for confounding. For purposes of this comparative effectiveness review, prospective comparative cohort studies which controlled for confounding and for which there was ≥ 80% follow-up and ≤10% difference in follow-up between treatments were considered “best evidence” in the absence of quality RCTs. In general, the above methods were consistent with the approach taken in the 2014 report. For study quality, both reports focus on comparability of groups with regard to measurements, patient retention (follow-up of at least 80%) and appropriate attention to confounders in analysis.

The SOE for all primary health outcomes was assessed by two researchers following the principles for adapting GRADE (Grades of Recommendation Assessment, Development and Evaluation)[23,96,97] as outlined by the Agency for Healthcare Research and Quality (AHRQ)[2]. The strength of evidence was based on the highest quality evidence available from comparative studies for a given outcome. In

determining the strength of body of evidence regarding a given outcome, the following domains were considered:

- **Risk of bias**: the extent to which the included studies have protection against bias
- **Consistency:** the degree to which the included studies report results that are similar in terms of effect sizes, range and variability.
- **Directness**: describes whether the evidence is directly related to patient health outcomes or comparisons of interventions are direct (head to head).
- **Precision:** describes the level of certainty surrounding the effect estimates.
- **Publication or reporting bias:** is considered when there is concern of selective publishing or selective reporting. This is difficult to assess particularly for nonrandomized studies.

When assessing the SOE for studies performing subgroup analysis, we also considered whether the subgroup analysis was preplanned *(a priori)* and whether a test for homogeneity or interaction was done.

Bodies of evidence consisting of RCTs are initially considered as High strength of evidence. In general, the GRADE and AHRQ methodologies initially consider nonrandomized studies as Low strength of evidence as such studies typically are at higher risk of bias due to lack of randomization and inability of investigators to control for critical confounding factors. In some instances (e.g. rare conditions, pediatric populations), RCTs may be unavailable, not feasible, not ethical or not substantially applicable to the target populations to be treated and use of high quality nonrandomized observational studies may provide the "best evidence" and may be considered to substitute for RCT evidence.[247] This does not, however, imply that the quality of nonrandomized studies is elevated only that such studies represent the best available evidence and that decision makers need to accept and consider the greater uncertainty of such evidence; one should not have greater confidence in the effect estimates from such studies. Observational studies with few methodologic limitations which control for risk of bias via study conduct or analysis may be initially considered as moderate versus low, particularly for harms and outcomes when such studies may be at lower risk of bias due to confounding.[30] There are also situations where studies (particularly observational studies) could be upgraded if the study had large magnitude of effect or if a dose-response relationship is identified and there are no downgrades for the primary domains listed above and confounding is not a concern.

The strength of evidence could be downgraded based on the limitations described above. There are also situations where studies (particularly observational studies) could be upgraded if the study had large magnitude of effect (strength of association) or if a dose-response relationship is identified and there are no downgrades for the primary domains listed above and confounding is not a concern. Publication and reporting bias are difficult to assess, particularly with fewer than 10 RCTs and for observational studies.[30,247] Publication bias was unknown in all studies and thus this domain was eliminated from the strength of evidence tables. The final strength of evidence was assigned an overall grade of high, moderate, low, or insufficient, which are defined as follows:

- High - Very confident that effect size estimates lie close to the true effect for this outcome; there are few or no deficiencies in the body of evidence; we believe the findings are stable.
- Moderate – Moderately confident that effect size estimates lie close to the true effect for this outcome; some deficiencies in the body of evidence; we believe the findings are likely to be stable but some doubt remains.

- Low – Limited confidence that effect size estimates lie close to the true effect for this outcome; major or numerous deficiencies in the body of evidence; we believe that additional evidence is needed before concluding that findings are stable or that the estimate is close to the true effect.
- Insufficient – We have no evidence, are unable to estimate an effect or have no confidence in the effect estimate for this outcome; OR no available evidence or the body of evidence has unacceptable efficiencies precluding judgment.

Similar methods for determining the overall quality (strength) of evidence related to economic studies have not been reported, thus the overall strength of evidence for outcomes reported in Key Question 5 was not assessed.

Primary outcomes for this report were OS and PFS, recurrence and treatment related toxicity or other harms. Strength of Evidence (SOE) was assessed for these primary outcomes only and details of other outcomes are provided in the full report. The results and SOE focus on the highest quality of evidence available. Where RCTs or higher quality evidence were available, these were used to assess the overall strength of evidence. In the absence of RCTs, the highest quality comparative observational studies were used to assess overall SOE. Evidence consisting of case series alone was considered insufficient as conclusions regarding comparative effectiveness or safety are not possible in the absence of a comparison with alternative treatments in groups of patients from the same underlying patient populations.

We compared overall conclusions and findings as reported in the previous report with findings in this update to the extent possible based on general qualitative concepts of AHRQ guidance on signal updates for systematic reviews[203], primarily based on the Ottawa Method.[252,256] Individual studies included in the prior report were not extensively evaluated by AAI. Considerations included:

- Comparison of the general quality of evidence of included comparative effectiveness studies on primary outcomes.
- Comparison of comparators used.
- Assessment of whether new evidence constitutes a major change in the evidence based on existence of opposing findings or major changes in effectiveness short of opposing findings based on the highest quality of evidence available (preferably from high quality systematic reviews or pivotal RCTs). Substantial changes in effect size (e.g. ≥50%) or changes in statistical significance beyond "borderline" changes (e.g. borderline p-values of 0.4 to 0.06) across studies of comparable quality were considered.
- Assessment of whether new evidence suggests substantial harm wherein risk of harm outweighs benefits.
- Assessment of whether new evidence provides high quality data on clinically important expansion of treatment (e.g. to new subgroups of patients) or clinically important caveat.

3.1.7 Analysis

Evidence was summarized qualitatively and quantitatively. As the majority of studies were observational and there was substantial heterogeneity across them with regard to patient populations, tumors/conditions studied, treatments and clinical methods, meta-analysis was not performed.
In the absence of adjusted effect size estimates, for dichotomous outcomes, crude risk ratios (RR) and 95% confidence intervals were calculated using either STATA or Rothman Episheet,[6,266] particularly for harms, if differences between treatments appeared to approach statistical significance for primary

outcomes/harms only. For instances with fewer than five observations per cell, exact methods were employed. These effect estimates cannot control for confounding. Where effect estimates that were adjusted for confounding were reported by study authors, they were preferred and reported. Risk differences were not calculated for observational studies as causality cannot be inferred.

Outcomes are detailed in the evidence tables in the appendices and/or the body of the report. Summary tables for case series are also found in the appendices.

4 Results

4.1 Number of Studies Retained & Overall Quality of Studies

Overall number of studies retained for this review

A total of 215 publications met inclusion criteria and form the basis for this review, including 205 publications (49 from comparative studies) that addressed efficacy, effectiveness or safety (one case series contributed data for both adults and pediatric populations separately[115]); six that addressed cost-effectiveness; and four that compared different treatment protocols or different dosing regimens for PBT and, to be consistent with the prior report, were included for context only.

A total of 56 publications were in pediatric tumors, including 13 retrospective comparative cohorts[15,31,32,71,72,89,95,131,132,148,221,245,265]; 41 case series[19,34,62,83,84,92,99,115,118-120,125,135,149-153,157,165,167,180,184-186,189,195,218,223,228,231,232,250,291,298,300,306,307,310,316,320]; and two cost-effectiveness studies[109,170]. The bulk of the evidence for this section was for the use of PBT in various pediatric brain tumors. A total of 159 publications were in adult tumors, including two RCTs (Liver and Lung cancer),[42,161] one quasi-RCT (Prostate cancer),[139] 33 retrospective comparative cohorts[13,33,35,37,48,69,76,77,94,108,111,129,163,164,168,173,178,196,206,216,233,238,244,251,255,258-260,274,282,303,317,326];115 case series (publications)[16,20,24,28,39,41,44-46,49,50,53,57,59,63,64,67,70,73,78-81,90,91,93,101,103,104,110,112-117,121-124,130,134,136,137,140,142,143,147,154,156,169,172,174-177,179,181,183,184,187,190,193,194,197,199,200,205,207,212-214,217,219,224,225,227,230,235,236,239,241-243,246,248,249,264,267,269-272,275-278,281,285,292,299,301,304,305,308,309,312,313,318,322-325,327]; four cost-effectiveness studies[159,171,192,253]; and four contextual studies (all in prostate cancer).[98,198,227,286] The majority of the evidence in adults was for the following cancers: Esophageal, Head and Neck, Brain, Lung, Ocular, and Prostate.

Overall quality of evidence base for this review

The overall quality of the available evidence base was considered poor; 46 comparative studies across 49 publications (observational or RCT) were identified, 43% of which were performed in same institution, the MD Anderson Cancer Center. Risk of bias assessment for included comparative studies is found in Main Appendix E. Comparative studies that provide a direct comparison of treatments in the same underlying patient population are considered to provide stronger evidence versus indirect comparisons of case series. Studies which indirectly compared results from separate case series were treated as case series and reported for PBT only. All case series were considered to be at high risk of bias and formal assessment was not performed.

Only two RCTs, one in Liver[42] and another in Lung[161] and one quasi-RCT in patients with Prostate cancer[139] met the inclusion criteria. Methodological issues in these studies included unclear concealment of allocation and lack of independent or blind assessment. (Main Appendix E). Few studies evaluated PBT for salvage therapy or as treatment for recurrence in answer to KQ2.

Comparative evidence for this report is primarily from non-randomized (observational) studies which were considered to be at moderately high risk of bias except where noted in the detailed description of results. Most studies were retrospective and a number of potential sources of bias must be considered when interpreting study findings. For purposes of this report, prospective comparative cohort studies

which controlled for confounding and for which there was ≥ 80% follow-up and ≤10% difference in follow-up between treatments were considered "best evidence" in the absence of quality RCTs. Few studies met all of these criteria.

In most instances, treatment groups were formed based on historical changes in methods of radiation therapy delivery, i.e. more conventional photon radiation therapy, including 3DCRT, was delivered to patients at a time prior to a switch to PBT as it became more the available. One consequence of the use of historically consecutive controls in these studies is differential length of follow-up by treatment group; historical groups receiving photon therapy had longer follow-up than those receiving PBT. This is a potential source of bias for comparing longer term benefits and harms in particular; there is potential bias related to survivorship. Studies using Kaplan-Meier or Cox regression methods for outcomes such as overall survival which account for person-time at risk and depending on the time frames may partially account for differential follow-up times. Completeness of follow-up and differential loss to follow was either not clearly reported or the criteria were not met as studies failed to delineate number of eligible persons/and or failed to adequately account for loss to follow-up, creating the potential for bias.

Across studies, blinded assessment of outcomes was not clearly reported and likely not done; for hard outcomes such as survival or mortality, this may not be of great concern.

Differences between treatment groups in patient characteristics, presentation, tumor stage, comorbidities, prior or concurrent treatments and surgical factors were noted in most studies. Although many studies evaluated possible confounding by such factors, there is the possibility of residual confounding or other biases that could influence results. In addition to potential confounding based on differences in patient characteristics (e.g. age, comorbidities) between treatment groups, treatment selection bias (confounding by indication), whereby individuals with more advanced or aggressive tumors are more likely to receive more intensive or aggressive treatments, presents another potential source of bias in the included studies; this concept may apply not only to primary treatments of interest but to co-interventions (e.g. chemotherapy) as well. Five studies used propensity score matching to reduce such bias by creating cohorts matched based on scores predicted from observed baseline values of specific confounding factors to balance treatment groups on such factors. While propensity score matching is a potentially useful method for adjusting for confounding variables and reducing treatment selection bias, there is the potential for creating biased effect estimates if groups are not well balanced and/or appropriate statistical methods for matched analysis are not used.[145,191,201] At minimum, studies using propensity scoring should specify and justify choice of covariates used for matching, describe statistical methods for propensity scoring and matched analyses and use processes to assess degree of balance between groups. Three of the propensity-score matching studies provided detail regarding evaluation of balancing, but only one appears to have used appropriate statistical methods for matched data. Five additional included studies used case matching to control for confounding but only two appear to have used appropriate statistical methods for matched data. With retrospective cohort studies in particular, despite use of such methods, there is still the potential for residual confounding and selection bias is still a potential concern.

Sample sizes across comparative studies varied (24 to 1,850), but most studies were small, with fewer than 50 participants per treatment arm. Sample sizes for rare tumor types were understandably smaller. Small sample sizes likely impacted the ability to detect rare events or statistical differences between treatment groups and are reflected in potentially inflated estimates of percentages for outcomes.

Comparison with 2014 report

The evidence base in the prior report primarily consisted of case series and focused on comparative studies for evaluation of benefits and harms as does this update. Comparative studies were primarily retrospective cohort studies. In general, the quality of comparative studies in the update report appears to be marginally better but varies somewhat by tumor category. Many studies published subsequent to the prior review had larger sample sizes, made direct comparisons of treatment groups and seemed to employ better methods for controlling for confounding and potential selection bias.

Many of the studies in the 2014 review used 3DCRT and some IMRT as a radiotherapy comparison with PBT; most of the studies in this update used IMRT and/or 3DCRT which may reflect some progression to more focused methods of RT delivery versus 2DCRT. The studies in the old report included a variety of comparators, many of which were not represented in the studies included in the update report. The prior report included carbon ion therapy as a comparator; it is not included in this review as it is not FDA approved. For some tumor categories, the comparators for studies included in the prior report were very different than comparators, which may reflect changes in clinical practice with time and may partially explain differences in findings between the 2014 report and this review. As an example, for ocular tumors, in the prior report, three studies compared PBT with surgical enucleation and one with transpupillary thermotherapy plus PBT. In this review, less invasive treatments (brachytherapy and stereotactic radiosurgery) were the comparators employed by included studies. Similarly for hepatocellular carcinoma, the interim RCT analysis included in this review compared PBT with transarterial chemoembolization (TACE) whereas the in the prior review, PBT combined with chemotherapy and carbon ion therapy were the comparators employed in separate studies. Thus, in drawing conclusions across both reviews for such instances, these differences need to be considered. For few tumor classifications RCT data were available in the previous report, but no new RCTs were identified for this review. In addition to heterogeneity in study design and implementation/comparators between included studies for the 2014 and 2019 reviews, specific tumor types and or stages studied in a given classification of tumor may differ between the 2014 and 2019 reports; use of prior or concurrent chemotherapy and other treatments across included studies may also differ within and between reports. Differences in evidence base, comparators and other factors are described with bulleted summary findings for the various tumor classifications.

Table 8 below provides a broad overview of the strength of evidence and direction of benefits for the 2014 review (based in their table ES2) compared with this 2019 review. (This overview does *not* connote any recommendations for policy). While for many tumor categories, general conclusions regarding benefits and harms are similar between the two reports, for some tumor types, general conclusions differ. These instances are described with the bulleted summary points for each tumor type.

Table 8. Summary of strength of evidence, direction of benefit and general comparison of the 2014 and 2019 report

Condition	Incidence (per 100000)	Number of Publications		Net Health Benefit vs. Comparators Type of Net Benefit (B, H) SOE		Impact of new studies (focus on comparative studies)*
		2014 Report†	2019 Report‡	2014 Report	2019 Report	2014 versus 2019 Report
Adults						
Cancer						
Bladder	20.3	CS=1	CS=1	NR	Insufficient	Similar conclusions
Bone	0.9	CC=1; CS=4	CS=8	Insufficient Low	Insufficient	Similar conclusions
Brain/Spinal	6.5	CC=2; CS=6	CC=5; CS=6	Incremental B: = H: ↓ Low	<u>PBT vs. photon</u> Unclear B: ↑ H: NR Low (curative); <u>PBT boost + photon vs. photon alone</u> Comparable B: = H: = Low (curative) Insufficient (salvage)	3 new retrospective comparative cohorts [2 curative (1 case-matched, 1 large propensity score-matched database) and 1 salvage] of different interventions and tumor types vs. 2014 report. The net health benefit for PBT vs. photon is unclear from 1 large data base study which did not report harms. For PBT boost + photon vs. photon alone, 1 cohort lead to different conclusions regarding harms. Evidence was insufficient for salvage therapy from 1 small cohort.
Breast	124.7	CS=4; Econ=3	CC=2 CS=4; Econ=1	Insufficient none	Unclear B: = H: NR Low	The net health benefit is unclear (addition of 1 large retrospective database study which did not report harms.)
Esophageal	4.6	CC=2; CS=7	CC=5; CS=2	Insufficient none	Incremental B: ↑ H: = Low	New retrospective comparative evidence [5 cohorts (2 propensity score-matched)], leads to different conclusions
GI	100.6§	CS=7	CC=1; CS=2	Insufficient none	Insufficient	Similar conclusions (1 small retrospective

Condition	Incidence (per 100000)	Number of Publications		Net Health Benefit vs. Comparators Type of Net Benefit (B, H) SOE		Impact of new studies (focus on comparative studies)*
		2014 Report†	2019 Report‡	2014 Report	2019 Report	2014 versus 2019 Report
						comparative cohort, inadequate evidence)
Gynecological	49.8	CS=2	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Head/Neck (oropharyngeal, nasopharyngeal, paranasl sinus, and oral cancers)	17.2§**	CC=1; CS=15; Econ=2	CC=7; CS=14; Econ=1	Insufficient low	Comparable B: = H: = Low	6 additional, larger, retrospective comparative cohorts lead to different conclusions
Head/Neck (Chondro-sarcoma of the skull base)		CC=1 CS=15	CC=1 CS=9	Insufficient low	Insufficient	Similar conclusions (1 small retrospective comparative cohort, inadequate evidence)
Liver	8.1	CC=3; CS=26	RCT=1; CC=1 CS=12; Econ=1	Comparable B: = H: = Low	<u>PBT vs. TACE</u> Incremental B: = H: ↓ Moderate <u>PBT vs. IMRT</u> Incremental B: = H: ↓ Low	RCT interim results with different comparator (TACE). Hospitalization was used as a surrogate for toxicity (see full report). PBT vs. IMRT, larger retrospective comparative cohort. Net health benefit vs. comparators across both reports is unclear.
Lung	60.5	CC=4; CS=19; Econ=2	RCT=1; CC=6††; CS=12	Comparable B: = H: = Low‡‡	Comparable B: = H: = Low	Similar conclusions; addition of a RCT and 5 retrospective comparative cohorts (1 large propensity score-matched database study).
Lymphomas	22.4	CS=1	CS=3	Insufficient none	Insufficient	Similar conclusions
Mixed/Various	N/A§	CC=3; CS=12	CS=3	NR	Insufficient	Similar conclusions
Ocular	0.9	RCT=1; CC=8; CS=45	CC=3; CS=22; Econ=1	Superior (Incremental) §§ B: ↑ H: ↓ Moderate	<u>PBT vs. BT alone</u> Inferior B: ↓ H: = Low	3 additional retrospective comparative cohorts (1 case-matched, and 1 large propensity score-matched database) with very different

Condition	Incidence (per 100000)	Number of Publications		Net Health Benefit vs. Comparators Type of Net Benefit (B, H) SOE		Impact of new studies (focus on comparative studies)*
		2014 Report†	2019 Report‡	2014 Report	2019 Report	2014 versus 2019 Report
					PBT + TSR vs. BT + TSR Incremental B: ↑ H: = Low PBT vs. SRS Insufficient	comparators. Prior report included primarily enucleation (4/7 studies) as comparator, also TTT (1 study); remaining 2 studies were indirect comparisons of case series. The net health benefit across all comparators (across both reports) is unclear.
Prostate	109.2	RCT=1; CC=9; CS=19; Econ=3	Quasi-RCT=1; CC=3; CS=11	Comparable B: = H: = Low‡‡	Comparable B: = H: = Low	Similar conclusions; addition of a quasi-RCT and 3 retrospective comparative cohorts (1 case-matched, 1 large propensity score-matched database)
Sarcomas	4.8§	CS=2	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Seminoma	4.0§	0	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Thymoma	0.2§	0	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Non-cancerous						
AVMs	1.0§	CS=6	0	Insufficient none	Insufficient (no studies)	Similar conclusions
Hemangiomas	2.0§	CC=1; CS=3	CS=2	Comparable B: = H: = Low	Insufficient	Similar conclusions
Pituitary Adenoma	NR§	CS=2	CS=1	N/A	Insufficient	Similar conclusions
Meningioma	2.0§	CC=2; CS=8	CS=3	Insufficient none	Insufficient	Similar conclusions
Pediatric						
Cancer						
All Cancer Types* **	18.3	CC=1; CS=41; Econ=3	CC=13; CS=41; Econ=2	Incremental B: = H: ↓ Low‡‡	See below	See below

Condition	Incidence (per 100000)	Number of Publications		Net Health Benefit vs. Comparators Type of Net Benefit (B, H) SOE		Impact of new studies (focus on comparative studies)*
		2014 Report†	2019 Report‡	2014 Report	2019 Report	2014 versus 2019 Report
Brain	3.1	---	CC=11; CS=25 Econ=2	N/A***	Incremental B: = H: ↓ Low	No comparative studies in the 2014 report; 6 new retrospective cohorts and 2 new prospective (1 propensity score-matched) cohorts suggest incremental net benefit of PBT for pediatric brain tumors
Bone	0.9	---	CS=1	N/A***	Insufficient	N/A
Head/Neck	NR§	---	CC= 1; CS=3	N/A***	Insufficient	N/A
Ocular	0.4	---	CC=1; CS=2	N/A***	Insufficient	N/A
Lymphoma	2.4	---	CS=2	N/A***	Insufficient	N/A
Rhabdomyo-sarcoma	NR§	---	CS=6	N/A***	Insufficient	N/A
Mixed/Various	NR§	---	CS=2	N/A***	Insufficient	N/A

AVM = Arteriovenous Malformation; **B = Benefits**; CC = Comparative Cohort; CS = Case Series; **H = Harms;** N/A = not applicable; IMRT = intensity-modulated radiation therapy (photons). NR = not reported; RCT = Randomized Control Trial; TTT = transpupillary thermotherapy.

*Due to lack of clarity in reported totals of studies, the study totals for the 2014 report here are derived from study lists in the appendix, and may differ from reported totals in body of report.

†All included studies were published subsequent to the prior report. Only studies that provided data on efficacy, effectiveness, safety or cost-effectiveness are included in this table (i.e., contextual studies are not included here).

‡When possible, incidence statistics were updated with more recent data from the National Cancer Institute's Surveillance, Epidemiology, and End Results Program (SEER) database and the American Cancer Society Cancer Statistics Center. Footnoted conditions were either obtained from the incidence values reported in the prior report, not acquirable through the reviewed databases (NR) or not applicable (N/A) because they represented a mixed population.

§Incidence is for head and neck cancers to include skull-base tumors (e.g., chondrosarcoma).

**The comparative cohort count includes the nonrandomized group from the RCT (Liao 2018).

††The prior 2014 PBT report had discrepancies between Table ES2 and Table 3 regarding the Strength of Evidence for Lung Cancer, Prostate Cancer, and Pediatric Cancers. AAI has made the decision to use the Strength of Evidence reported in Table ES2.

‡‡Authors of the 2014 report list the net health benefit as "superior" in their executive summary table. In the report body authors state "Limited, low-quality evidence suggests comparable rates of harm for PBT relative to treatment alternatives in patients with ocular tumors" which suggests that the net health benefit may be more incremental.

§§In the 2014 report, assessment of pediatric cancer was not separated by cancer types.

4.2 Pediatric Tumors

Overview across all pediatric studies

No comparative studies in pediatric populations were described in the 2014 report. No RCTs were identified for this update. Overall, 10 comparative cohort studies (across 13 publications)[15,31,32,71,72,89,95,131,132,148,221,245,265] and 41 case series (publications)[19,34,62,83,84,92,99,115,118-120,125,135,149-153,157,165,167,180,184-186,189,195,218,223,228,231,232,250,291,298,300,306,307,310,316,320] published subsequent to the prior report were identified in pediatric populations to address one or more of the first four Key Questions. Two full economic studies were also identified for Key Question 5 (Table 12). Most of the new studies identified were in patients with brain tumors. Eight of the comparative studies (11 publications)[31,32,71,72,95,131,132,148,221,245,265] reported on pediatric brain tumors. An overview of risk of bias concerns of included comparative studies is presented in the section "Risk of bias for included studies". All were considered to be at moderately high risk of bias. Risk of bias assessment for included comparative studies is found in Main Appendix D. All case series are considered at high risk of bias.

Results are organized by general tumor category. Key Questions 1-5 are addressed sequentially within each category.

Table 9. Overview of included studies in pediatric patients by tumor category

Tumor	Comparative, # studies (# publications)	Case series*, # publications
Bone	0	1[310]
Brain	8 studies (11 publications)[31,32,71,72,95,131,132,148,221,245,265] 2 Economic[109,170]	25[19,34,62,83,84,92,99,118-120,125,135,149-151,167,180,189,218,228,232,250,291,307,320]
Head and neck	1 (Safety)[89]	3[165,231,300]
Lymphoma	0	2[115,316]
Ocular	1 (Salvage)[15]	2[195,223]
Soft-tissue (sarcoma)	0	6[152,153,157,186,298,306]
Various/mixed	0	2[184,185]

*There is overlap in patient populations across case series for pediatric brain and various/mixed tumor types

Key points (across all pediatric tumors)

- **Pediatric brain tumors**:
  - The bulk of the comparative evidence from studies published subsequent to the 2014 report was for the use of PBT in various pediatric brain tumors. Eight comparative cohort studies at moderately high risk of bias compared PBT with treatment alternatives.
    - Three studies compared PBT with IMRT
    - Two studies compared patients who received PBT with those who received IMRT and/or 3DCRT

        - One study indicated PBT was compared to photon RT with no further specification and one study indicated that those in the comparison group received either 2DCRT or 3DCRT
        - One study compared craniospinal PBT and focal PBT with surgery.
    - Benefits in terms of OS, PFS and tumor recurrence were generally similar between PBT and other forms of radiation therapy across four comparative studies (Low SOE). Some differences may be clinically important.
    - Regarding toxicities and harms, hypothyroidism was less common with PBT versus other RT. Low SOE) Many other toxicities (including other endocrine-related toxicities) tended to be less frequent in those receiving PBT vs other RT, however statistical significance was generally not reached, likely due to study sample sizes and possibly residual confounding. (Low SOE) Some differences may be clinically important. One prospective cohort study reported declines for full scale intelligence quotient (FSIQ) and processing speed index scores when craniospinal PBT was compared with surgery but no differences between focal PBT and surgery for any score. The clinical relevance of the declines was not described. One retrospective cohort reported no difference between PBT and photon therapy for FSIQ scores (Low SOE for all outcomes.)
    - While two poor-quality full economics studies suggest that PBT may be cost-effective for treatment of pediatric brain or CNS tumors vs other types of radiation, the limitations of these studies need to be considered.
    - None of the included studies evaluated differential effectiveness or safety.
- **Other pediatric tumors:**
    - Evidence for effectiveness and safety was considered to be insufficient for all other pediatric tumors. Studies published subsequent to the 2014 report were identified for the following pediatric tumor categories: head and neck, soft tissue (rhabdomyosarcoma), ocular, lymphoma, bone and one study of mixed tumor types. Evidence was primarily from case series, with only two small comparative (one for salivary gland tumors, the other salvage treatment in ocular tumors) identified.
    - No full-economic studies or studies designed to evaluate differential effectiveness or safety were identified.
- **Comparison with 2014 report**.
    - The 2014 report identified only case series, with exception of one poor quality comparative study of secondary cancer in patients who had undergone RT for ocular tumors. The 2014 report generalized results across pediatric tumors and concluded that the net health benefit for PBT was considered to be incremental versus other forms of radiation therapy based on theoretical considerations that benefits would be comparable but harms of PBT would be lower compared with other forms of RT.
    - This updated report focuses on results from 10 comparative studies published subsequent to the previous report, all but two of which compared PBT with other forms of radiation therapy or surgery for treatment of pediatric brain tumors. The overall body of evidence for PBT use in pediatric tumors in this update is of somewhat higher quality based on the availability of comparative studies compared with the previous report. Most studies compared PBT with more contemporary types of RT, specifically IMRT.
        - For pediatric brain tumors, the overall health benefit for PBT was considered to be incremental versus other forms of radiation therapy: benefits (i.e. OS, PFS) were considered to be comparable between treatments but toxicities/harms tended to be less common following PBT vs. other forms of RT. (SOE LOW).

- While two additional small comparative studies were identified, one employing PBT for salvage therapy for ocular tumors and the other employing PBT for treatment of rare salivary gland tumors, the overall strength of evidence was considered to be insufficient.

4.2.1 Brain, Spinal, Paraspinal Tumors

Key points

- Eight cohort studies (11 publications) compared PBT with other treatment alternatives (primarily other forms of radiation therapy) in persons with pediatric brain tumors with a curative intent; all were considered to be at moderately high risk of bias.
    - Across the four small comparative studies (6 publications) that provided data on effectiveness there were no statistically significant differences in OS at any time point which may be a reflection of sample sizes and/or residual confounding. Some differences may be clinically important. (Low SOE)
    - Across the seven comparative studies (10 publications) that reported on toxicities and harms, risk of hypothyroidism and other endocrine toxicities tended to be lower with PBT versus other forms of radiation, however statistical significance was not generally achieved, in part due to small sample sizes; the role of residual confounding may also contribute. Some differences may be clinically important. One prospective cohort study reported declines for full scale intelligence quotient (FSIQ) and processing speed index scores when craniospinal PBT was compared with surgery but no differences between focal PBT and surgery for any score. The clinical relevance of the decline was not described. Another retrospective cohort reported no difference in FSIQ scores between PBT and photon radiation therapy. (Low SOE)
- Economic: Two poor quality full economic studies, one in patients with pediatric medulloblastoma, the other which included various CNS tumors, suggest that PBT may be cost-effective versus conventional radiotherapy.
- There were no comparative studies in patients with spinal or paraspinal tumors.
- No studies evaluated differential effectiveness or safety
- Limited information from case series does not provide sufficient information to evaluate effectiveness or radiation safety of PBT. (SOE Insufficient for all case series)

The 2014 report included only case series for this tumor category and results across all pediatric tumors were generalized. This updated review provides higher quality new evidence from seven comparative studies where PBT was compared with newer forms of RT such as IMRT in most studies for treatment of pediatric brain tumors.

Key Question 1 (Effectiveness, curative intent)

Description of included studies

Four comparative cohort study populations (five publications) that compared PBT for curative intent with other radiation therapies for ependymoma[95,245], meduloblastoma[72,148], and craniopharyngioma[32] reporting on effectiveness were identified. One prospectively enrolled patients for PBT as part of a phase II study, however the control group was from a separate institution.[72] Five additional comparative studies reported on toxicities/safety only[31,131,132,221,265] and are described further in Table 11 below. Tumors in all comparative studies were in the pediatric brain. In addition, 25 case series (publications) of PBT across various tumor types were identified. [19,34,62,83,84,92,99,118-120,125,135,149-151,167,180,189,218,228,232,250,291,307,320] Summaries of case series are found in the appendices.

Across the four comparative studies reporting on effectiveness of PBT for curative intent, three were clinical cohorts (N range 52 to 88) and one an administrative data study (N=783 with data on outcomes of interest). Reported ages ranged substantially from 2.5 years old to 9 years old and the proportion of males ranged from 43% to 67%. (Table 10 and Table 11) Comorbidities or presenting symptoms were reported in one study. Surgical intervention prior to radiation therapy was reported for all patients in the clinical cohort.

Two studies compared PBT to IMRT[32,245], one study compared PBT with a group that received either IMRT or 3DCRT[72]; the administrative data study only provided comparative data for PBT versus other forms of radiation therapy (2DRT, 3dCRT, or IMRT) [148]. Across studies, median total radiation doses ranged from 23 to 57 Gy for PBT and 23 to 54 Gy for other radiation therapy forms.

Two publications on the same underlying population of patients in the youngest age group (2.5 years old) with ependymoma compared PBT with IMRT and reported on different outcomes, one focusing on MRI evaluation[95] (N = 72) and the other focusing on primary clinical outcomes [245](N= 79). Across these reports 14 to 15% of PBT versus 20% to 24% of IMRT recipients had chemotherapy prior to radiation therapy. Most tumors were infratentorial and Grade III anaplastic (>80%) tumors; gross total tumor resection was slightly higher for PBT recipients (93-97%) than for IMRT recipients (76%-80%).

The two comparative studies in pediatric patients with medulloblastoma used substantially different study methods, one used clinical data from separate institutions[72] and the other an administrative data/registry study.[148] In the cohort[72], following primary tumor resection, all participants received craniospinal irradiation (CSI), chemotherapy and either involved field (IF) or posterior fossa (PF) boost radiation therapy (RT). Discrepancies between groups included age (6.2 years old vs. 8.2 years), tumor histology, residual tumor following surgery and location of RT boost. Authors report results for a propensity matched cohort; however, details of propensity matching and balancing verification were not provided. The administrative/registry data study compared PBT (n=117), IMRT (n=157) and 2D/3DCRT (n=1003), however survival analysis included only patients diagnosed prior to 2010 (N=783 across all treatments, numbers of patients per treatment are not provided for this subset of data). All patients received chemotherapy, but information on prior surgical intervention was not reported. Factors associated with increased likelihood of PBT use included higher median income, private insurance status, and higher education. Histology was not correlated with likelihood of receiving PBT and younger age (< 3 years) was marginally insignificant. Only income remained a significant predictor of receiving PBT in multivariate analysis of sociodemographic factors.

The fourth retrospective cohort study in pediatric patients with craniophyaryngioma compared PBT with IMRT.[32] As the majority (57%) of patients received PBT for definitive or adjuvant treatment it is included in KQ1. Authors report no differences by RT intent (salvage vs. definitive or adjuvant) were observed in 3-year OS, CFFS, or NFFS rates (p=0.294 OS, p=0.412 CFFS, and p=0.951 NFF). Differences in presenting symptoms between treatment groups included presence of headache (76% vs. 48%), visual defects (52% vs. 81%) and endocrinopathies (19% vs. 39%) as well as radiation intent (post-operative 38% vs. 48%, definitive treatment 19% vs. 10%).

All studies were considered to be at moderately high risk of bias.

Table 10. Study and patient characteristics from comparative studies reporting on effectiveness only and effectiveness and safety: Pediatric brain tumors

	Effectiveness only			Effectiveness and Safety							
Author (year)	Kopecky 2017			Bishop 2014		Eaton 2016a/2016b†		Gunther 2015‡		Sato 2017‡	
Characteristics	PBT (n=117)**	Photon RT (n=157)**	Photon RT (n=1003)**	PBT (n=21)	Photon RT (n=31)	PBT (n=45)	Photon RT (n=43)	PBT (n=37)	Photon RT (n=35)	PBT (n=49)	Photon RT (n=38)
Patient Characteristics											
Males, % (n)	55%	67%	66%	43%	45%	56%	67%	59%	54%	61%	55%
Age, years; median (range)	8.4 (0 to 18)			9.1 (NR)	8.8 (NR)	6.2 (3.3 to 21.9)*	8.2 (3.4 to 19.5)*	2.62*	6.08*	2.5 (0.5 to 18.7)*	5.7 (0.4 to 16.5)*
Comorbidities/Presenting symptoms											
Headaches	---	---	---	76%*	48%*	---	---	---	---	---	---
Visual defects	---	---	---	52%	81%	---	---	---	---	---	---
Endocrinopathies	---	---	---	19%	39%	---	---	---	---	---	---
Tumor Characteristics											
Subtype	Medulloblastoma			Craniopharyngioma		Medulloblastoma		Ependymoma		Ependymoma	
Radiation Treatment											
Radiation Intent											
Curative	100%			57%		100%		100%		100%	
Salvage/Recurrence	0%			42%		0%		0%		0%	
Technique	---	IMRT	2DCRT or 3DCRT	Passive Scatter	IMRT	3D Conformal	IMRT or 3DCRT	---	IMRT	---	IMRT
Tumor Bed Boost	---	---	---	---	---	62%	54%	---	---	---	---
Posterior Fossa Boost	---	---	---	---	---	29%	27%	---	---	---	---
Posterior Fossa & Tumor Bed Boost	---	---	---	---	---	9%	20%	---	---	---	---
Median total dose (Gy)	54			50.4	50.4	23.4	23.4	57.2	55.9	55.8	54
Additional Treatments											
Prior to Radiation											
Gross Total Resection	---	---	---	24%	3%	---	---	97%	80%	93%*	76%*
Sub-total Resection	---	---	---	43%	35%	---	---	3%	20%	---	---
Any resection	---	---	---	---	---	100%	100%	---	---	---	---
Cyst drainage, fenestration, shunting	---	---	---	33%	61%	---	---	---	---	---	---

	Effectiveness only			Effectiveness and Safety							
Author (year)	Kopecky 2017			Bishop 2014		Eaton 2016a/2016b†		Gunther 2015‡		Sato 2017‡	
Chemotherapy	---	---	---	---	---	---	---	14%	20%	15%	24%
Adjunctive/Concomitant											
Chemotherapy	100%	100%	100%	---	---	100%	100%	14%	0%	---	---
Study Design	**Retrospective Comparative Cohort**			**Retrospective Comparative Cohort**		**Prospective Comparative Cohort**		**Retrospective Comparative Cohort**		**Retrospective Comparative Cohort**	
Follow-up, months (% followed)	**54 (60.2%††)**			**33.1 (CD§)**	**106.1 (CD§)**	**74.4 (CD§)**	**84 (CD§)**	**40.6 (CD§)**		**31.2 (CD§)**	**58.8 (CD§)**
Risk of Bias	**Moderately High**			**Moderately High**		**Moderately High**		**Moderately High**		**Moderately High**	

3DCRT = three-dimensional conformal radiotherapy; CD = cannot be determined; Gy = gray; IMRT = intensity modulated radiation therapy; NR = not reported; PBT = proton beam therapy
*Indicates that there is statistically significant difference between the two groups
†Study and patient characteristics are drawn from Eaton 2016a
‡Gunther 2015 and Sato 2017 appear to be in the same study population overall with each publication reporting on different outcomes
§Follow-up and differential loss to follow-up cannot be determined (number eligible not provided, number excluded and loss to follow-up not described)
**Analyses of interest for this report were only reported in 783 patients, demographic data is provided for the larger group of 1277 patients; differential follow-up cannot be determined
††Authors account for legitimate patient exclusions to obtain 1300 patients meeting inclusion criteria, however an additional 517 were excluded from survival analysis based on diagnosis after 2009 based on database use guidelines to allow at least 5 years follow-up for all patients but do not indicate to which treatment groups the 517 belonged

Table 11. Study and patient characteristics from comparative studies reporting on safety only: Pediatric brain tumors

	Safety Only										
Author (year)	**Bielamowicz 2018****		**Paulino 2018****		**Song 2014**		**Kahalley 2016**		**Kahalley 2019**		
Characteristics	PBT (n=41)	Photon RT (n=54)	PBT (n=38)	Photon RT (n=46)	PBT CSI (n=30)	Photon CSI (n=13)	PBT (n=90)	Photon RT (n=60)	PBT CSI (n=22)	PBT focal RT (n=31)	Surgery (n=40)
Patient Demographics, %											
Males, % (n)	75.9%	68.3%	74%	70%	53%	62%	60%	55%	59.1%	45.2%	52.5%
Age, years; median (range)			7.6 (2.9 to 14.5)	9 (3 to 18)	10 (2 to 18)	11 (3 to 18)	Mean: 9.2 (1.7 to 18.2)	Mean: 8.1 (1.2 to 18)	10 (2.2 to 17.8)	8.4 (1.0 to 16.5)	9.3 (2.2 to 18.6)
Comorbidities											
Posterior Fossa Syndrome	---	---	13%	15%	---	---	---	---	40.9%*	3.2%*	7.5%*
Tumor Characteristics, %											
Subtype											
Medulloblastoma	100%	100%	100%	100%	100%	100%	---	---			
Ependymoma	---	---	---	---	---	---	4%	22%	0%*	19.4%*	0%*
Medulloblastoma/PNET	---	---	---	---	---	---	38%	47%	77.3%*	3.2%*	0%*
Glioma	---	---	---	---	---	---	22%	13%	4.5%*	51.6%*	80%*
Germ Cell Tumor	---	---	---	---	---	---	19%	5%	13.6%*	9.7%*	0%*
Other	---	---	---	---	---	---	17%	7%	4.5%*	3.2%*	10%*
Radiation Treatment											
Treatment Intent											
Curative	100%		100%		72%		100%		100%		
Salvage/Recurrent	0%		0%		28%		0%		0%		
Technique	Passive Scatter	3DCRT + IMRT boost	Passive Scatter	3DCRT + IMRT boost	---	---	Passive scatter: 90% PBS: 10%	3D-CRT: 8.3% IMRT: 45% 3DCRT+IMRT boost: 46.7%	---	---	---
Posterior Fossa Boost	---	---	0%	13%	---	---	---	---	---	---	---
Posterior Fossa + Tumor Bed Boost	---	---	0%	63%	---	---	---	---	---	---	---
Tumor Bed Boost	---	---	100%	24%	---	---	---	---	---	---	---
Median total dose (Gy)	Mean: 55.3	Mean: 55.4	Range, 54 to 55.8	Range, 54 to 55.8	51.8	53.2	Mean: 54	Mean: 54	54	50.4	---

Final

	Safety Only										
Author (year)	**Bielamowicz 2018****		**Paulino 2018****		**Song 2014**		**Kahalley 2016**		**Kahalley 2019**		
Additional Treatments											
Prior to Radiation Treatment											
Craniotomy	---	---	---	---	---	---	87%	97%	---	---	---
Any tumor resection	100%	100%	100%	100%	---	---	---	---	---	---	100%
Cyst drainage, fenestration, shunting	---	---	34%	52%	---	---	---	---	---	---	---
Chemotherapy	---	---	---	---	87%	77%	---	---	0%	0%	0%
Concurrent/Adjuvant											
Chemotherapy	100%	100%	100%	100%	---	---	---	---	---	---	---
Timing NOS											
Shunt Placement	---	---	---	---	---	---	30%	50%	40.9%*	3.2%*	7.5%*
CSI Radiation	---	---	---	---	---	---	57%	52%	---	---	---
Study Design	**Retrospective Comparative Cohort**		**Retrospective Comparative Cohort**		**Retrospective Comparative Cohort**		**Retrospective Comparative Cohort**		**Prospective Comparative Cohort**		
Follow-up, months (% followed)	**56.4* (CD§)**	**121.2* (CD§)**	**55.5 (86%)**	**65.5 (73%)**	**55.5 (CD§)**	**65.5 (CD§)**	**32.4* (all pts, 73%)**	**64.8* (all pts, 73%)**	**NR (74.5%)**		
Risk of Bias	**Moderately High**		**Moderately High**		**Moderately High**		**Moderately High**		**Moderately High**		

2DCRT = two-dimensional conformal radiotherapy; 3DCRT = three-dimensional conformal radiotherapy; CD = cannot be determined; Gy = gray; IMRT = intensity modulated radiation therapy; NR = not reported; PBS = Pencil Beam Scanning; PBT = proton beam therapy; PNET = Primitive neuroectodermal tumor; pts = patients

*Indicates that there is statistically significant difference between the two groups

†Analyses of interest for this report were only reported in 783 patients, demographic data is provided for the larger group of 1277 patients; differential follow-up cannot be determined

‡Authors account for legitimate patient exclusions to obtain 1300 patients meeting inclusion criteria, however an additional 517 were excluded from survival analysis based on diagnosis after 2009 based on database use guidelines to allow at least 5 years follow-up for all patients but do not indicate to which treatment groups the 517 belonged

§Follow-up and differential loss to follow-up cannot be determined (# eligible not provided, patient selection methods not clear)

**Paulino 2018 and Bielamowicz 2018 are from the same underlying population.

Results

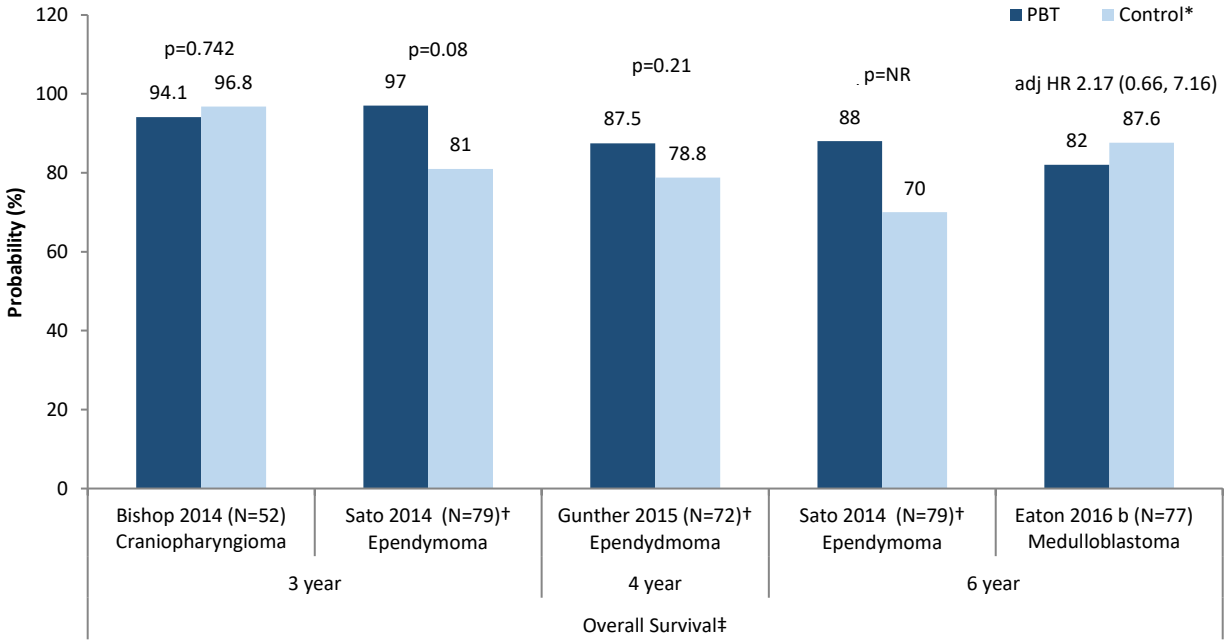
Survival outcomes and recurrence

In pediatric patients with brain tumor, for probability of overall survival, PBT resulted in similar OS (three studies, Bishop, Kopecky, Eaton) or slightly greater (one study, two citations, Sato, Gunther) OS compared with IMRT or CRT but no statistically significant differences were observed across tumor types or time frames. Small sample sizes may have contributed to the failure to find statistical differences in most of the studies; residual confounding may also impact results. Figure 4 summarizes reported survival probability for four of the studies. The fifth study, Kopecky 2017, reported an unadjusted HR of 0.99 (95% CI 0 .41 to 2.4) for the comparison of PBT with 2D/3D-CRT. Figure 4 and Abstraction Appendix N.

Progression-free survival (PFS) in one study of patients with medulloblastoma[245] tended to be better following PBT versus IMRT at 3 (82% vs. 60%) and 6 years (82% vs. 38%), but statistical significance was not reached at 3 years and not reported for 6 years (survival estimated from author graph). Recurrence-free survival was similar between PBT and IMRT in one study in patients with medulloblastoma[72] (78.8% vs.76.5%, adjusted HR 1.13 (95% CI 0.5, 3.41). Figure 5 and Abstraction Appendix N.

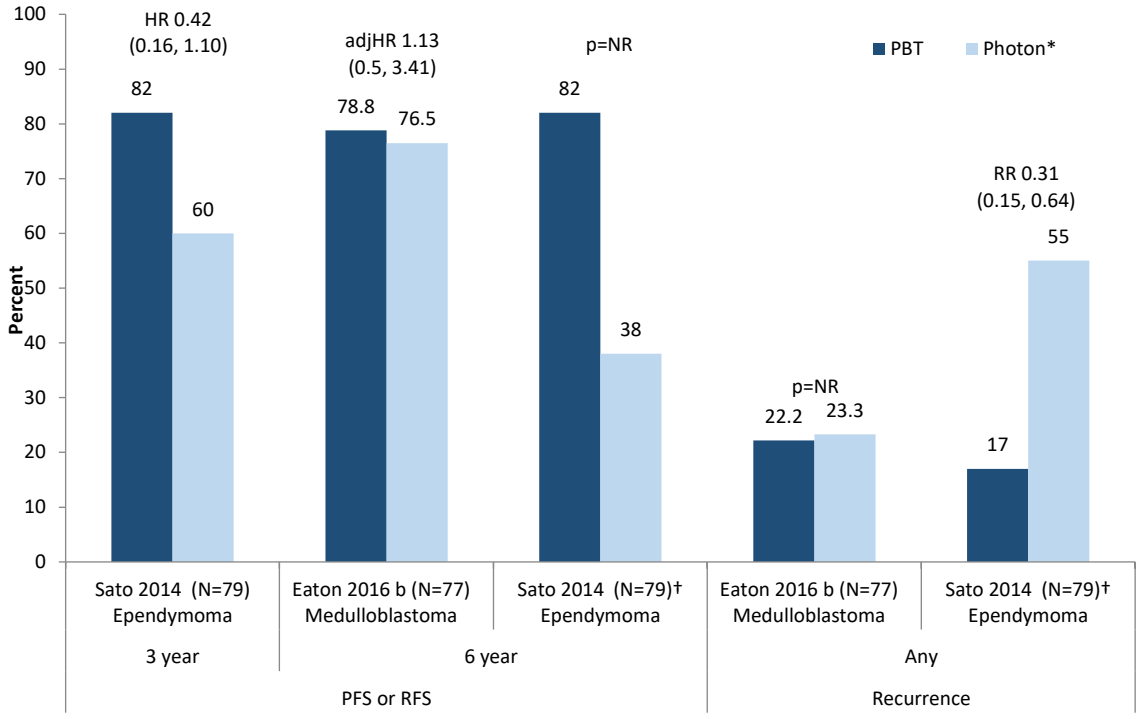
Lower disease-related mortality for those receiving PBT (4.9%) versus IMRT (31.6%) was reported in one study in patients with ependymoma.[245] Mortality was not reported in the other studies.

Figure 4. Probability of overall survival in comparative studies of PBT versus other type of radiation therapy in children with brain tumors



*PBT was compared with IMRT in Bishop, Gunther and Sato and with either 3D CRT or IMRT in Eaton
†Sato and Gunther report on the same underlying patient population. Sato 6 year estimates from author's graph
‡One additional study (Kopecky 2017) reported a HR of 0.99 (95% CI 0 .41 to 2.4) for the comparison of PBT with CRT but

Figure 5. Progression-free survival and tumor recurrence in comparative studies of PBT versus other type of radiation therapy in children with brain tumors



*PBT was compared with IMRT in Sato and with either 3D CRT or IMRT in Eaton
†6 year estimates for Sato estimated from author's graph

Across case-series, overall survival following PBT for brain tumors at 2 years ranged from 65% (in two small series of atypical teratoid rhabdoid tumor (ATRT)) to 90.5% in one larger series in patients with a variety of tumor types. OS was generally high at all time other time frames (range 83% to 100%) across case series enrolling patients with a variety of tumors. At 2 years, the two small series reported PFS ranging from 48% to 66% in patients with ATRT. Probabilities for progression-free survival were somewhat lower than for OS across other tumor types for other time frames (Main Appendix F, Table F58).

Findings across case series for mortality and progression/relapse/treatment failure are in Main Appendix F, Table F59.

Other outcomes – comparative studies

Recurrence or relapse was reported in one comparative study of pediatric medulloblastoma and was similar following PBT (22.2%) and 3DCRT or IMRT (22.3%).[72] In one study of pediatric ependymoma, recurrence was markedly lower following PBT (7/41 or 17%) versus IMRT (21/38 or 55%), p = 0.005 (Figure 5).[245]

Health related quality of life (HRQoL) measured via the parent-proxy version of the Peds QL core module was reported in one study for which the PBT and photon RT groups were from separate institutions[320], thus the comparison is considered indirect. The common tumor type in both groups was

medulloblastoma, followed ependymoma/high-grade glioma. Across tumor types, the total domain score as well as summary domain scores for physical, psychological, emotional and social functioning were higher in the PBT group compared with the photon group. Conclusions are however limited as the patients are from different institutions.

Key Question 3 (Safety)

Description of included studies

Seven comparative cohort studies at moderately high risk of bias compared PBT other forms of radiation therapy for treatment of pediatric brain tumors and reported on a variety of toxicities, harms and/or adverse events across various tumor types.[31,32,71,131,221,245,265]. An eighth prospective longitudinal cohort study compared focal and CSI PBT with surgery and reported on neurocognitive outcomes.[132] Tables 10 and 11 in KQ 1 above provide a summary of all comparative studies in pediatric patients with brain tumors. Five of the studies did not report data on effectiveness (KQ 1 or 2) and reported only safety outcomes[31,131,132,221,265]. In addition, 20 case series[19,34,62,83,84,92,99,118-120,149-151,167,180,189,232,291,307,320] provided data on safety outcomes (Main Appendix F, Tables F60 to F68). Detailed data abstraction for case series is found in Abstraction Appendix N.

Across the eight comparative cohort studies that provided data on safety, four included only patients with medulloblastoma[31,71,221,265], one included only patients with ependymoma[245], one included only patients with craniophayngioma[32] and two included different tumor types[131 132](ependymoma, glioma, germ cell, and others). In most studies, there were substantial differences between treatment groups with regard to the length of followup that should be considered when interpreting longer-term outcomes. Sample sizes were small to moderate (N = 43 to 93) and ages ranged from 2.5 years old to 11 years old at time of diagnosis/treatment across studies. (Table 10 and Table 11)

Across the four studies in patients with medulloblastoma, sample sizes ranged from 43 to 88, mean ages ranged from 6.2 years to 11 years with significant differences in age by treatment group in 2 studies[71,221] and the proportion of males ranged from 53% to 76%. All patients in three studies received concurrent or adjunctive chemotherapy[31,71,221] and 87% of PBT recipients vs. 77% of photon therapy recipients had chemotherapy prior to radiation in the third study[265]. Locations receiving boost radiation were similar between treatment groups in one study[71], but differed between treatment groups in the other study that reported on boost location; in the latter study all patients in the PBT group received boost to the tumor bed vs. only 24 % of those in the 3DCRT plus IMRT group.[221] Reported median total PBT radiation dose ranged from 23 Gy in one study[71,72] to between 53 Gy[265] and 56 Gy[221] in other studies.

One retrospective cohort study (N = 150) compared PBT with photon RT which included, three-dimensional conformal (8.3%), IMRT (45.0%), or three-dimensional conformal plus IMRT tumor bed margin boost (46.7%) in a population with various tumor types.[131] The population was predominately male with mean age at time of RT of 9 years old; ependymoma was the most common tumor type overall. Tumor type/histology, mean Lansky/Karnofsky performance scores and use of craniotomy and ventriculoperitoneal shunt differed between treatment groups. Total median radiation was 54 Gy for both treatment groups. A prospective cohort study by the same author (N-93) compared CSI PBT, focal PBT and surgery only in a population with various tumor types[132] was 52% male with a mean age at diagnosis of 9 years old. There was a substantial difference in tumor histology across the three treatment groups; All but one patient with medulloblastoma/primitive neuroectodermal tumors received CSI PBT and none had surgery. Similarly none of the patients with ependymoma or germ cell

tumors had surgery. Conversely, the majority of patients with glioma received surgery. Authors also report significant differences between groups with regard to tumor location, use of ventriculoperitoneal shunt and presence of posterior fossa syndrome. Authors statistically adjusted for these variables. In the PBT groups, total RT dose to the tumor was 54 Gy for CSI higher versus and 50.4 GY for focal PBT. Baseline neurocognitive scores did not differ between treatment groups and authors report that. (Table 11)

Two publications on the same underlying population of patients with ependymoma from the same institution reported on different outcomes.[95,245] One focused on MRI evaluation[95] (N = 72) and correlation with patient symptoms; data for this publication are found in the appendices. The other reported data on primary safety/toxicity outcomes[245] (N= 79) and the findings are included in this section. Differences between treatment groups in age (2.5 years vs. 5.7 years), frequency of total gross resection (93% vs. 76%), proportion of patients receiving chemotherapy and length of follow-up are noted. Similarly two publications on the same underlying population of patients with medulloblastoma at another institution reported on different toxicities; one reported on hypothyroidism[31] (N=95) and the second on other radiation-related toxicities [221](N=84). Both compared passive scatter PBT with 3DCRT with IMRT boost and are included in the description above.

In the eighth study in patients with craniopharyngioma (N=52), mean age was around 9 years old with slightly more females enrolled.[32] Differences between treatment groups with regard to presence of headache, visual defects and endocrinopathies at baseline and treatments prior to radiation are noted and it is not clear if these were included in adjustment for confounding.

Results

Across comparative studies, the frequency of most toxicities and adverse events was similar to or lower in patients receiving PBT compared with other forms of radiation therapy but statistical significance was not uniformly reached; some differences may be clinically important. All studies were at moderately high risk of bias; adjusted effect size estimates were reported for some outcomes.
We attempted to focus on common Grade 3 or 4 toxicities that might be attributed to RT and evidence from studies comparing PBT to other forms of RT. In the absence of such a comparison, firm conclusions regarding improvements in radiation-related safety of PBT are limited.

Endocrine-related (late) toxicities

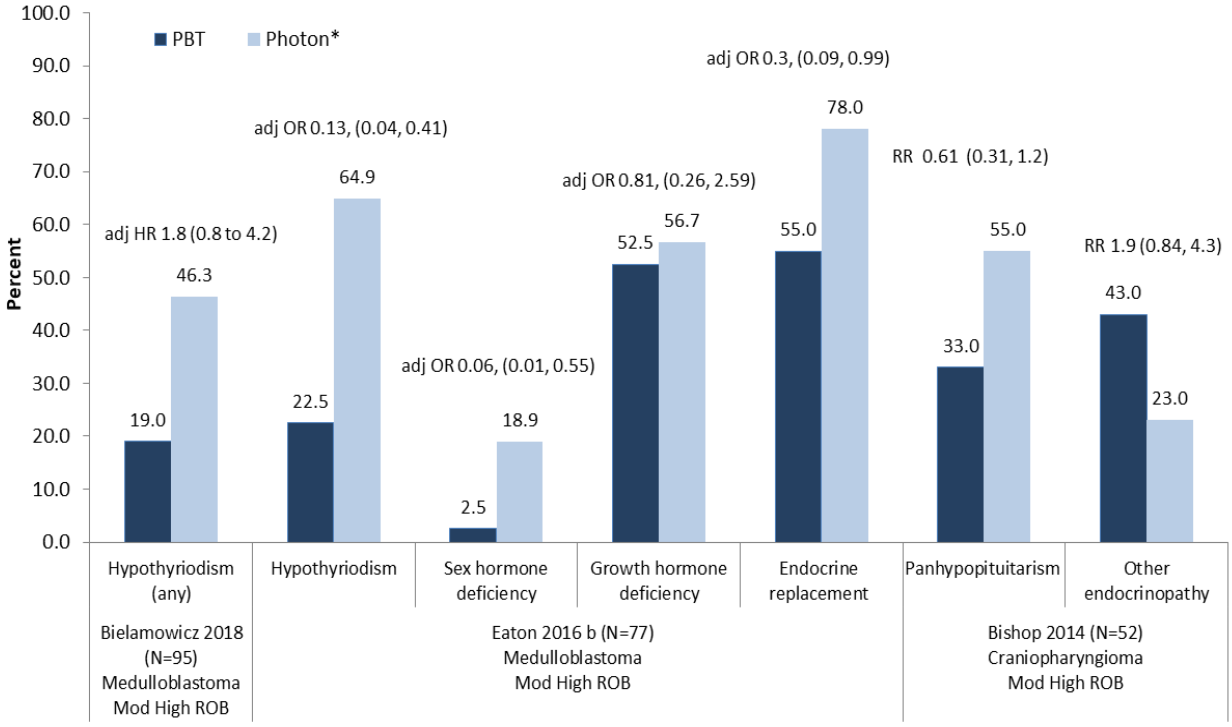
Endocrine-related (late) toxicities were reported in three studies.[31,32,71] Overall, most endocrinopathies across the three studies were less common in PBT recipients compared with other forms of radiation therapy.

In two studies in patients with medulloblastoma, hypothyroidism was less common following PBT compared with 3DXRT with IMRT[31] or with 3DCRT or IMRT, however statistical significance was only seen in one study[71,72] which may be function of sample size (Figure 6). In addition to reporting on any hypothyroidism, the most recent study also reported lower frequency of primary hypothyroidism (7.3% vs. 20.4%, adj HR 2.1, 95% CI 0.6 to 7.7) and central hypothyroidism (9.8% vs. 24.0% ,adj HR 2.2, 95% CI 0.7 to 6.6)[31] following PBT versus 3DXRT with IMRT boost.

With regard to other endocrine-related toxicities, one study in patients with medulloblastoma[71] found that PBT was associated with significantly lower risk of sex hormone deficiency and need for endocrine

replacement in multivariate analysis compared with other modes of radiation therapy delivery (IMRT or 3DCXRT) however, no association between PBT and lower risk of growth hormone deficiency was found. In the study of patients with craniopharyngioma, panhypopitiutitarism was less common in PBT recipients (33%) versus those receiving IMRT (55%) but other types of endocrinopathy (including growth hormone deficits, adrenal insufficiency, sexual hormone deficiencies) were somewhat more common with PBT (43% vs. 23%), however, baseline differences in the presence of endocrinopathies and it is unclear if analysis included adjustment for this. The small sample size may partially explain failure to reach statistical significance and observation of high percentages.[32] (Figure 6)

Figure 6. Endocrine-related (late) toxicities and adverse events reported in comparative studies of PBT



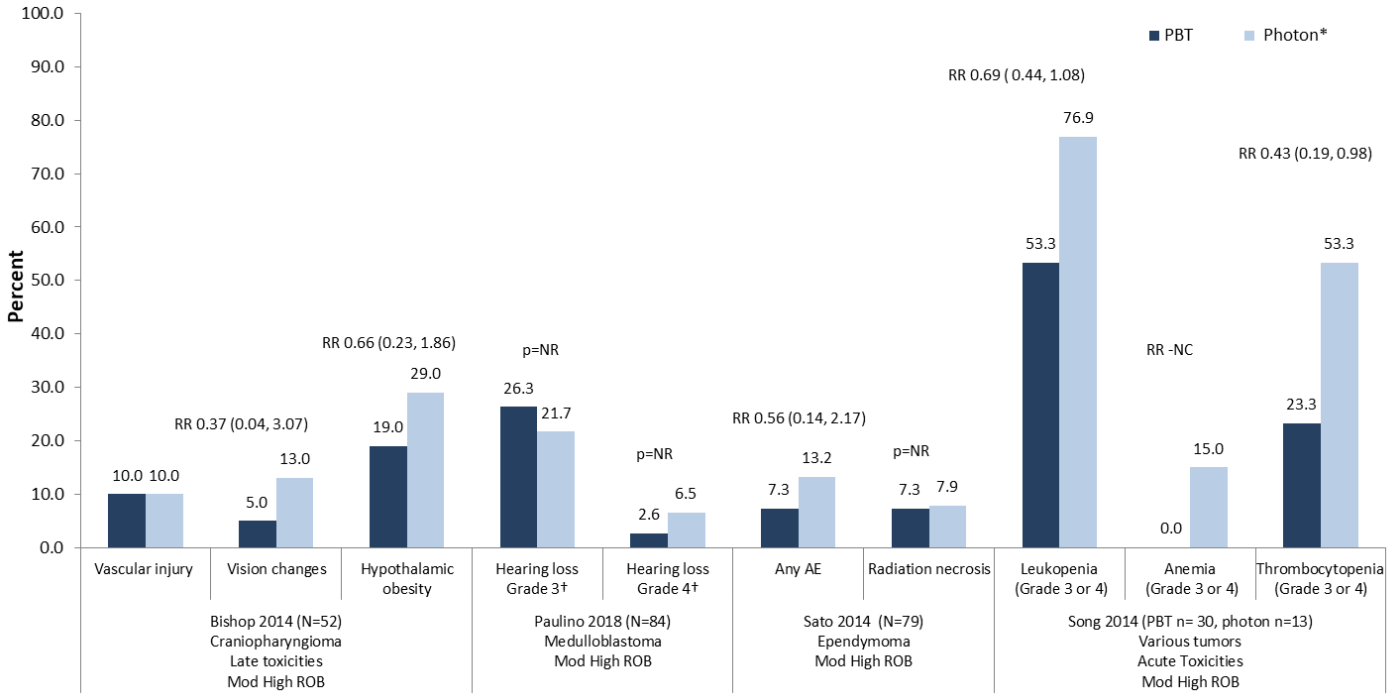
*PBT was compared with3D CRT with IMRT boost in Bielamowicz with IMRT in Bishop and with either 3D CRT or IMRT in Eaton

Across four case series of PBT, endocrine-related toxicities secondary to PBT were common. One large series (N = 179) in patients with ependymoma[118] follow-up of 38 months reported late Grade 2 + hormone deficiency in 7.3% of patients with growth hormone deficiency being most common. The next largest series (N = 59) in patients with medulloblastoma[320] reported 3, 5 and 7 year cumulative incidence of any hormone deficiency as 27% (16-39%), 55% (41% to 67%) and 63% (48% to 75%) with growth hormone deficiency and thyroid insufficiency being most common. In one small series (n= 29)[92] in patients with low-grade glioma a rate of 50% for any endocrine deficiency at 10 years was estimated. One small series (N=25) reported growth hormone deficiency in 8% (2/25) of patients with ependymoma[167]. (Main Appendix F, Table F60)

Other reported toxicities/harms

There were no commonly reported outcomes across comparative studies. The frequency of most toxicities were similar for PBT and photon therapy; for those that were less common in the PBT group, statistical significance was not reached for any outcome (Figure 7).

Figure 7. Toxicities and adverse events reported in comparative studies of PBT versus other type of radiation therapy in children with brain tumors



* Bishop and Sato PBT vs. IMRT; Paulino, passively scattered PBT vs. 3DCRT with IMRT boost; Song PBT vs. photon (not specified); †CTCAE grade; crude RR calculated by AAI.

One small cohort (n = 30 PBT, n=13 photon)[265] enrolling patients with various brain tumor types reported on acute toxicities; 72% of patients received radiation therapy for curative intent, thus the study is included here in KQ 1. Although the frequencies of grade 3 or 4 leukopenia, anemia and thrombocytopenia was lower less following PBT, the study evaluated 30 patients in the PBT group but only 13 patients in the photon therapy group and none of the differences were statistically significant; 84% of patients had some form of chemotherapy which may have contributed to toxicity. (Figure 7)

Acute hematologic toxicities reported in the largest case series in patients with medulloblastoma (N= 59, 52 had concurrent chemotherapy) were as follows: Neutropenia Grade 3 (32%) and Grade 4 (8%), Grade 3 anemia (5%), Grade 3 thrombocytopenia (3%), lymphopenia Grade 3 (17%) and Grade 4 (12%).[320] (Main Appendix F, Table F61)

With regard to late toxicities, in one small comparative cohort study in patients with ependymoma (N = 42)[32], the frequency of vascular injury was similar for PBT and IMRT, but vision changes and hypothalamic obesity were less common with PBT; although statistical significance was not reached due in part to sample size, differences may be clinically important. Grade 3 and Grade 4 CTCAE hearing loss were not statistically different for PBT and photon therapy in a cohort of patients with ependymoma

(N=84)[221] and frequency of radiation necrosis was similar in a cohort of patients with ependymoma (N=79)[245]. This same study reported similar frequency of any adverse event (7.3% or 3/41 patient vs. 13.2% or 5/38 patients) for PBT and IMRT and that no strokes or cavernomas occurred in the PBT group compared with 1 each in the IMRT group. Sample sizes in all studies may preclude detection of rare events. (Figure 7)

Vascular injury (case series)

The three-year cumulative rate of serious vasculopathy in the largest case series (N=644) was 2.6% and the frequency of stroke with permanent neurological deficits was 1.2%.[99] Late Grade 2+ vasculopathy occurred in 3.4% (6/179) of patients in the next largest series.[118] This is lower than the 10% for both PBT and IMRT described in the small comparative study by Bishop (Main Appendix F, Table F62).

Brainstem injury/necrosis (case series)

One comparative study (N= 79) in patients with ependymoma reported similar frequency of radiation necrosis between PBT and IMRT (7.3% and 7.9% respectively, Figure 7). One small series reported late Grade 3 radiation necrosis in 7.7% (4/52) patients and that use of >3 chemotherapeutic agents was associated with radiation necrosis.[149] The frequencies brain stem injury and necrosis were generally low in the largest of the case series. In the largest series (N=516)[83], Grade 3 or 4 brain stem injury occurred in 0.6% of patients. Cumulative incidence of grade 3 or higher brainstem toxicity in the next largest series (N = 313) was 2.1% (± 0.9%).[120] Across five case series of at least 100 persons, Grade 3+ brainstem necrosis or injury occurred in less than 2% and most usually in less than 1% of PBT recipients.[34,83,84,118,120] (Main Appendix F, Table F63 and F64)

Hearing loss

Grade 3 and Grade 4 hearing loss were not statistically different for PBT and photon therapy in a cohort of patients with medulloblastoma (N=84)[221] (Figure 7). Four case series reported on hearing loss. (Main Appendix F) The largest series (N=179)[118] reported late Grade 2+ hearing loss regarding hearing aids in 6.1% of patients with ependymoma. The next largest series (N=45)[320] reported Grade 3-4 hearing loss in 15.6% of patients with medulloblastoma with 3 and 5-year cumulative incidence of ototoxicity in 12% (4% to 25%) and 16% (6% to 29%) of patients respectively. Hearing loss in the other two small series were reported in 4% (2/50)[19] and 8.7% (2/23)[167] patients (Main Appendix F, Table F65).

Neurocognitive measures

Neurocognitive measures evaluated in various studies included the following: Full Scale Intelligence Quotient (FSIQ) from Wechsler Intelligence Quotient scales with Mental Development Index (MDI) as well as component scores for the FSIQ namely, the Perceptual Reasoning Index (PRI); Processing Speed Index (PSI); Verbal Comprehension Index (VCI); and Working Memory Index (WMI). Studies do not describe thresholds for changes that may be clinically significant.

A recent prospective cohort study (N=93)[132] in patients with various brain tumors evaluated neurocognitive measures across three treatment groups, CSI PBT, focal PBT and surgery up to 6 ears post-treatment. There was no direct comparison with contemporary XRT. Data were available for 74.5% of enrolled patients. After adjusting for baseline differences previously described, FSIQ and PSI score declines with time were significantly greater for the CSI PBT group compared with the surgery only

group. Beta coefficients, 95% CIs and p-values respectively were -2.1 (-3.8 to -0.3), p = 0.019 and -2.6 (-4.7 to -0.3), p=0.020). No statistically significant changes over time in VCI, PRI or WMI were reported for the comparisons of either CSI PBT or focal PBT compared with surgery (or when CSI and focal PBT were compared). Authors report that all index scores remained stable with time for the focal PBT and surgery groups while declines were seen in the CSI PBT groups. While authors suggest that this supports neurocognitive sparing with focal PBT, the impact of potential residual confounding is unclear. The clinical significance of the changes is not described.

In the same author's prior retrospective cohort study[131] (N=150) in a different patient population with various brain tumors, there were no statistically significant differences in IQ score changes with time between PBT and photon radiation therapy (beta coefficients -0.7 (95% CI -1.6 to 0.2), vs. -1.1 (95%CI -1.8 to -0.4), p = 0.509,) even though scores in the PBT group were persistently higher in the PBT group by 8.7 points[131] and those in the photon group experienced a decline of 1.1 IQ points per year versus no decline in the PBT group. Authors to not provide a threshold for changes that may be clinically meaningful. Similarly there were no statistically significant differences between treatment groups for subanalyses of patients who received craniospinal radiation (beta coefficients -0.8 vs. -0.9, p = 0.89) or those who received focal RT (beta coefficients 0.6 (95% CI -2.0 to 0.8) vs. -1.6 (95% CI -3.0 to -0.2); p = 0.34). IQ data were available for 150 out of 205 eligible patients (73%). Sample size and use of global IQ measures may have precluded detection of clinical or statistical differences between groups. (Abstraction Appendix N)

Changes from baseline in IQ scores from the FSIQ) from Wechsler Intelligence Quotient scales with Mental Development Index (MDI) from Bayley scales were reported in six case-series[92,167,218,228,291,320] in pediatric patients who had PBT to treat various brain tumors. Component scores of the FSIQ were also reported and included PRI, PSI, VCI and WMI in some series (see above for acronyms). While none of the studies describe what may constitute a clinically significant decrease in these scores, one study indicates that scores of less than 69.7 are considered to be at risk for impairment.[291] The usual categorization of average scores on the Wechsler Intelligence scale for children is as follows: low-average 80-89, average 90-109, high-average 110 to 119.[314] (Additional detail in Table 1)

Across case series, results suggest that PBT may impact IQ scores, processing speed and other neurocognitive measures; however the clinical significance of some of the changes is not clear and conclusions are limited in the absence of data from studies comparing PBT to other forms of radiation in particular and control for potential confounding factors. Reporting was based on usually small subsets of patients in whom neurocognitive tests were performed. All case series were considered to be at high risk of bias. A summary of these outcomes is found in Main Appendix F, Table F66.

All but one study[218] from Korea was from the same institution, Massachusetts General Hospital (MGH) and, based on patient enrollment dates, it is likely that there is overlap in study populations. The largest series by Pulsifer 2018 (N=114 with IQ data) appears to have the most complete data and includes various brain tumors. Three of the publications from MGH measured IQ in patients treated with PBT for specific tumor types, including ependymoma (N=14)[167], low-grade glioma (N=12)[92] and medulloblastoma (N=54)[320]; the other publication (N=65) also included various tumor types provided some information on IQ and cognitive function but focused on a mediation model. The Pulsifer series reported a decrease of approximately 2.9 IQ points between baseline assessment and followup (mean follow-up interval of 3.6 years) that was statistically significant; however the clinical significance is not clear. Subanalysis suggests that younger patients (<6 years old) had a significant decline and that regardless of age, patients receiving craniospinal radiation (CSI) seemed to be particularly vulnerable. Authors report the mean

baseline and follow-up scores for the whole population were considered to be average. Medulloblastoma was the most common tumor (35%) in this population. The study by Yock from the same institution, which focused on patients with medulloblastoma, reported a statistically significant annual decline in FSIQ of 1.5 points per year at a median follow-up of 5.2 years.[320] Again, the clinical importance of this decrease is not clear. Both studies note decrements in processing speed between baseline and follow-up as does the series by Ventura from the same institution. While the larger Pulsifer series[228] noted decreases in working memory and perceptual reasoning between baseline and follow-up, the Yock series in patients with medulloblastoma found no significant change. The finding of no substantial differences in scores from baseline to followup in the series of ependymoma[167] and low grade glioma[92] may in part be due to small sample sizes; conclusions regarding the impact of PBT are not possible.(Appendix F, Table F66)

The study from Korea reported on 20 patients with intracranial germ cell tumors of which 10 received CSI, 10 received whole ventricle irradiation.[218] At 1-2 years post-PBT, authors report that scores at follow-up for all neurocognitive domains were not significantly different from baseline overall. Patients who had CSI tended to have lower follow-up scores compared with baseline while those who had whole ventricle irradiation tended to have higher follow-up scores versus baseline, but differences were not statistically significant. Authors report that overall, scores were lower than expected when compared with a normal population. Given the small sample size, firm conclusions are not possible.

Case series – general toxicities by grade

Three case series (N=105) reported on acute toxicity.[180,307,320] Across studies and tumor types, 100% of patients experienced acute Grade 1 or 2 toxicity and the frequency of ≥ Grade 3 toxicities ranged from 0% to 83%. The frequency of ≥ Grade 3 late toxicities across four case series (N=340) [19,83,307,320] ranged from 1.9% to 13.6%. Across three case series[119,120,189]which did not specify the timing of toxicity, 1.3% to 7.1% of patients experienced treatment-related toxicity. (Main Appendix F, Tables F67 and F68 in)

Key Question 5 (Economic)

Two poor quality full economic studies met the inclusion criteria; both were cost-utility analyses (CUA). One poor quality (QHES 50/100 points) formal cost-effectiveness study in patients with medulloblastoma was funded by the Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST program) and the National Cancer Center Research and Development Fund.[109] The other poor quality (QHES 48/100 points) study included various CNS tumors[170] and evaluated various PBT/photon dose pairs to evaluate thresholds for cost-effectiveness (Mailhot Vega 2015). Primary concerns were limited parameters modeled, clinical data derived from case series of treatment options, lack of transparent methodology, and inadequate sensitivity analyses. Table 12 and Main Appendix E. Both studies modeled hypothetical pediatric populations using Markov models based on data from case series to evaluate the cost-effectiveness of PBT versus conventional radiotherapy and reported incremental cost-effectiveness ratios (ICER) to reflect the marginal cost of adding one quality adjusted life year (QALY) to a patient's life when using PBT versus conventional radiotherapy (photon therapy).

Key points

Pediatric Medulloblastoma:

One poor quality CUA was conducted in Japan using a payer perspective concluded that PBT is more cost-effective from a payer perspective than conventional x-ray therapy for pediatric patients with medulloblastoma at a WTP threshold of $46,729/QALY (JPY 5 million/QALY).[109]

- ICER ranged from $21,716/QALY using EQ5-D, $11,773/QALY using HU13, and $20,150/QALY using SF-6D.
- Sensitivity analyses: The discount rate, incidence of hearing loss in average risk patients and cost of PBT were most influential; all simulations yielded ICERS < $46,729/QALY) threshold. The probability of WTP being < $46,729/QALY in simulations was 99.5% overall.
- Limitations:
  - Inadequate description of PBT costs; incomplete delineation of operational costs
  - Clinical outcomes data are from case series for conventional photon therapy and IMRT ; PBT dose information was derived from eight patients who had PBT; no comparative studies on effectiveness or harms
  - Limited outcomes considered: no inclusion of long-term outcomes related to motor/physical or intellectual challenges or long-term health challenges or costs
  - Utilities based on hearing aid use in adult populations in western countries and may not be applicable to this study population

Pediatric CNS tumors:

One poor quality study (Mailhot Vega 2015) modeled various PBT radiation doses versus different expected equivalent photon radiation doses to evaluate thresholds for cost-effectiveness in pediatric patients with various CNS tumors.[170] Authors concluded that averting growth hormone deficiency (GHD)-related treatment costs may make PBT a cost-effective and possibly cost-saving strategy for many PBT doses versus photon therapy.

- ICERs ranged based on proton-photon dose combinations; many combinations, particularly at lower doses of PBT, were cost-effective or cost-saving at a WTP of $50K/QALY. PBT was not cost-effective at the highest PBT dose (30 Gray [Gy]) compared with photon therapy, however.
- Sensitivity analyses were limited; assuming no utility difference between GDH and health) resulted in fewer instances where PBT was cost-effective versus the base-case scenario, again particularly at higher PBT doses (25, 30 Gy); authors state that PBT may continue to be cost-effective until the cost difference was $725,000 or $580,000 greater than photon therapy in 4 and 12 year olds respectively
- Limitations:
  - Completeness of PBT operational costs is not clear; detail is not provided
  - Use of identical costs for pediatric and adult GH treatment may not accurately represent the true costs
  - Limited outcomes were included in model; no modeling of long-term toxic effects (e.g. auditory or cognitive effects) other than GHD; use of utilities for adult populations may not reflect those for this population
  - Clinical data from were case-series, some not specific to PBT; no long term, comparative data to validate assumption of no difference in treatments for tumor control or other factors or true impact of PBT doses

- Methodology, modeling parameters and assumptions not transparently described or documented; only ICERs reported (not components of ICERs such as cost/QALY of each treatment modality)
- Sensitivity analyses were limited and their basis/rationale not well described

Detailed Results

Table 12 summarizes characteristics and findings from pediatric cost-effectiveness studies.

Medulloblastoma

Study characteristics and framework

One poor quality CEA evaluated the cost-effectiveness of PBT compared to conventional x-ray therapies for treatment of pediatric medulloblastoma.[109] The costing year was 2012. The study adopted a payer perspective and modeled a hypothetical cohort of 6 year old patients over a lifetime horizon using a Markov model. Costs included treatment, hearing tests, hearing aid fitting tests, and hearing aids. The costs for PBT and conventional x-ray therapies were $26,943.90 and $3,082.20, respectively. PBT-specific clinical data were not used aside from dose information from eight patients who had undergone PBT. Data on outcomes were from case series of conventional radiation therapy.

Base Case Results

Treatment with PBT cost $28,937.00/QALY while conventional x-ray therapies cost $7,541.00/QALY. PBT and conventional x-ray therapies resulted in 23.44 and 22.46 QALYs using EQ-5D utilities, 22.78 and 20.96 QALYs using HU13 utilities, and 23.38 and 22.32 QALYs using SF-6D utilities, respectively. The ICER varied substantially based on the utility measure used. The ICERS was $21,716/QALY using EQ-5D, $11,773/QALY using HU13, and $20,150/QALY using SF-6D.

Both one-way sensitivity analyses and probabilistic sensitivity analyses (PSA) were conducted.

Sensitivity Analyses

Results from a one-way sensitivity analysis showed that the variables with the greatest impact on the ICER were the discount rate, incidence of hearing loss, and treatment costs for PBT. ICERs estimated from authors' figures ranged from about $4,000/QALY to about $50,000/QALY. PSA results demonstrated that the overall probability of PBT being cost-effective at a WTP threshold of $46,729/QALY (JPY 5 million/QALY) was 99.51% (96.95% for EQ-5D, 100% for HU13, and 98.72% for SF-6D).

Conclusions and Limitations

The authors concluded that PBT is more cost-effective than conventional x-ray therapy at a WTP threshold of $46,729/QALY. Additionally, given the sensitivity of the model to PBT costs, it is would be even more cost-effective if PBT costs decrease.

The primary limitation of this study is that several important costs and outcomes were excluded from the model. For example, there was no inclusion of long-term outcomes related to motor/physical or intellectual challenges or other long-term health challenges or costs. Indirect costs were not included and the description of included costs was inadequate. In addition, utilities were based on hearing aid use among adults in western countries and may not be applicable to this study population. Clinical parameters were derived from case series data that were not specific to PBT. Finally, these results may not be applicable to the United States' health care system. The QHES score for this study was 50/100 points.

Central nervous system tumors

Study characteristics and framework

One poor quality CUA modeled the cost-effectiveness of various PBT radiation doses versus different expected equivalent photon radiation doses in grays (Gy) delivered to the hypothalamic region. A Markov model consisting of two states, healthy and growth hormone deficiency (GHD) was used with a 60 year time horizon from a payer /healthcare system perspective. Hypothetical models for patients exposed to different proton-photon doses values in 4 year old and 12 year old patients were reported to evaluate dose-pair thresholds which may be cost effective. Radiation exposure to the hypothalamus is associated with endocrine deficiencies, and increasing radiation dose is associated with increased risk of GHD. The prevalence of GHD among pediatric cancer survivors is estimated to be 35.6%.[182] GHD is associated with increased cardiovascular risk factors[58] and cognitive impairment.[226] Dosimetric data for photon therapy and the basis of modeling GHD were taken from a study of 192 pediatric patients between 1997 and 2008[182] who had conformal radiation therapy for a variety of tumors; the most common tumors were ependymoma (46%) and low grade glioma (26%). On average, patients receiving more than 60 Gy of radiation developed GHD within 12 months, patients receiving 25 to 30 Gy developed GHD within 36 months, and patients receiving 15 to 20 Gy developed GHD within 60 months.[182] A cumulative dose of 16.1 GY was associated with a 50% risk of GHD at 5 years.[182]

Results

Cost effectiveness and cost-savings related to PBT varied by proton-photon dose comparisons. In both 4 year old and 12 year old patients, PBT was projected to be cost-effective at a WTP of $50,000/QALY or cost saving across a range of doses compared with the expected equivalent photon radiation doses. Cost-effectiveness at higher proton doses and for smaller differences between proton and photon doses tended to be lower, i.e., the ICER tended to increase. The highest modeled dose of PBT (30 Gy) was not cost-effective compared with photon therapy at any expected equivalent dose (35 to 60 Gy). At a PBT dose of 30 Gy, for those exposed at 12 years old, ICERs ranged from $512,400/QALY to $103,300/QALY, while for those exposed at 4 years old, ICERs ranged from $430,200/QALY to $96,200/QALY (All estimates are from author's figures.) Authors report that at a difference of 10 Gy between PBT and photon therapy, PBT may be cost effective; this appears to be dependent on dose. Authors conclude that proton therapy may be more cost effective for scenarios in which the radiation dose to the hypothalamus can be spared, but may not be cost effective with regard to GHD when PBT plans also deliver a high dose to the hypothalamus.

Sensitivity Analyses

Sensitivity analyses were limited and rationale for parameters evaluated was not well-documented. Specific drivers of cost other than setting utilities to a value of 1 and examining the differential cost between PBT and photon therapy are not evaluated. Sensitivity analysis setting the GHD utility to 1 (no utility difference between GDH and health) resulted in fewer instances where PBT was cost-effective versus the base-case scenario, again particularly at higher PBT doses (25, 30 Gy). The sensitivity analyses based on differences in PBT versus photon costs describes thresholds below which PBT would be cost-saving with an assumption that PBT cost is $40,600 less expensive than photon therapy. Authors evaluated the influence of PBT costs, concluding that PBT may continue to be cost-effective until the cost difference was $725,000 or $580,000 greater than photon therapy in 4 and 12 year olds respectively.

Conclusions and Limitations

Authors' models of different estimated PBT versus photon dose combinations suggest thresholds of dose and differences between PBT and photon costs at which PBT may be cost-effective at a WTP of $50,000/QALY and possibly cost saving in hypothetical cohorts of children with CNS tumors. PBT may be cost-effective at lower dose combinations and over a broad range of costs, with fewer instances of cost effectiveness at higher PBT-photon dose combinations in all scenarios. A number of limitations need to be considered. First, the only clinical data available are from case-series, some not specific to PBT. Second, models included utilities and costs for GHD based on those for adults, which may not be appropriate for this population as modeled. Details of cost basis for all parameters were not provided. While GHD may represent an important adverse effect of radiation therapy to the hypothalamus, other comorbidities, including cognitive impairment, hearing loss or hypothyroidism that would potentially affect both cost and utility were not included. The QHES score for this study was 48/100 points.

Table 12. Summary of economic studies comparing PBT with conventional RT in pediatric patients with brain or CNS tumors

	Hirano 2014	Mailhot Vega 2015
Population (condition)	6 year old patients with medulloblastoma	Pediatric patients with CNS tumors; cohorts exposed at age 4 or age 12
Intervention(s)	PBT (following chemotherapy)	PBT (timing, use as sole therapy unclear)
Comparator(s)	Conventional radiation therapy	Conventional radiation therapy
Country	Japan	USA
Funding	The Funding Program for World-Leading Innovative R&D on Science; Technology (FIRST program); National Cancer Center Research and Development Fund	Medical student grant from Conquer Cancer Foundation
Study design	CUA	CUA
Perspective	Payer	Health care system
Time horizon	Lifetime	60 years
Analytic model	Decision analysis (Markov model stated; no specifics provided)	Markov model with 2 health states

	Hirano 2014	Mailhot Vega 2015
Effectiveness outcome	QALYs	QALYs
Effectiveness outcome components	Hearing loss and death; QOL	QOL
Source for effectiveness data	Case-series data on PBT effectiveness Prior research using HRQOL measures before and after hearing aid use in adults	Prior literature (case series); not specific to PBT use in children with CNS tumors
Costing year	2012	2012
Currency	USD and JPY (1 USD = 107 JPY)	USD
Discounting	3%	3%
Components of cost data	Radiation cost (X-ray or proton), hearing test, hearing aid fitting test, hearing aid	Cost of GHD (medication and management costs), cost of photon or proton therapy
Cost sources	Table of Medical Service Fees in Japan (2012) PBT cost derived from median treatment fee of medical institutions in Japan	*Red Book* for GH, CMS for management, "institutional experience" and prior literature for PBT vs photon costs; little detail provided
Sensitivity analysis	One-way PSA: Monte Carlo simulations using 10,000 iterations	Conducted based on varying proton vs. photon doses to the hypothalamus
QHES	50	48
Results:		
Cost / QALY of intervention	EQ-5D: $28,937/23.44 = $1,235/QALY HU13: $28,937/22.78 = $1,299/QALY SF-6D: $28,937/23.38 = $1,238/QALY	NR
Cost / QALY of comparator(s)	EQ-5D: $7,541/23.46 = $321/QALY HU13: $7,541/20.96 = $360/QALY SF-6D: $7,541/22.34 = $338/QALY	NR
ICER	EQ-5D: $21,716/QALY HU13: $11,773/QALY SF-6D: $20,150/QALY	Dose-dependent; Range for 12 year-old patients: dominant (photon dose ≥ 15 Gy with proton dose 0 Gy, photon dose ≥25 Gy with proton dose 5 Gy, photon dose ≥30 GY with proton dose 10 Gy, photon dose ≥35 Gy with proton dose 15 Gy, photon dose ≥50 with proton dose 25 Gy) to $512,400/QALY (photon dose 35 Gy, proton dose 30 Gy) Range for 4 year-old patients: dominant (photon dose ≥ 10 Gy with proton dose 0 Gy, photon dose ≥20 Gy with proton dose 5 Gy, photon dose ≥25 GY with proton dose 10 Gy, photon dose ≥30 Gy with proton dose 15 Gy, photon dose ≥35 Gy with proton dose 20 Gy, photon dose ≥45 Gy with proton dose 25 Gy) to $430,200/QALY (photon dose 35 Gy, proton dose 30 Gy)

	Hirano 2014	Mailhot Vega 2015
One-way SA	Most influential parameters: discount rate, hearing loss incidence, and treatment costs for proton irradiation of average-risk group (range of ICERs not reported)	Not done
Other SA	Results from Monte Carlo simulations: most trials yielded values <$46,729/QALY, the willingness to pay threshold Probability of willingness-to-pay <$46,729/QALY is 99.51%	PBT is cost-effective until costs $580k more than photon therapy for 4 year-olds or costs more than $725k than photon therapy for 12 year-olds; GHD utility = 1: fewer instances where PBT was cost-effective particularly at higher PBT doses (25, 30 Gy).
Author's Conclusion	At a cost-effectiveness threshold of $46,729/QALY (JPY 5 million/QALY), PBT is more cost-effective than conventional X-ray therapy	Proton therapy may be more cost effective for scenarios in which radiation dose to the hypothalamus can be spared, but PBT may not be cost effective when tumors involve or are directly adjacent to the hypothalamus and radiation dose is high
Limitations	• Inadequate description of PBT costs; incomplete delineation of operational costs • Clinical outcomes data are not from comparative studies • Limited outcomes considered: no inclusion of long-term outcomes related to motor/physical or intellectual challenges or long-term health challenges or costs • Indirect costs not included • Utilities based on hearing aid use, not specific to post-radiation population of children • Utilities derived from western countries and adult populations; may not be applicable to this study population; ICER varies by utility used • Doses of radiation derived from small cohort (8 patients) • May not be applicable to US	• Inadequate description of PBT costs; • Limited parameters included in model; no modeling of long-term toxic effects (e.g. auditory or cognitive effects) other than GHD • Data from case-series; no long term comparative data to validate assumption of no difference in treatments for tumor control or other factors or true impact of PBT doses or lifetime horizon • Basis of PBT operational costs not detailed • Sensitivity analyses were limited • Did not model death • Methodology, modeling parameters and assumptions not transparently described or documented; only ICERs reported (not components of ICERs) • Detailed costing basis was not provided • Utility weight derived from adult study; assumes costs of therapy for adults and children are similar

PBT = proton beam therapy; CNS = central nervous system; CUA = cost-utility analysis; QALY = quality-adjusted life-year; JPY = Japanese yen; GHD = growth hormone deficiency; HRQOL = health-related quality of life; CMS = Centers for Medicare and Medicaid Services; PSA = probabilistic sensitivity analysis; QHES = Quality of Health Economic Studies; SA = sensitivity analysis; ICER = incremental cost-effectiveness ratio; Gy = Gray (unit of absorbed dose); WTP = willingness-to-pay

Key Question 2 (Effectiveness, salvage therapy) and Key Question 4 (Differential Effectiveness and Safety)

No studies that met inclusion criteria were identified.

4.2.2 Head and Neck (including Skull-base)

Key points

- There is insufficient evidence from three case series to evaluate the effectiveness and PBT in in pediatric patients with head and neck tumors. (Insufficient SOE)
- Evidence from one small cohort study which evaluated the safety of PBT in pediatric patients with primary salivary gland tumors suggests that acute Grade 2 or 3 mucositis may be significantly less common with PBT versus other forms of RT; the frequency of other acute toxicities was similar between groups. The evidence was considered to be insufficient due to study limitations and small sample size. (Insufficient SOE)
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness of PBT in this population.
- The 2014 report made conclusions across all pediatric tumor categories. Evidence from three additional case series of head or neck cancers included in this update report is insufficient to draw conclusions about the effectiveness PBT. Evidence from one small study compared PBT with other forms of RT was considered to be insufficient to draw conclusions about toxicities.

Key Question 1 (Effectiveness, curative intent)

Description of included studies

One small retrospective study (N = 24) in children with salivary gland tumors compared adjuvant PBT with adjuvant photon RT[89] and reported on acute toxicities but not on primary outcomes for KQ 1; it is included in KQ 3.

Three small case series in patients receiving PBT for different tumors types were identified.[165,231,300] (Abstraction Appendix N)

Results

OS and PFS in the series of PBT in patients with skull-based chordoma (N=18)[231] at 5 years were 54% and 57% respectively; both OS and PFS were 57% at years 10 and 20. One very small series in patients with esthesioneuroblastoma (n=8)[165] reported 87.5% OS at 5 years. The third series (N=69)[300] enrolled patients with various tumor types (rhabdomyosarcoma, Ewing sarcoma, others) and reported 93% 1 year OS across all patients. (Main Appendix F, Table F69, F70, and F71)

Key Question 3 (Safety)

Description of included studies

One small comparative retrospective cohort study (N = 24) in children with salivary gland tumors (which are rare) compared adjuvant PBT (n = 11) with adjuvant photon RT (n=13)[89] and reported on acute

toxicities as did two case series[165,300]. (Abstraction Appendix N). In the cohort study PBT recipients were slightly younger versus RT recipient and had substantially shorter duration of follow-up. Between the treatment groups, there were a similar proportion of males, similar proportion receiving chemotherapy and reasonably similar distribution regarding tumor. Most tumors were in the area of the parotid and most common tumor type was mucoepidermoid. Both groups received similar mean total radiation. Table 13.

Table 13. Study and patient characteristics from comparative studies reporting on effectiveness or safety: Pediatric head and neck tumors

	Safety Only	
	Grant 2015*	
Characteristics	PBT (n=13)	Photon† or electron RT (n=11)
Patient demographics		
Males, % (n)	46%	45%
Age, years; median (range)	13 (6 to 18)	15 (7 to 18)
Tumor characteristics		
Subtype		
Mucoepidermoid carcinoma	54%	45%
Adenoid cystic carcinoma	23%	18%
Adenocarcinoma	15%	0%
Acinic cell carcinoma	0%	18%
Pleomorphic adenoma	8%	0%
Myoepithelioma	0%	9%
Undifferentiated carcinoma	0%	9%
Tumor Grade		
Low/Intermediate	54%	45%
High	15%	27%
Unknown	31%	27%
Radiation Treatment		
Technique	Passive Scatter or IMPT	Electron Beam Therapy or IMRT
Median total dose (Gy)	60	60
Additional Treatments		
Prior to RT		
Submandibular glad resection	16.7%	
Superficial parotidectomy	29.2%	
Neck dissection	66.7%	
Concurrent/Adjuvant		
Concurrent Chemotherapy	7.9%	9.1%
Study Design	**Retrospective Comparative Cohort**	
Follow-up, months (% followed)	**8† (100%)**	**92† (100%)**
Risk of bias	**Moderately High**	

Gy = Gray; IMPT = Intensity Modulated Proton Therapy; IMRT = Intensity Modulated Radiation Therapy; PBT = Proton beam therapy; RT = radiation therapy

*Adjuvant PBT , 8 had passive scatter PBT, 5 had modulated PBT; Adjuvant RT, 8 had electron beam RT, 3 had IMRT.

†Indicates a statistically significant difference between groups

Results

In the cohort study[89], fewer patients in the PBT group experienced the following: Frequency of Grade 2 and 3 dermatitis (7/13 vs. 6/11), dysphagia (0 vs. 3/11) and otitis externa (1/13 vs. 2/11) was similar between groups. Grade 2/3 acute mucositis was significantly less common following PBT versus photon RT (IMRT or electron) (6/13 vs. 10/11, RR 0.51 (0.27, 0.94). Authors also report that total body integral radiation dose was substantially lower with PBT versus photon RT.

In the largest case series of patients (N=69)[300] with various tumor types, most acute toxicities were Grade 1 with frequencies ranging from 1% (dehydration, taste change) to 61% (radiation dermatitis). Grade 3 toxicities occurring in >1% of patients included anorexia (22%), dysphagia (7%), oral mucositis (4%) and dry mouth (3%); Grade 3 dehydration, nausea, mucosal infection and radiation dermatitis occurred in 1% of patients. In the small (n=8) series of patients with esthesioneuroblastoma, there were 5, 18 and 5 acute toxic events reported as Grade 1, 2, and 3 respectively. Grade 2 radiation dermatitis was the most common and was experienced by 5/8 patients.[165] (Main Appendix F, Table F72)

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential effectiveness and safety), Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.2.3 Lymphoma

Key points

- There is insufficient evidence from 2 case series to evaluate the effectiveness or safety of PBT in in pediatric patients with lymphoma.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness of PBT in this population.

Key Question 1 (Effectiveness, curative intent)

Description of included studies

Two case series[115,316] reported on patients with pediatric Hodgkin Lymphoma. It is unclear whether there may be some overlap in patient populations in these studies given overlap in authors and institutions. PBT was used for consolidation treatment following chemotherapy in the larger series which included 59 pediatric patients.

Results

At 3 years, relapse-free survival was 87% overall and was 100%, 83% and 87% for favorable early stage, unfavorable early stage and advanced stage disease respectively in the larger series (N= 59).[115] The small case series (N=22)[316] reported 2 and 3-year OS as 94% and PFS for both years as 86% with 3/22 (13.6%)

of patients experiencing recurrence. All patients had also received chemotherapy. Recurrence was experienced by 10.2% and 13. 6% of patients in the large and small series respectively.

Key Question 3 (Safety)

Results

Two small case series provided limited information on safety or harms. No patients experienced grade 3 toxicities and no clinically meaningful pneumonitis was reported in the large series.[115] The small case series (N=22) also reported that no PBT-related Grade 3 or 4 (early or late) toxicities occurred.[316]

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety), Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.2.4 Ocular Tumors

Key points

- There is insufficient evidence from 2 small case series to evaluate the effectiveness PBT in pediatric patients with ocular tumors. (Insufficient SOE)
- In one small retrospective cohort study in pediatric patients where PBT was used primarily for salvage therapy in pediatric patients with retinoblastoma, acute toxicities were more common with PBT, however, statistical significance was not reached likely due to sample size and/or confounding. Evidence was considered insufficient due to study limitations and sample size. (Insufficient SOE)
- No studies meeting inclusion criteria were identified that evaluated differential effectiveness and safety or cost-effectiveness of PBT in this population.
- The 2014 report reported conclusions across all pediatric tumor categories. One poor quality study in patients with ocular tumors comparing PBT with photon therapy was included and result suggested that cumulative incidence of secondary tumors may be lower following PBT. In this update, evidence from two additional case series of head or neck cancers included in this update report is insufficient to draw conclusions about the effectiveness PBT for treatment of ocular tumors. Similarly, evidence from one small study compared PBT for salvage treatment with other forms of RT was considered to be insufficient to draw conclusions about toxicities.

Key Question 1 (Effectiveness, curative intent)

Description of included studies

No comparative studies were identified. Two case small series of PBT for treatment of pediatric ocular tumors were identified, one in patients with uveal melanoma (N=43)[223], the other in patients with retinoblastoma[195].

Results

In patients with uveal melanoma, 5, 10 and 15 year relative survival rates were 93%, 93% and 85% respectively with corresponding metastasis rates of 8%, 11% and 19%.[223] In patients with retinoblastoma, authors report that no patients died or developed metastases and visual acuity was considered good in 47% of 30 eyes and moderate in 23%.[195]

Key Question 2 (Effectiveness, salvage therapy)

Description of included studies

Evidence for the use of PBT as salvage therapy in pediatric patients with tumors is sparse. Only one comparative study reported on the use of PBT for salvage therapy in pediatric patients and focused on ocular tumors.[15]

The small retrospective cohort study in pediatric patients with retinoblastoma (N = 39, 47 eyes) treated with radiation therapy[15] for salvage intent, compared PBT (16 eyes) with photon or electron radiation therapy (ERT, n=27 eyes) or brachytherapy (BRT, n = 4 eyes). Mean ages were < 2 years old in all groups but ranged from 3 months to >10 years, 49% were male. More PBT patients (93.8%) than RT patients (51.5%) were treated as salvage procedures; PBT patients had more advanced disease and more intensive treatment overall. Mean radiation dose was lower for PBT compared with photon therapy, (median total doses 36 Gy versus 45 Gy). Length of follow-up for PBT was substantially shorter (3 years) versus ERT (10 years). Authors did not evaluate or adjust for potential confounding. Table 14 below provides a summary of patient and study characteristics.

Table 14. Study and patient characteristics from comparative studies reporting on effectiveness or safety: Pediatric ocular tumors

	Effectiveness and Safety		
	Agarwal 2016		
Characteristics	**PBT (n=16 eyes)**	**Photon RT (n=27 eyes)**	**Brachytherapy (n=4 eyes)**
Radiation Intent			
Postoperative	8.5%		
Curative	29.8%		
Salvage	61.7%		
Patient demographics			
Males, % (n)	49%		
Age, years; median (range)	**1.9 (0.9 to 4.3)**	**1.4 (0.25 to 10.4)**	1.8 (0.83 to 4.9)
Tumor characteristics			
Lateral Disease			
Trilateral	3%		
Bilateral	77%		
Unilateral	21%		
ICSS			
B	17%		
C	8.5%		

	Effectiveness and Safety		
	Agarwal 2016		
Characteristics	**PBT (n=16 eyes)**	**Photon RT (n=27 eyes)**	**Brachytherapy (n=4 eyes)**
D	42.6%		
E	14.9%		
Extraocular	10.6%		
Unknown	6.4%		
Radiation Treatment			
Technique	Passive scatter	---	---
Median total dose (Gy)	36.0	45.0	45.0
Additional Treatments			
Chemotherapy	72%		
Study Design	**Retrospective Comparative Cohort**		
Follow-up, months (% followed)	36 (all pts, 97.4%)	120 (all pts, 97.4%)	60 (all patients, 97.4%)
Risk of bias	**Moderately High**		

Gy = Gray; ICSS = International Classification System Stage; IMRT = intensity modulated radiation therapy; NR = Not Reported; PBT = proton beam therapy; RT = radiation therapy; SBRT = stereotactic body radiation

Results

Authors provide limited data on effectiveness of PBT for salvage therapy in very young patients (< 2 years old) with retinoblastoma. They[15] report OS of 97% across treatment groups and enucleation-free survival of 38.5% for PBT versus 54.5% with other forms of radiation therapy in patients with stage D or E disease. Enucleation, considered under safety for this report, was done for a number of reasons (local disease progression, intraocular hemorrhage, painful glaucoma or factors leading to inability to examine the eye), some of which may not be related to radiation safety.

Key Question 3 (Safety)

Description of included studies

The retrospective cohort study described in KQ2 reported limited comparative information by treatment groups (PBT, RT, brachytherapy) for toxicities in a population where salvage was the primary intent.[15] Patient and study characteristics are summarized in Key Question 2, Table 14. Detailed results are found in Abstraction Appendix N.

Two case small series of PBT for treatment of pediatric ocular tumors were identified, one in patients with uveal melanoma (N=43)[223], the other in patients with retinoblastoma (N= 49, 60 eyes)[195] reported on toxicities, harms or safety. Detailed results are found in Abstraction Appendix N.

Results

In the retrospective cohort[15], study acute toxicities were more common with PBT (93.8%) versus other RT (74.1%), however, statistical significance was not reached likely due to sample size and/or confounding. Skin erythema was most common acute toxicity (data not provided separately by

treatment). Salvage was the primary use for PBT (93.8%) versus RT (51.5%) and PBT recipients overall had more intensive treatment. Enucleation was performed in 37.5% (6/16 eyes) in the PBT group compared with 29.6% (8/27 eyes) of those receiving other forms of RT; reasons for enucleation were not provided by treatment group and not all may be related to radiation exposure. Median time to enucleation from the end of salvage treatment was 11.5 months; details by treatment group are not provided. Authors do not report other late/long-term toxicities separately by treatment. For PBT recipients versus ERT or BT, the following late/long-term complications based on the number of eyes were reported: Cataracts (5 vs. 10), vitreous hemorrhage (3 eyes vs. 4), radiation retinopathy (2 eyes vs. 3), change in visual acuity (0 vs. 4 eyes) and strabismus (1 eye vs. 2). Follow-up in the PBT group was substantially shorter (3 years versus 10 years) than for other RT types, precluding conclusions regarding comparability of long-term toxicities including secondary malignancies.

Two case small series of PBT for treatment of pediatric ocular tumors were identified, one in patients with uveal melanoma[223], the other in patients with retinoblastoma[195] reported on toxicities, harms or safety. (Abstraction Appendix N) Enucleation was performed in 12% of melanoma patients and 18% of eyes in patients with retinoblastoma. In those with melanoma, events included lens opacity (39%), need for pseudophakia (16%) and retinal detachment (21%)[223]. In patients with retinoblastoma, visual acuity was rated poor in 30% of patients and complications requiring correction occurred in 22% of patients[195].

Key Question 4 (Differential Effectiveness and Safety), Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.2.5 Soft Tissue Sarcomas

Key points

- There is insufficient evidence from six small case series to evaluate the effectiveness or safety PBT in in pediatric patients with soft-tissue tumors. (Insufficient SOE)
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness of PBT in this population.
- The 2014 report made conclusions across all pediatric tumor categories and consisted primarily of case series. Evidence from six additional case series in pediatric patients with rhabdomyosarcoma included in this update report is insufficient to draw conclusions about the effectiveness or safety PBT for these tumors.

Key Question 1 (Effectiveness, curative intent)

Description of included studies

No comparative studies were identified. Six small case series evaluated PBT for the treatment of pediatric soft tissue tumors (rhabdomyosarcoma)[152,153,157,186,298,306] across various time frames.

Results

Across 3 series (N = 179) [152,157,306], 5-year OS ranged from 73% to 80.6%. Probability of PFS in one series was 81.6% at 1 year and 72.4% at two years[186]; a separate series reported PFS of 72% at 5 years[306].

Across five series [153,157,186,298,306], the proportion of patients experiencing recurrence or progression ranged from 16.7% to 25.6%. (Main Appendix F, Table F73)

Limited information on mortality and disease progression/recurrence is reported across five case series. [153,157,186,298,306]Disease-related mortality ranged from 9.1% to 23.1% and proportion of patients experiencing disease progression or recurrence ranged from 16.7% to 25.6% (Main Appendix F, Table F74)

One series reported that health-related quality of life was improved at 4 years versus baseline based on proxy assessments by parents, however data were available for only 34/83 patients.[157]

Key Question 3 (Safety)

Description of included studies

No comparative data are available and there is limited information from case series (Main Appendix F, Table F75)

Results

The frequency of acute ≤ grade 3 (i.e. grade 1, 2 or 3) radiation-induced toxicities in one series (N =55) was 16.4%, however authors report that no late Grade 3+ radiation-induced toxicities occurred.[186] The most common acute Grade 3 toxicity in the largest series (N=83) was mucositis (12%); Grade 3 skin toxicity occurred in 3.6%. of patients.[157] In another series, across tumor types and radiation sites, acute Grade 3 radiation dermatitis occurred in 9% of patients and in those receiving PBT and among those receiving PBT to the head and neck, odynophagia was the most common acute Grade 3 toxicity (9.7%, 3/31 patients).[153]

Regarding late toxicities, the two largest series reported development of cataracts in approximately 14% of patients.[298,306] The most common adverse event in one small series (N=39)[306] was facial hypoplasia (20%); the need for growth hormone replacement, and chronic head and neck structure congestion each had a frequency of 13% in this series; this series also reported 5-eyear toxicity ≥ grade 3-free survival of 95%. Risk of secondary malignancy across two case series (N = 39 and 83) were 0% and 2.4% up to 44 months following radiation. [157,306] The limited length of follow-up and small sample sizes may have precluded detection of additional secondary malignancies.

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety), Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.2.6 Other Tumors (Bone, Mixed Tumors)

Key points

- There is insufficient evidence from 1 small case series to evaluate the effectiveness or safety of PBT in in pediatric patients with Ewing Sarcoma or from another large series of patients with various tumor types. (Insufficient SOE)

- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness of PBT in this population.
- The 2014 report made conclusions across all pediatric tumor categories and consisted primarily of case series. Evidence from two additional case series included in this update report is insufficient to draw conclusions about the effectiveness or safety PBT.

Key Question 1 (Effectiveness, curative intent)

Description of included studies

Single case series evaluating PBT were available for the following pediatric tumor types. Details are found in Abstraction Appendix N.

Results

Bone (Ewing sarcoma in spine): One small case series (N = 28) reported 5-year OS of 83% (95% CI 69.1%, 96.9%) and 5-year metastasis-free survival of 76.4 % (95%CI 60.1%, 92.7%).[310]

Various tumors: One case series enrolled pediatric patients with various tumor types from various anatomic locations (N=343)[184]; approximately 25% had PBT for recurrent disease. The most common radiation sites were central nervous system (37%) and head or neck (30%). Across all patients, OS decreased between 1 year (82.7%) and 10 years (58.7%). OS was lowest in patients with neuroblastoma (72% at 1 year, 57.6% at 5 years) and highest in those with brain tumors (91.4% at 1 year and 81.7% at 5 years.) (Main Appendix F, Table F76)

Key Question 3 (Safety)

Description of Included Studies

Limited data from single case series provided data on PBT toxicity in various pediatric tumor types. (Abstraction Appendix N)

Results

Bone (Ewing sarcoma): One small case series (N = 38)[310] reported that 52.6% of patients had at least one late toxicity; all but two events were classified as Grade 1 or 2 however.

Various tumors: One case series reported findings across two publications (Mizumoto 2016/2017). In the initial report (N=343)[184], across all patients, the frequency of toxic effects was less than 2% with few grade 3 or 4 events reported. Overall, 2% of patients developed a second malignancy. In the subgroup of patients followed ≥5 years (median of 8 years, N=62)[185] late toxicity event of ≥ Grade 3 at 5, 10 and 20 years were 6%, 17% and 17% respectively.

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety), Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3 Adult Tumors

A total of 38 comparative studies were included in the 2014 report: 2 RCTs (1 in ocular tumors and 1 in prostate), 28 cohort studies [6 prospective (4 of which were PBT vs. carbon ion) and 22 retrospective], and eight non-contemporaneous case series. Additionally, 209 single arm case series were included. The tumor types with the most comparative evidence in the previous report were prostate cancer (10 studies) and ocular tumors (8 studies). The studies in the old report included a variety of comparators, many of which were not represented in the studies included in this 2019 review (

Table 15). For example, carbon ion therapy was excluded as a comparator because it is currently not available in the United States and comparisons of PBT to PBT plus another treatment were also not included. Half of the studies evaluating ocular tumors that were included in the previous report compared PBT to enucleation; no such studies were identified in the update report, likely due to advancements in eye and vision sparing radiation techniques (such as photon and proton).

This 2019 re-review found 36 comparative studies that met inclusion criteria and provide evidence on efficacy or effectiveness and/or safety: two RCTs (Liver and Lung cancer),[42,161] one quasi-RCT (Prostate cancer)[139] and 33 retrospective comparative cohorts[13,33,35,37,48,69,76,77,94,108,111,129,163,164,168,173,178,196,206,216,233,238,244,251,255,258-260,274,282,303,317,326] Additionally, 108 case series (across 109 publications) were included. Table 16 below provides the number and type of study for each specific tumor category. The tumor types with the most comparative evidence in this re-review were head and neck (8 studies), lung (6 studies) and esophageal (5 studies). Most studies compared PBT to more updated forms of radiation therapy such as IMRT (primarily) or 3DCRT (

Table 15). Additionally, four cost-effectiveness studies (one each in breast, head and neck, liver, and ocular tumors)[159,171,192,253] and four studies included for context only (all in prostate cancer)[98,198,227,286] were identified that met inclusion criteria.

As stated previously (see section 4.1), with the exception of the two RCTs, all comparative studies were considered to be of moderately high risk of bias due to the retrospective study design and concerns regarding blinding, loss to follow-up and controlling for confounding. Risk of bias assessment for included comparative studies is found in Main Appendix E. All case series are considered at high risk of bias. The quality of the included cost-effectiveness evaluations varied widely; common methodological concerns across the poorer quality studies were the use of non-validated health outcome measure/scales and lack of transparency regarding the structure of the economic model. Quality assessment for cost-effectiveness studies is found in Main Appendix E.

Results are organized by general tumor category. Key Questions 1-5 are addressed sequentially within each category.

Table 15. Overview of comparators (by tumor type) for adult populations evaluated in the 2014 report and in this 2019 re-review

	2014 report		2019 re-review	
	# studies	**Comparator* (vs. PBT)**	**# studies**	**Comparator* (vs. PBT)**
Bone tumors	1	**PBT + photon + surgery**	0	-----
Brain, Spinal, Paraspinal	2	IMRT (1) Photon RT (1)	5	IMRT (2) Photon RT (3) FSRT (1)
Breast	0	----	2	Photon +/- electron boost (1) WBI (1)
Esophageal	2	IMRT (2) 3DCRT (2)	5	IMRT (4) 3DCRT (1) XRT (1)
GI (pancreas)	0	----	1	Photon
Head and neck	2	IMRT (1) Endoscopy (1) **PBT + photon (1)**	8	IMRT (7) Surgery alone (1)†
Liver	3	**Carbon ion (1)** Photon (1) **PBT + chemotherapy (1)**	2	TACE (1 RCT) IMRT (1)
Lung	4 (1 NCCS)	**Carbon ion (1)** IMRT (3) (1 NCCS) 3DCRT (3) (1 NCCS)	7‡	IMRT (5) (1 RCT) 3DCRT (1) Various photon (1)
Ocular	8 (2 NCCS)	Enucleation (4) **PBT + TTT (2) (1 RCT, 1 NCCS)** **PBT + endoresection (1 NCCS)** **PBT + chemotherapy (1)** **PBT + laser (1 NCCS)**	3	Brachytherapy + TSR (1) Brachytherapy alone (1) Stereotactic radiosurgery (1)
Prostate	10 (4 NCCS)	IMRT (4) (2 NCCS) 3DCRT (1 NCCS) Brachytherapy (2 NCCS) Conservative (1) EBRT (2) **PBT + EBRT (1)** Watchful waiting (1) Photon (1 RCT)	4	IMRT (3) Photon alone (1 quasi-RCT)
Noncancerous conditions	3	Photon (3) **PBT + Photon (1)**	0	-----
Mixed cancer types	3 (1 NCCS)	Photon (1 NCCS) **Carbon ion (2)**	0	-----

FSRT: Fractionated stereotactic radiation therapy; NCCS – non-contemporaneous case series; TACE: transcatheter arterial chemoembolization; TSR = trans-scleral resection; WBI = whole breast irradiation with X-rays

*Some studies had more than one comparator arm. Parentheses indicate how many studies reported on that comparator.

†One study of skull-base chondrosarcoma (all other head and neck tumors were non-skull-base, e.g., oropharyngeal, nasopharyngeal, sinonasal).

‡Includes the nonrandomized cohort from the RCT.

Table 16. Summary of included studies in adult patients by tumor type: 2019 re-review

Tumor	Comparative*			Case series*		
	Total	Curative	Salvage	Total	Curative	Salvage
Bladder	0	0	0	1[271]	1[271]	0
Bone	0	0	0	8 16,49,63,121,130,241,264,267	8 16,49,63,121,130,241,264,267	0
Brain	5 13,37,94,129,196	3[13,37,129]	2 94,196	6 24,70,137,175,188,197‡	5 24,70,175,188,197‡	1[137]
Breast	2[48,274] 1 Economic[171]	2[48,274]	0	4 41,57,214,292	4 41,57,214,292	0
Esophageal	5 77,164,173,255,317	5 77,164,173,255,317	0	2[122,269]	2[122,269]	0
Gastrointenstinal (Pancreas)	1[168]	1[168]	0	2[112,142]	2[112,142]	0
Head and neck	8[33,111,178,238,251,259,260,326] 1 Economic[253]	8 33,111,178,238,251,259,260,326	0	23 59,64,67,78,81,91,93,103,104,177,179,193,199,224,239,267,272,281,301,308,309,324,325	18 59,64,67,78,81,91,93,177,193,199,267,272,281,301,308,309,324,325	5 103,104,179,224,239
Liver	2 (1 RCT)[42,244] 1 Economic[159]	2 (1 RCT)[42,244]	0	12 79,80,90,113,114,140,143,183,187,213,318,322	8 79,80,90,114,183,187,213,318	4 113,140,143,322
Lung	6 (1 RCT) 108,161,206,233,282,303	5 (1 RCT) 108,161,206,233,282	1[303]	12 44,46,101,123,136,156,174,190,205,212,242	11 44,45,101,123,136,156,174,190,205,212,242	1[46]
Lymphoma	0	0	0	3[115,117,200]	3[115,117,200]	0
Ocular	3[35,163,258] 1 Economic[192]	3[35,163,258]	0	22[28,134,147,154,176,217,219,225,230,235,236,243,246,248,249,275-278,305,312,313]	21[28,134,147,154176,217,219,225,230,236,243,246,248,249,275-278,305,312,313]	1[235]
Prostate	4 (1 quasi-RCT) 69,76,139,216	4 (1 quasi-RCT) 69,76,139,216	0	11 (12 publications)[20,39,50,53,110,116,124,172,181,227,270,285§**	11 (12 publications)[20,39,50,53,110,116,124,172,181,227,270,285**	0
Hemangiomas (benign)	0	0	0	2[169,323]	2[169,323]	0
Other benign tumors†	0	0	0	4[73,197,299,304]‡	3[73,197,299]‡	1[304]
Various/mixed	0	0	0	3[194,207,327]	3[194,207,327]	0
TOTAL	**41§ (4 economic)††**	**38§ (4 economic)**	**3**	**114 (115 publications)‡****	**101(102 publications)‡****	**13**

* Unless otherwise indicated, all comparative studies were retrospective cohort/observational studies.
†Includes meningioma and pituitary adenoma.

‡One of the case-series is included in the count for both Brain Tumors and Other Benign Tumors. This series included both malignant (WHO grade 2/3) and benign (WHO grade 1) meningiomas[197]; data for malignant tumors is described in the section on Brain Tumors and data for benign tumors is described in the section on Other Benign Tumors.
§Includes the nonrandomized cohort from the RCT (Liao 2018), which is described in the same publication.
**Bryant/Colaco 2015 (prostate) are one study published across two articles.
††Additionally, four comparative studies (2 RCTs and 2 retrospective cohort studies), all in men with prostate cancer, were included for contextual purposes only and are not included in the count here.

4.3.1 Bladder Cancer

Key Points

- There is insufficient evidence from one case series to evaluate the effectiveness or safety of PBT for bladder cancer in adults.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.

Description of included studies

No comparative studies of PBT for the treatment of bladder cancer that met inclusion criteria were identified.

Only one retrospective case-series (high-risk of bias) that met inclusion criteria was identified which evaluated patients with muscle invasive bladder cancer who were treated with trimodal bladder-preserving therapy (maximal transurethral resection, small pelvis photon irradiation, intra-arterial chemotherapy) for curative intent; patients with a complete response underwent subsequent PBT.[271] A total of 70 patients (74% male, median age 65 years) received PBT (52% of all patients enrolled); they had stage 2 (27%) or 3 (73%) cancer without nodal involvement or metastases. Total PBT dose was 36.3 GyE given in 11 fractions; total dose of radiation (to include photons) to the pelvis was 77.7 Gy in 34 fractions. Median follow-up was 3.4 years.

Results

Key Question 1 (Effectiveness, curative intent)

Survival outcomes

The 3-, 5-, and 10-year probabilities, respectively, of overall survival were 90%, 82% and 78% and of progression-free survival were 80%, 77% and 73%.[271] Metastatic disease was noted during the course of follow-up in eight (11%) patients (four pelvic lymph node, three lung and one peritoneum).

Key Question 3 (Safety)

The authors do not distinguish between toxicities specifically attributed to radiation therapy/PBT versus concurrent treatments (e.g., chemotherapy); it is unclear to what degree PBT was associated with the following events.

Acute grade 3 hematological toxicities were seen in 26% (18/70) of patients (leukopenia predominately, 21%); there were no grade 4 hematological events.[271] One acute grade 4 thromboembolic event occurred; no other non-hematological toxicities grade 3 or higher were observed. Regarding late

toxicity, two patients (3%) had a grade 3 urinary tract obstruction, but were resolved by conservative treatments; there were no grade 4 late toxicities reported.

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety), and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.2 Bone Tumors

Key Points

- There is insufficient evidence from seven case series to evaluate the effectiveness or safety of PBT for bone tumors in adults.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.

Description of included studies

No comparative studies of PBT for the treatment of bone tumors that met inclusion criteria were identified.

Seven case-series (N = 33 to 126) evaluating PBT for the treatment of primarily chordomas and chondrosarcomas of the spine that met inclusion criteria were identified.[16,49,63,121,130,241,264,267] Two studies[49,63] also included a small number of patients (<20% of both populations) with osteosarcomas, other sarcomas, or unknown types of sarcomas. Tumor location along the spine varied by study and primarily included thoracic, lumbar, and sacral tumors. Consistent with the prior report, studies that were majority cervical or skull-based chordomas/chondrosarcomas can be found in the section on Head and Neck cancers. Two studies with a minority (20% and 46%) of patients with cervical bone tumors are included in this section.[130,264] Median patient age ranged from 53 to 71 years and the majority were male (52.5% to 72.5% across 6 studies). One study reported all patients received resection prior to RT.[241] Treatment intent was curative in three studies[16,49,130] and either curative (primarily) or salvage in four studies.[63,121,241,264,267] In two studies[16,264,267], PBT was the sole intervention and was delivered via PBS technique in one[264,267] study; PBT technique not reported in the other study. In the remaining four studies patients received a combination of PBT and photon RT. PBT delivery technique was passive scatter in two of these studies[63,241]; the other studies did not report PBT technique. When reported, photon RT technique varied across studies. Total median radiation dose ranged from 70.2 to 74 Gy (across 6 studies). Median follow-up periods varied widely across the studies, range 12.9 to 87.6 months.

All case series are considered to be high risk of bias.

Results

Key Question 1 (Effectiveness, curative intent)

Primary and other outcomes

The probability of 5-year OS was reported by five case series, two[49,130] in treatment for purely curative intent (N range, 40 to 50) and three[63,241,264,267] in mixed treatment intent for curative and salvage (N range, 50 to 126). Across the studies for treatment for purely curative intent, the range of probabilities were 81.9% to 88.7%. Across the treatment for mixed curative and salvage treatment, the range of probabilities were 81% to 84%.

One study[16] in patients receiving treatment for purely curative intent reported probability of 3-year progression free survival to be 89.6%. Three studies[63,241,264,267] (N range, 50 to 126) for mixed curative and salvage treatment reported probability of 5-year local control which ranged from 61% to 81%.

For other outcomes reported across the case series, see Appendix F.

Key Question 3 (Safety)

Unless designated below as radiation-related, authors do not distinguish between toxicities specifically attributed to radiation therapy/PBT versus concurrent treatments (e.g., surgery); it is unclear to what degree PBT was associated with those events.

Radiation-related secondary malignancies occurred infrequently across four case series[121,130,241,264] (N range, 40 to 126) ranging from 0% to 2% (median F/U range, 44.4 to 65.5 months). Grade ≥3 sacral/vertebral fracture was reported in 6 studies[16,63,121,130,241,264] (N range, 33 to 126; Median F/U range, 37 to 87.6 months) ranging from 2% to 25%. Two of these studies[121,130] did not report grade of fracture, including the study reporting 25%, which could be cause for the high proportion of patients with fracture reported on in this study. The frequency of bone/soft tissue necrosis (any grade) and radiation-related spinal cord injury ranged from 0% to 5.9% (3 studies; N range, 40 to 126)[121,130,241] and from 0% to 1.5% (3 studies; N range, 40 to 68)[49,63,130], respectively. One study[241] (N=126) reported that 7.1% of patients presented with radiation-related deterioration in neurological status.

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety), Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.3 Brain, Spinal, Paraspinal Tumors

Key Points

- Results were inconsistent across two retrospective case-matched cohorts evaluating adult patients with different types of brain tumors undergoing treatment for curative intent. In one retrospective cohort, there was no statistical difference in the probability of 1-3 year OS and 1-2 year PFS following photon RT plus a PBT boost versus photon RT alone in patients with high-grade glioblastoma; those receiving PBT boost tended to have higher PFS but lower OS versus

those receiving photon alone and differences may be clinically meaningful. One large database study of primarily high-grade glioma reported statistically higher 5-year overall survival following PBT alone versus photon RT alone. (Low SOE for both comparisons).

- One small retrospective cohort study in patients with metastatic CNS disease found no statistical difference between salvage PBT compared with photons in the probability of 6-month OS or of CNS relapse; at 1 year, OS was better in the PBT group but statistical testing was not done and sample size was small (Insufficient SOE).
- For safety, no statistical differences were seen between groups in the frequency of acute grade 3 toxicity across both studies or of radiation necrosis (1 study of curative intent) or severe CNS toxicity (1 study of salvage therapy) over the late term (Low SOE for curative intent; Insufficient SOE for salvage therapy).
- No studies meeting inclusion criteria were identified that evaluated differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of included studies

A total of five retrospective comparative cohort studies that met inclusion criteria were identified that compared PBT with photon radiation therapies in adult patients with various brain or spinal tumors. Three studies evaluated radiation therapies for curative intent[13,37,129] and two for salvage therapy[94,196] (Table 17 and Table 18). In addition, six case series were identified across various tumor types; five evaluated PBT for curative intent[24,70,175,188,197] and one for salvage therapy.[137] For the reasons stated in the previous section, all comparative cohort studies were considered to be at moderately high risk of bias; however, the three studies[13,37,129] of curative intent did control for confounding. All case series are considered high risk of bias.

Table 17. Study and patient characteristics from comparative studies reporting on effectiveness and safety: Adult Brain, Spinal, & Paraspinal Tumors

	Effectiveness and Safety			
	Adeberg 2017		Gunther 2017	
	Photon + PBT Boost (n=66)	Photon RT (n=66)	PBT (n=14)	Photon RT (n=23)
Patient Characteristics				
Males, % (n)	63.6%	57.6%	57%	65%
Age, years; median (range)	57.9 (20 to 77)	57.9 (21.6 to 77.9)	37 (26 to 51)	39 (28 to 45)
Tumor Characteristics, %				
Subtype				
Glioblastoma	95.4%	95.4%	---	---
Astrocytoma	3%	3%	---	---
Oligodendroglioma	1.6%	1.6%	---	---
CNS Involved Tumors	---	---	100%*	
Gross Residual Tumor Size				
<1.5 cm^2	74%	81%	---	---
≥1.5 cm^2	26%	19%	---	---
Radiation Treatment Characteristics				
Treatment Intent	Curative (100%)		Salvage (78%)/Curative (22%)	
Technique	Boost	---	Passive Scatter	---
Median total dose (Gy)	Photon: 50 Proton Boost: 10	60	21.8	24
Gy per fraction	Photon: 2.0 Proton Boost: 2.0	2.0	---	---
Additional Treatments				
Treatment Prior to Radiation				
Biopsy Only	9.7%	6.6%	---	---
Subtotal Resection	80.3%	84.8%	---	---
Gross Total Resection	---	---	---	---
Concurrent/Adjuvant Treatment				
Chemotherapy	93.9%†	87.9%†	Yes§	Yes§
Stem Cell Transplant	---	---	100%	
Study Design	**Retrospective Matched Pairs Comparative Cohort**		**Retrospective Comparative Cohort**	
Follow-up (% followed)	15 months (NR)	15 months (NR)	8 months (54%)‡	8 months (54%)‡
Risk of bias	**Moderately High**		**Moderately High**	

cm = centimeters; Gy = Gray; IMRT = intensity modulated radiation therapy; PBT = Proton Beam Therapy; RT = radiation therapy.
*To include acute Lymphoblastic leukemia, acute myeloblastic leukemia, chronic lymphocytic leukemia, chronic myelocytic leukemia, lymphoma (not otherwise specified), and myeloma.
†chemotherapy could have been delivered as either adjuvant or concurrent treatment.
‡Seventeen (46%) patients died; no difference in loss-to follow-up groups.
§ Authors report that patients typically received multiple salvage chemotherapy regimens prior to radiation, but no actual data is reported.

Table 18. Study and patient characteristics from comparative studies reporting on effectiveness only or safety only: Adult Brain, Spinal, & Paraspinal Tumors

	Safety only		Effectiveness only						
	Bronk 2018		Mozes 2017			Jhaveri 2018*			
	PBT (n=34)	Photon RT (n=65)	PBT (n=27)	Photon RT (n=16)	Photon RT (n=23)	PBT (n=170)	Photon RT (n=49,405)	PBT (n=161)	Photon RT (n=161)
Patient Characteristics									
Males, % (n)	64.7%	64.6%	14.8%	31.3%	26.1%	58.6%		59.7%	59.7%
Age, years; median (range)	All patients: 48 (24 to 94)		---	---	---	Mean (SD): 57.3 (13.96)		Mean (SD): 49.4 (0.88	Mean (SD): 49.4 (14.51)
Tumor Characteristics, %									
Tumor Subtype									
Intracranial Meningioma	---	---	100%			---			
Astrocytoma	26.5%	35.4%	---			---			
Oligodendroglioma	73.5%	64.6%	---			---			
High Grade Glioma	---	---	---			91.2%		26.7%	26.7%
Low Grade Glioma						9.8%		73.3%	73.3%
Mean initial TV ± SD, cm^3	---	---	26.1 ± 22.2	37.3 ± 29.5	26.7 ± 23.1	---			
Stage									
I	---	---	63%	44%	39.1%	---		---	---
II	52.9%	27.7%	0%	19%	17.4%	---		---	---
III	47.1%	72.3%	0%	12%	8.7%	---		---	---
Unknown	---	---	37%	25%	34.8%	---		---	---
Radiation Treatment Characteristics									
Treatment Intent	Curative (100%)		Residual and Recurrent (67%)/Curative (33%)			Curative (100%)			
Technique	Passive scatter: 85.3%	IMRT	---	IMRT	FSRT	PBT (n=170) 3DCRT (n=5,196)		---	---

	Safety only		Effectiveness only						
	Bronk 2018		Mozes 2017			Jhaveri 2018*			
	PBT (n=34)	Photon RT (n=65)	PBT (n=27)	Photon RT (n=16)	Photon RT (n=23)	PBT (n=170)	Photon RT (n=49,405)	PBT (n=161)	Photon RT (n=161)
	Scanning beam: 14.7%					IMRT (n=20,215) Photon RT NOS (n=23,994)			
Median total dose (Gy)	Oligodendroglioma: 54 Astrocytoma: 50.4	Oligodendroglioma: 57 Astrocytoma: 57	56	56	56	---	---	---	---
Gy per fraction			1.8 or 2.0	1.8 or 2.0	1.8 or 2.0	---	---	---	---
Additional Treatments									
Prior to RT									
Gross Total Resection	64.7%	66.2%	---	---	---	12.2%		---	---
Subtotal Resection	35.3%	33.8%	---	---	---	11.9%		---	---
Biopsy	---	---	---	---	---	9.8%		---	---
Other	---	---	---	---	---	55%		---	---
Unkown	---	---	---	---	---	11.1%		---	---
Chemotherapy						83.6%		---	---
Adjuvant/Concurrent Treatment									
Adjuvant chemotherapy	2.9%	20.0%	---	---	---	---		---	---
Concurrent chemotherapy	52.9%	55.4%	---	---	---	---		---	---
Study Design	**Retrospective Comparative Cohort**		**Retrospective Matched Pairs Comparative Cohort**			**Retrospective Comparative Cohort**		**Retrospective Matched Pairs Comparative Cohort**	
Follow-up (% followed)	34 months (NR)	46 months (NR)	24 months (NR)	24 months (NR)	24 months (NR)	50.3 (NR)	62.3 (NR)	---	---
Risk of bias	**Moderately High**		**Moderately High**			**Moderately High**			

cm = centimeters; CNS = Central Nervous System; FSRT = fractionated stereotactic radiation therapy; Gy = Gray; IMRT = intensity modulated radiation therapy; KQ = Key Question; PBT = Proton Beam Therapy; RT = radiation therapy

*Jhaveri 2018 provides data from an overall cohort as well as a matched pairs cohort; data for both cohorts are reported here.

Key Question 1 (Effectiveness, curative intent)

Two comparative studies of PBT for curative intent provided data for effectiveness[13,129] (Table 17 and Table 18); the third study reported only safety outcomes and is described in Key Question 3 below (Table 18).[37] One study performed a single-center, matched-pair analysis of adults with glioblastoma who underwent photon radiotherapy (RT) followed by a proton boost (n=66) versus photon RT alone (n=66).[13] The second study queried the National Cancer Databases (NCBD) for patients with primary glioma treated with either PBT (N=170) or proton RT (N=49,405).[129] In both studies, the patients had high-grade forms of the disease (91% high-grade in the study of glioma). Median patient age across the studies was 58 and 59 years and the majority were male (60% and 65%). In the single-center study of glioblastoma, the majority of patients (83%) had subtotal surgical resection upon study entry compared to only 12% in the database study of glioma; other surgical interventions specified in the latter study included gross total resection (12%) and others (55%) PBT technique was not specified in either study. Median total radiation dose was 60 Gy in both groups in both studies and all patients received chemotherapy.

Additionally, three case series evaluated the effectiveness of PBT for curative intent across different types of brain tumors.[24,188,197] One of these series included both malignant (WHO grade 2/3) and benign (WHO grade 1) meningiomas[197]; only data for those patients with malignant tumors is described here and information regarding the benign population can be found in the section on Other Benign Tumors.

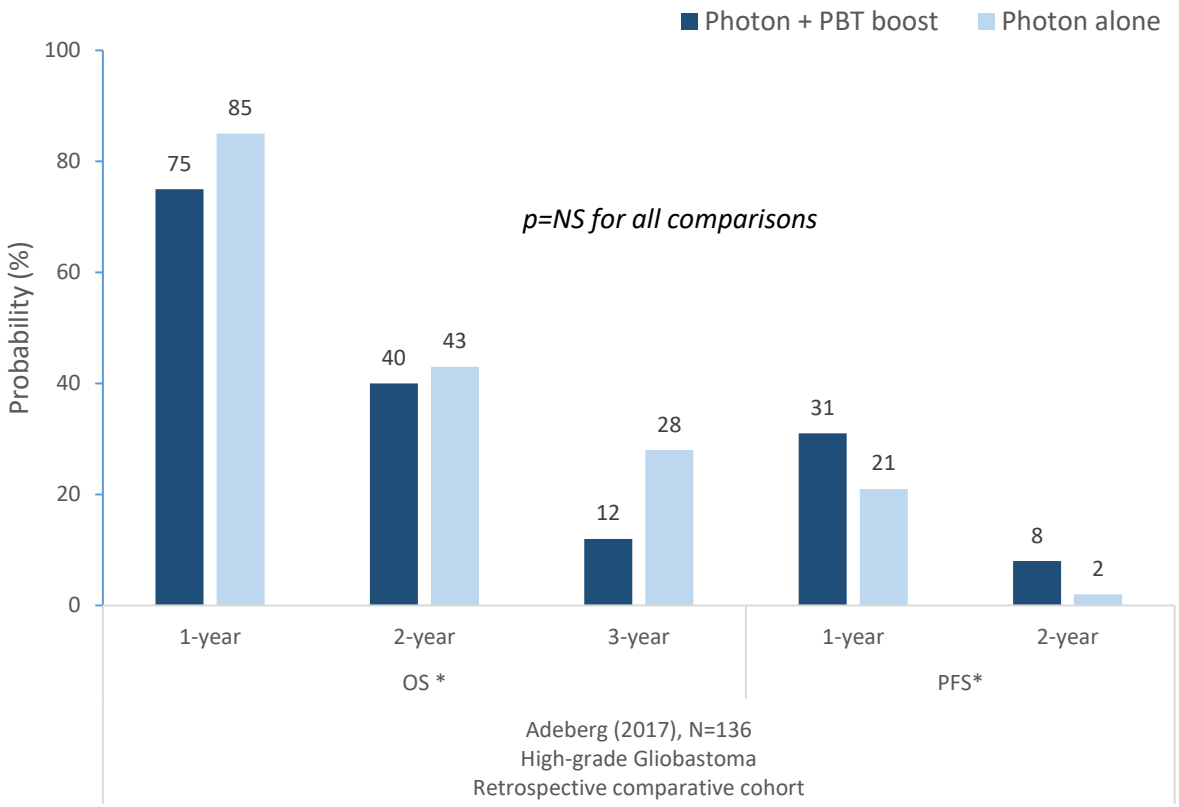
Survival outcomes

Comparative studies

In the single-center study of high-grade glioblastoma, PBT boost plus photon resulted in a slightly lower probability of overall survival (OS) through but a slightly greater probability of progression-free survival (PFS) compared with photon RT alone but no statistically significant differences were observed at any timepoint (Figure 8).[13] Sample sizes may have contributed to the failure to find statistical differences; differences may be clinically meaningful.

In the large database study of primary glioma, PBT was associated with greater OS on multivariate analysis compared with any photon RT (adjusted HR 0.66, 95% CI 0.53 to 0.83); the association persisted when PBT was compared with 3D-CRT, IMRT and other photon RT individually as well as when patients were stratified by high-and low-grade glioma (see Table C3 in Abstraction Appendix for details). The authors also conducted a propensity score-matching analysis (N= 161 in each group) to further control for selection bias and reported a significantly greater probability of 5-year OS following PBT compared with any photon RT (46.1% vs. 35.5%, p=0.0009) and with IMRT (p=0.01) and 3D-CRT (p=0.007), specifically. Of note, the PBT group had a significantly lower median follow-up period than the photon group and this difference was not controlled for: 50.3 versus 62.3 months (median 62.1 months for all patients).

Figure 8. Probability of OS and PFS in a Retrospective Cohort Study of PBT versus Photon Therapy for Adults with Brain Tumors.



*All data estimated from figures/graphs in the article. Sample sizes may have contributed to the failure to find statistical differences; differences may be clinically meaningful.

Case series

The probabilities of 5-year OS following PBT were similar across two case series involving malignant meningiomas or mixed tumor types (primarily medulloblastoma and germ cell tumors): 81% and 84%.[24,197] In the third case series evaluating patients with glioblastoma multiforme, the 2 and 5 year OS probabilities were low: 48% and 30%, respectively.[188]

The probability of 5-year PFS or local control following PBT was 68% as reported by two case series, one involving patients with primarily medulloblastoma and germ cell tumors[24] and the other malignant meningiomas.[197] In the case series evaluating glioblastoma multiforme, the probabilities of PFS were 37% and 12% at 1 and 2 years, respectively.[188]

Key Question 2 (Effectiveness, salvage therapy)

Two small retrospective comparative studies provided data for effectiveness (Table 17 and Table 18).[94,196] Additionally, one small case series (n=16; 58% female; median age 21 years) evaluating PBT for primarily salvage therapy in patients with central neurocytomas was included.[137]

One matched-pairs cohort study (N=66) included patients with inoperable, residual or recurrent intracranial meningioma undergoing predominately salvage treatment (67%) with PBT versus with IMRT or FSRT.[196] Tumor grade was unknown in 37% of PBT patients and 31% of photon patients; the remaining 63% in the PBT group had benign tumors versus 41% in the photon groups [the remaining photon patients had malignant tumors of grade 2 (18%) and grade 3 (10%)]. Patient age was not reported; there were more females in the PBT group compared with the photon groups (85% vs. 72%). PBT technique was not reported; all groups received a median total dose of radiation of 56 Gy (RBE). Whether or not patients received concurrent or adjuvant receipt of chemotherapy was not reported; neither were previous treatments. Follow-up was 24 months.

The second cohort study (N=37) included patients with CNS metastases from hematological malignancies, primarily acute lymphoblastic leukemia (49%) and acute myeloblastic leukemia (22%). Median patient age was 38 years and 62% were male. All patients were undergoing primarily salvage (78%) craniospinal radiation therapy prior to stem cell transplantation, either passive scatter PBT (total dose 21.8 Gy) or photon therapy (total dose 24 Gy). Authors report that patients typically received multiple salvage chemotherapy regimens prior to radiation, but no actual data is reported and timing is unclear. Median follow-up was 8 months.

Survival and relapse outcomes

Comparative studies

The cohort in patients with CNS metastases reported no statistical difference between PBT and photon in the probability of 6-month OS (79% vs. 70%); at 1 year, as estimated from the graph provided in the article, OS was 70% vs. 38%, respectively, but statistical testing was not done.[94] One patient in the PBT group (7%), compared with none in the photon group, experienced CNS relapse at 5 months (this patient also had concurrent systemic relapse and died from disease); there was no statistical difference between groups in the risk of CNS relapse (p=1.0).

Case series

In the small case-series, the probability of 5-year PFS was 100% following salvage (primarily) PBT for central neurocytomas.[137]

Other outcomes

The matched-pairs cohort study in patients with intracranial meningioma reported only change in tumor volume (TV) following PBT compared with photon therapy.[196] All groups showed statistically significant reduction in absolute TV shrinkage compared with baseline; however there were no statistical differences between groups at either 1 or 2 years. Mean change at 2 years versus baseline for PBT compared with IMRT and FSRT, respectively, was -3.7 ± 4.6 cm^3 vs. -4.3 ± 4.1 cm^3 and -7.0 ± 14.7 cm^3 (corresponding relative TV: 86.2% ± 9.2% vs. 69.4% ± 17.7% vs. 77.0% ± 14.6%).

Key Question 3 (Safety)

Comparative studies of curative intent

In addition to the case-matched study included in the section on effectiveness, a second comparative cohort study was identified that reported on the rate of pseudoprogression following PBT versus IMRT.[37] This study included 99 patients with low grade and anaplastic glioma; most were male (64%) with a median age of 48 years (Table 18). Total radiation doses were 57 Gy (RBE) (PBT) and 54 Gy (IMRT). PBT was delivered with either passive scatter (primarily) or scanning beam technique. Half of the patients received adjuvant chemotherapy (54%) with no differences between groups; fewer patients in the PBT group received concurrent chemotherapy (3% vs. 20%, $p<0.01$).

In the case-matched cohort of patients with high-grade glioblastoma, no cases of treatment-related grade ≥3 toxicity were observed in the PBT boost group compared with five cases in the photon only group ($p<0.02$) (

Table 19).[13] All patients in this study were also receiving chemotherapy. No statistical differences between groups were seen in the proportion of patients experiencing either worsening of preexisting symptoms or new deficits following treatment; most patients with pre-therapeutic deficits showed a stable deficit level after radiotherapy. No radiation necrosis outside the treatment field was reported in either group.

Pseudoprogression (assessed via MRI) was reported by both studies,[13,37] with similar frequencies between the PBT and photon groups

Table 19). In the study by Bronk et al. authors report that 3-year PFS and OS were significantly improved in patients with pseudoprogression (regardless of radiation modality received). The significance of these finding is unclear.

Table 19. Safety Results from Retrospective Comparative Cohort Studies in Adults with Brain, Spinal, or Paraspinal Tumors.

Author, Year, N, Tumor Type Study Design, RoB	Outcome	Timing	Photon + PBT boost, % (n/N)	Photon alone, % (n/N)	p-value
Adeberg 2017 (N = 136) High-grade Glioblastoma Retrospective cohort Moderately high	Acute Toxicity (any), ≥Grade 2	≤3 mos.	9% (6/66)	14% (9/66)	NR
	Acute Toxicity (any), Grade 3	≤3 mos.	0% (0/66)	7.5% (5/66)	p=0.02
	Neurocognitive deficits*	NR (median f/u 15 mos.)			
	Worse (vs. baseline)		3% (2/66)	6% (4/66)	p=NS
	New		9% (6/66)	2% (2/66)	p=NS
	Sensorimotor deficits*				
	Worse (vs. baseline)		3% (2/66)	5% (3/66)	p=NS
	New		11% (7/66)	14% (9/66)	p=NS
	Seizures*				
	Worse (vs. baseline)		0% (0/66)	0% (0/66)	NA
	New		2% (1/66)	6% (4/66)	p=NS
	Radiation necrosis (outside treatment field)		0% (0/66)	0% (0/66)	NA
	Pseudoprogression		8% (4/66)	8% (4/66)	NA
Bronk 2018 (N=99) Oligodendroglioma or astrocytoma† Retrospective cohort Moderately high	**Outcome**	**Timing**	**PBT, % (n/N)**	**IMRT, % (n/N)**	**p-value**
	Pseudoprogression	NR (median f/u 42 mos.)	14.7% (5/34)‡	13.8% (9/65)‡	p=1.0

CI = confidence interval; f/u = follow-up; NS = not statistically significant; PBT = Proton beam therapy; RoB = risk of bias.
*Authors describe these as/along with toxicity. As baseline in the PBT vs. photon groups, neurocognitive deficits, sensorimotor deficits, and seizures were presents in 30% (20/66) vs. 42% (28/66), 39% (26/66) vs. 30% (20/66), and 6% (4/66) vs. 3% (2/66), respectively.
†Similarly, no statistical difference between radiation modalities for the subgroups of patients with oligodendrogiloma [16% (4/25) vs. 14.3% (6/42)] and astrocytoma [11.1% (1/9) vss. 13% (3/23)].

Case series of curative intent

Five case series involving a variety of different brain tumors reported safety outcomes following PBT.[24,70,175,188,197] Acute toxicities of grade ≥3 ranged from 8% to 17% across four studies evaluating patients with glioma, medulloblastoma (primarily), and glioblastoma multiforme[24,70,175,188] while in the case series evaluating patients with malignant meningioma the frequency was 1%.[197] Late toxicities were reported by two case series. In one study of malignant meningiomas the frequency of late toxicities of grade ≥3 was 4% and included one case of brain edema and three cases of brain necrosis (one of which resulted in death).[197] In the second study (glioblastoma multiforme) 24% of patients presented with radiation necrosis.[188]

Comparative studies of salvage therapy

Only the cohort study evaluating patients with CNS metastases from hematological malignancies undergoing pre-stem cell transplantation craniospinal radiation therapy reported safety outcomes (Abstraction Appendix C, Tables C2 and C3).[94] For acute toxicity, the frequency of mucositis (of any grade) was lower following PBT (n=14) versus photons (n=23) (7% vs. 44%, p=0.03; RR 0.16, 95% CI 0.02 to 1.1); however, there was no difference between groups in severe/grade 3 mucositis (7% vs. 9%; one and two events, respectively). No statistical differences were seen between groups, respectively, in the frequency of any gastrointestinal toxicity (29% vs. 30%), any CNS toxicity (21% vs. 13%), or infection (57% vs. 35%). Over the long-term, only one case of severe neurotoxicity (characterized by diffuse demyelination and necrosis, neurocognitive impairment, lower extremity weakness, incontinence, difficulty swallowing) was reported and it occurred in the PBT group (7% vs. 0%, p=0.38). Sample size may have played a role in the lack of statistical significance.

Case series

There were no grade ≥3 PBT-related toxicities in the small case series evaluating central neurocytomas. Four patients (25%) experienced permanent mild-to-moderate concentration impairment. The most common PBT-related adverse events (i.e., fatigue, alopecia, and radiation dermatitis) were transient (resolved within 6 months post-PBT).[137]

Key Question 4 (Differential Effectiveness and Safety) and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.4 Breast Cancer

Key Points

- There is low strength of evidence from one retrospective comparative database study that there is no statistical difference in the probability of OS at 5 years between PBT versus photon with or without electron boost therapy for treatment of breast cancer.
- One moderate quality cost-utility study (QHES 73/100) concluded that, compared with photon therapy, PBT was not cost effective in women without cardiac risk factors (CRF) or PBT mean heart radiation doses <5 Gy. PBT is more likely to be cost-effective for patients with higher risk of coronary heart disease (CHD) and for younger patients (40 or 50 years old versus 60 years old); authors indicate a societal perspective, however indirect societal costs were not described.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy or differential effectiveness and safety in this population.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of included studies

A total of six studies evaluating PBT for curative intent for the treatment of breast cancer that met inclusion criteria were identified: one retrospective comparative database study,[48] one retrospective comparative cohort study[274], and four case series.[41,57,214,292]

In addition, one cost-utility analysis comparing PBT with conventional radiation therapy for breast cancer that met inclusion criteria was identified.[171]

Results

Key Question 1 (Effectiveness, curative intent)

Two comparative studies that met inclusion criteria that evaluated the effectiveness of PBT for breast cancer treatment for curative intent were identified. One large, retrospective comparative database study (National Cancer Data Base)[48] evaluated patients with non-metastatic breast cancer treated with adjuvant PBT (n=871) versus photons/photons plus electron boost (n=723,621) following either breast-conserving surgery (80%) or mastectomy (20%), Table 20. The vast majority of patients were female (99%) with a median age of 60 years; most disease was stage 0 to I (56%). Aside from the breast, additional lymph node irradiation was indicated in 22% of patients. Other treatments received included chemotherapy in 46% and endocrine therapy in 69% of the population. The median total radiation doses were similar between the PBT and photon/electron groups [60.0 Gy(RBE) vs. 60.4 Gy, respectively]. Median follow up was significantly longer for the PBT group: 74.6 vs. 62.2 months.

The second study was a cross-sectional survey study in disease-free survivors of early breast cancer who were greater than 5 years post-diagnosis (N=129) and compared adjunctive partial breast PBT versus whole body radiation (WBI)[274], Table 20. Surveys were mailed to 180 eligible patients who had had PBT or WBI; 142 surveys were completed (79% response rate). Baseline differences across most characteristics were similar between PBT and WBI with the exception of less time since diagnosis (means of 7.44 versus 6.23 years respectively) in the WBI group and a higher ratio of Caucasian to non-Caucasian patients in the PBT group. Authors do not report adjusting for potentially confounding variables. Median follow up was significantly longer for the PBT group: 84 versus 72 months. For the reasons stated in the previous section, all comparative cohort studies were considered to be at moderately high risk of bias; however, the large database study did control for confounding.[48] All case series are considered high risk of bias.

Table 20. Study and patient characteristics from comparative studies reporting on effectiveness only and effectiveness and safety: Breast Cancer

	Effectiveness Only			
	Chowdhary 2019		Teichman 2018	
	PBT (n=871)	Photon RT* (n=723,621)	PBPT (n=72)	WBI with photons (n=57)
Patient Characteristics				
Males, %	0.6%			
Age, years; mean ± SD	59 (NR)	60 (NR)	72.5 (53 to 94)†	70 (46 to 86) †
Tumor characteristics				
Left laterality, %	54.1%	50.5%	56.9%	50.9%
Stage				
0	13.5%	10.1%	20.8%	21.1%
I	44.8%	46.2%	66.7%	66.7%
II	23%	27.3%	12.5%	12.3%
III	15.2%	12.8%	---	---
Unknown	3.6%	3.6%	---	---
Comorbidites				
Charlson-Deyo comorbidity score				
0	90.7%	86.4%	---	---
1	7%	11.5%	---	---
≥2	2.3%	2.2%	---	---
Radiation Treatment Characteristics				
Technique	---	---	---	---
Median total dose (Gy)	60	60.4	40 CGE	50 Gy to entire breast + 10 Gy boost to the tumor bed
No. fractions	---	---	10 daily	---
Additional Treatment				
Prior to Radiation				
Surgery	100%		100%	
Breast-conserving surgery	76.6%	79.9%	---	---
Mastectomy	23.3%	20%	---	---
Not Specified	0.1%	0.1%	100%‡	
Timing NOS				
Chemotherapy	42.9%	45.7%	0%	0%
Endocrine Therapy	63.9%	68.9%	---	---
Lymph node irradiation	23.7%	22.2%	---	---
Study Design	**Retrospective Comparative Database Study**		**Retrospective Comparative Cohort**	
Median Follow-up, months (% followed)	**74.6 (NR)**	**62.2 (NR)**	**84 (93.5%, all patients)**	**72 (93.5%, all patients)**
Risk of bias	**Moderately High**		**Moderately High**	

CGE = Cobalt Gray Equivalent; Gy = Gray; NOS = not otherwise specified; NR = not reported; PBT = proton beam therapy; PBPT = Partial Breast Proton Therapy; RT = radiation therapy; SD = standard deviation; WBI = Whole breast irradiation

*Patients could have received either Photon RT alone or Photon RT + electron boost.

†At time of follow-up

‡Most commonly a lumpectomy

Additionally, four case-series evaluating PBT for the treatment of breast cancer that met inclusion criteria were identified.[41,57,214,292] Sample sizes ranged from 30 to 100 and median patient age from 49 to 63 years (across 3 studies)[41,57,292]; one study included two (2%) male patients.[292] The primary histological subtype was invasive ductal carcinoma (3 studies); one case series included patient ductal carcinoma in-situ or invasive adenocarcinoma.[214] All studies included patients with stage I-III disease and in most cases patients were receiving PBT therapy post-mastectomy or lumpectomy. PBT technique varied across the studies and included 3-D uniform scanning (primarily), passive scattering (multi beam technique), and pencil beam scanning. Total PBT dose ranged from 34 to 50.4 Gy (RBE). The majority of patients had received either neoadjuvant or adjuvant chemotherapy. Follow-up periods varied widely across the studies, range 6 to 60 months. All case series are considered to be high risk of bias.

Primary and other outcomes

Comparative studies

The database study reported similar 5-year probabilities (unadjusted) of overall survival following PBT compared with photon/electron boost therapy: 91.9% vs. 88.9%; adjusted HR 0.85 (95% CI, 0.68 to 1.07), p=0.12.[48] A second multivariate analysis was conducted after stratifying for factors associated with increase heart doses; PBT, relative to photons/electrons, was not associated with OS for any of the stratified subsets (Abstraction Appendix Table D3).

The cross-sectional survey study compared responses to quality of measures for those who had received partial breast PBT with those who had had WBI.[274] Multiple validated measures (with multiple domains) and analysis modifications of some measures were used in addition to an investigator-designed instrument to assess quality of life and treatment satisfaction among early breast cancer survivors. Many of the domains for various measures may suggest that QOL may be somewhat higher for PBT recipients versus WBI in this population (see Data Abstraction Appendix D for details). For the Breast Cancer Treatment Outcome Scale (BCTOS), 9/22 domains were statistically significant, favoring PBT over WBI and author modeling. Similarly 5/10 domains related to body image significantly favored PBT and PBT recipients reported less fatigue on the brief fatigue inventory compared with WBI recipients. In interpreting these results, several factors should be considered. First, authors do not report use of statistical methods to correct for multiple comparisons, thus some significant results could be spurious and there was likely substantial correlation between some measures. A number of confounding factors may have impacted patient's recall and perceptions including time from diagnosis, the overall impact of all treatments provided (including surgical procedures), patient's beliefs regarding treatment effectiveness (most women who had received PBT had been enrolled in a formal clinical trial which may have led to an impression of receiving the best, newest treatment), recall of the impact of going through radiation therapy and others.

Case series

One case series (N=100)[41] reported 5-year probabilities of overall survival (95%) and disease-free survival (94%). Another study (N=91)[292] reported mortality (as opposed to OS); over a median follow-up period of 1.3 years, six patients (7%) died, five of whom had relapsed. A third case series (N=30)[57] reported only that no patient experienced disease progression or recurrence during follow-up (median 9 months).

Two case series reported the incidence of distant metastases which ranged from 3% (1/30)[57] to 11% (10/91)[292] over median follow-up periods of 9 months and 1.3 years, respectively.

Key Question 3 (Safety)

Unless designated below as radiation-related, authors do not distinguish between toxicities specifically attributed to radiation therapy/PBT versus concurrent treatments (e.g., surgery); it is unclear to what degree PBT was associated those events.

During the acute period, radiation dermatitis grade ≥3 occurred infrequently across all four case series, ranging from 0% to 5%.[41,57,214,292] In two case series[57,292], no incidences of acute grade 3 esophagitis were seen and grade 3 breast/chest wall pain was rare (0% and 1%).

Regarding late toxicities, the incidence of rib fracture was rare, ranging from 0% (2 studies)[41,57] to 2% (1 study)[292]; in the latter study, these two cases were uncomplicated and occurred at 13 and 39 months. Fat necrosis occurred infrequently as reported by two studies (one case each, 1% to 2% of patients)[41,214] over 6 and 12 months of follow-up; drainage was required in one case. In one case series,[292] three (3%) patients had clinically evident lymphedema at final follow-up (median 15 months) and required compression sleeves and/or pumps. The remaining late events as reported by two case series were generally mild and uneventful (e.g., telangiectasia, skin thickening, retraction/asymmetry).[41,214]

In two case series, there were no reported cases of lung toxicity or cardiac toxicity during follow-up.[41,57]

Key Question 5 (Economic)

One moderate quality CUA[171] (QHES 73/100) comparing PBT with conventional radiation therapy met the inclusion criteria (

Table 21). Study funding was not reported. The Markov model for hypothetical cohorts of women treated with PBT versus conventional radiotherapy for breast cancer was based on case series of PBT treatment and a case-control study evaluating risk of ischemic heart disease following conventional radiation therapy to estimate incremental cost-effectiveness ratios (ICER) reflecting the marginal cost of adding one quality adjusted life year (QALY) to a patient's life when using PBT versus conventional radiotherapy (photon therapy). The purpose of the modeling was to identify mean heart doses (MHDs) at which photon therapy would remain cost-effective compared with an average proton plan that yields a MHD of 0.5 Gy. Primary limitations of the study included lack of clarity or detail regarding model inputs, PBT clinical data from case-series, and costing of PBT.

Key points

One moderate quality study (QHES 73/100) concluded that, compared with photon therapy, PBT was not cost effective in women without cardiac risk factors (CRF)or PBT mean heart radiation doses <5 Gy. PBT is more likely to be cost-effective for patients with higher risk of coronary heart disease (CHD) and for younger patients (40 or 50 years old versus 60 years old); authors indicate a societal perspective, however indirect societal costs were not described.

- ICER varies by presence of CRF, dose combination and age: in 50 year old women without CRFs, ICERs estimated from author figures ranged from approximately $890,000/QALY (lowest doses) to approximately $90,000/QALY (highest doses) and for 50 year old women with at ≥1 CRF, from $90,000/QALY to $49,000/QALY.
- Sensitivity analyses: ICER overall range from $49,757/QALY to $161,285/QALY, suggesting substantial variation in cost-effectiveness.
- Limitations:
  - Unclear methodology for Markov modeling: transition probabilities, timing of transitions, and other modeling assumptions were not presented
  - Outcomes other than CHD and death were not modeled and utilities for specific states were not detailed
  - Clinical and radiation dose reduction data specific to PBT are based on small case series.
  - Costing for PBT was described in supplemental material but it is not clear if the costs capture all aspects of operation
  - Components of treatment costs for CHD were not reported and it is not clear that all applicable costs were included; CHD models included PCI but not coronary artery bypass grafting
  - Life-time horizon modeled, but comparative data on long-term data on PBT are not available.
  - Radiation-related risk for ischemic heart disease in women receiving RT between 1958 and 2001 may not reflect the impact of newer RT methods

Detailed results

Study characteristics and framework

One moderate quality CEA intended to identify the doses at which photon therapy would remain cost-effective compared to a mean heart dose (MHD) from PBT of 0.5 Gy for women with breast cancer (Mailhot Vega 2016). The costing year was 2012. A lifetime horizon was used. The study adopted a societal perspective; however, indirect societal costs were not described. Six hypothetical cohorts were

modeled: patients 40 years old with and without cardiac risk factors (CRFs), patients 50 years old with and without CRFs, and patients 60 years old with and without CRFs. Costs included treatment of CHD, including inpatient PCI and outpatient PCI. The costs for PBT and photon therapy were $21,933 and $13,552, respectively. The basis for radiation-related toxicity to the heart was from a population-based case-control study of 2,168 women who underwent radiotherapy for breast cancer between 1958 and 2001 in Sweden and Denmark.[61] The impact of newer forms of radiation therapy over this time frame is unclear. Data on dosimetry for PBT and photon therapy were based on a systematic review of MHDs published from 2003 to 2013 which included limited information on PBT.[273]

Base Case Results

For women without CRFs, PBT was not cost-effective at a WTP threshold of $50,000/QALY. At a WTP threshold of $100,000/QALY, PBT was cost effective at 10 Gy for 40 year-old women without CRFs and 9 Gy for 50 year-old women without CRFs. For women with ≥1 CRF, PBT was cost-effective at a WTP threshold of $50,000/QALY beginning at doses of 9 Gy and 10 Gy for women aged 50 years and 60 years, respectively. At a WTP threshold of $100,000/QALY, PBT was cost-effective for women with ≥1 CRF at the following doses: ≥6 Gy for 40 and 60 year-old women and ≥5 Gy for 50 year-old women. Depending on proton-photon dose combinations, in 50 year old women without CRFs ICERs estimated from author figures ranged from approximately $890,000/QALY (lowest doses) to approximately $90,000/QALY (highest doses) and for 50 year old women with at ≥1 CRF, from $90,000/QALY to $49,000/QALY.

Sensitivity Analyses

A probabilistic sensitivity analysis (PSA) was conducted.

Results from a PSA demonstrated that there were no doses at which PBT was cost-effective for women without CRFs at a WTP threshold of $50,000/QALY. However, at a WTP threshold of $100,000/QALY, PBT was cost-effective for women with CRFs at 7 Gy for 50 year-old women and at 9 Gy for 40 and 60 year-old women. For women with ≥1 CRF, PBT was cost-effective at a WTP threshold of $50,000/QALY at 9 Gy for 40 year-olds, 7 Gy for 50 year-olds, and 8 Gy for 60 year-olds. At a WTP threshold of $100,000/QALY PBT was cost-effective for women with ≥1 CRF at 5 Gy for 40 year-olds, 4 Gy for 50 year-olds, and 5 Gy for 60 year-olds. Additional sensitivity analyses were reported in supplemental material with a primary focus on women with ≥1 CRF based on "real-world" cases at specific lower proton/photon dose pairs. In women with ≥1 CRF, ICERs varied by age across the dose pairs chosen as follows: 40 year olds ($65,039/QALY to $161,285/QALY), 50 year olds ($49,757/QALY to $116,740/QALY) and 60 year olds ($60,282/QALY to $147,093/QALY). ICERs decreased when lowering the discount rate to 1% or 0% per annum and difference between age groups was more pronounced.

Conclusions and Limitations

The authors concluded that PBT is more likely to be cost-effective for patients with higher risk of CHD (≥1 CRF) and for younger patients. This is due to the lower risk of CHD at lower radiation doses and the accrual of more time living without CHD among younger patients. The purpose of the modeling was to identify MHDs at which photon therapy would remain cost-effective compared with an average proton plan that yields an MHD of 0.5 Gy.

The limitations of this study include unclear methodology for Markov modeling; transition probabilities, timing of transitions over the years, and other modeling assumptions were not presented. Data specific

to PBT and potential reduction in radiation exposure comes primarily from very small case series in breast cancer patients. Authors assume that PBT and photon therapy are equally effective, but comparative data to support this assumption were not described and the assumption was not included in sensitivity analyses. Full results of the PSA were not described. The costing method for PBT operation was described in supplemental material, but it is not clear to what extent these costs capture all aspects of operation. The components of treatment costs for CHD were not reported; costs for PCI were based on CMS CPT Codes and procedure allowable charges and it is not clear that all applicable costs were included. CHD treatment models included PCI (once per lifetime) only but not coronary artery bypass grafting. Outcomes other than CHD and death were not modeled and utilities for specific states were not detailed. The QHES score for this study was 73/100 points.

Table 21. Summary of the economic study comparing PBT with conventional RT in breast cancer patients

	Mailhot Vega 2016
Population	Women with breast cancer aged 40, 50, or 60; with or without CRFs
Intervention(s)	PBT (timing, intent unclear)
Comparator(s)	Photon therapy
Country	USA
Funding	NR
Study design	CUA
Perspective	Societal (as stated by authors; data for this perspective not provided; appears to be healthcare system perspective)
Time horizon	Lifetime
Analytic model	Markov model (stated, details unclear)
Effectiveness outcome	QALYs
Effectiveness outcomes	Risk of CHD, PCI, death from CHD, inpatient PCI
Source for effectiveness data	Prior literature (case series, population-based case-control study, SEER database)
Costing year	2012
Currency	USD
Discounting	3%
Components of cost data	Cost of treatment strategy, cost of CHD
Cost sources	Microcosting for treatment strategy; *Red Book* and CMS allowable costs for CHD; PCI estimates for in- vs. outpatient from New York University Hospital Operations department, representing average Medicare reimbursements
Sensitivity analysis	PSA: Monte Carlo simulations using 50,000 iterations
QHES	73/100
Results:	
Cost / QALY of intervention and comparators	NR
ICER	ICERS varied by dose, presence of cardiac risk factors and age: Range for 50 year old women: with no CRF $890,000/QALY (lowest doses) to $90,000/QALY (highest doses); with ≥1 CRF, $90,000/QALY to $49,000/QALY (estimates from author figures). Doses cost-effective at $50,000/QALY In women with no CRFs: none In women with ≥1 CRF: PBT cost-effective beginning at MHD 9 Gy and 10 Gy for women aged 50 years and 60 years Doses cost-effective at $100,000/QALY In women with no CRFs: MHD 10 Gy for 40 year-old women, 9 Gy for 50 year-old women

	Mailhot Vega 2016
	In women with ≥1 CRF: MHD ≥6 Gy for 40 and 60 year-old women and MHD ≥5 Gy for 50 year-old women
One-way SA	Not done
Other SA	No scenarios in which PBT is cost-effective at $50,000 in women w/o CRFs; PBT cost-effective at $100,000 for women w/o CRFs in all age groups (7 Gy for 50 year-old women, 9 Gy for 40 and 60 year-old women) For women with ≥1 CRF, ICERs ranged from $49,757/QALY to $161,285/QALY depending on age and dose
Author's Conclusion	For women without CRFs, PBT was not cost-effective at a WTP threshold of $50,000/QALY. PBT is more likely to be cost-effective for patients with higher risk of CHD and for younger patients.
Limitations	• Unclear methodology for Markov modeling: transition probabilities, timing, and other modeling assumptions were not presented • Outcomes other than CHD and death were not modeled; utilities for specific states were not detailed; • Lifetime horizon, but no comparative long-term data • PBT: not clear that costs capture all aspects of operation • Components of CHD treatment costs not reported; unclear whether all applicable costs were included; CHD models included PCI but not coronary artery bypass grafting • No Comparative data for effectiveness and harms; Data from case series on PBT and a case-control study of radiation-related risk for ischemic heart disease in women receiving RT between 1958 and 2001 (impact of newer RT methods is unclear)

CHD = coronary heart disease; CMS = Centers for Medicare and Medicaid Services; CRF = cardiac risk factor; CUA = cost-utility analysis; Gy = Gray (unit of absorbed dose); ICER = incremental cost-effectiveness ratio; IMRT = intensity-modulated radiation therapy; PBT = proton beam therapy; PCI = percutaneous coronary intervention; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life-year; QHES = Quality of Health Economic Studies; QOL = quality of life; RT = Radiation therapy; SA = sensitivity analysis; SEER = Surveillance, Epidemiology, and End Results Program; WTP = willingness-to-pay.

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety)

No studies that met inclusion criteria were identified.

Esophageal Cancer

Key Points

- Five retrospective comparative cohort studies that evaluated the effectiveness and safety of PBT compared with photon RT for curative intent in adult patients with esophageal cancer that met inclusion criteria were identified.
- With the exception of OS at 1 year which was similar between groups, probabilities of OS and PFS/DFS were greater following PBT versus IMRT or 3D-CRT over 1 to 5 years follow-up in two studies; however, statistical significance was achieved in only the largest study (Low SOE).
- Mortality (as opposed to OS) was reported by two studies with no statistically significant differences seen between the PBT and the photon groups (IMRT, 3D-CRT, XRT) (Low SOE for the large, higher quality study; Insufficient SOE for the small, poorer-quality study).
- For the comparison of PBT versus IMRT, with the exception of grade 4 radiation-induced lymphopenia (2 studies) and any would event (1 study) which were less common with PBT, all other RT-related and treatment-related toxicities did not differ statistically between groups. For PBT versus 3DCRT or XRT, with the exception of GI events, PBT was associated with a statistically less treatment-related toxicity (i.e., pulmonary, cardiac, and wound events; grades ≥2 or not specified) across three studies (Low SOE for all).
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of included studies

Five retrospective comparative cohort studies[77,164,173,255,317] that compared PBT with photon radiation therapies for curative intent in adult patients with esophageal cancer were identified (Table 22); one study[255] reported on safety only and will be described further under Key Question 3 below. Of the cohort studies, three evaluated PBT as part of definitive chemoradiotherapy[77,173,317] and two as part of trimodal therapy (in conjunction with surgery and chemotherapy)[164,255]. In addition, two case series of PBT for curative intent were identified.[122,269] No studies evaluating PBT for salvage therapy for the treatment of esophageal cancer that met inclusion criteria were identified.

For reasons stated previously (see section 4.1), the comparative cohort studies were all considered to be moderately high risk of bias; however, four of the five did control for confounding.[77,164,255,317] All case series were considered to be at high risk of bias (Appendix D).

Table 22. Esophageal Cancer in Adults: Study Characteristics and Demographics for Studies Comparing PBT versus Photon RT for Curative Intent

	Effectiveness and Safety									Safety only	
	Fang 2018		Lin 2017			Makishima 2015		Xi 2017		Shiraishi 2018	
	PBT (n=110)	Photon RT (p=110)	PBT (n=111)	Photon RT 1 (n=214)	Photon RT 2 (n=255)	PBT (n=25)	Photon RT (n=19)	PBT (n=132)	Photon RT (n=211)	PBT (n=136)	Photon RT (n=136)
Patient Characteristics											
Males, % (n)	94%	77%	89%	82%	87%	---	---	82%	79%	90%	87%
Age, years; median (range)	70 (41–86)	69 (44–84)	>65 years: 32% ≤65 years: 68%	>65 years: 36% ≤65 years: 64%	>65 years: 26% ≤65 years: 74%	---	---	>67 years: 30% ≤67 years: 71%	>67 years: 62% ≤67 years: 72%	63 (26–76)	60 (26–82)
Comorbidities											
Coronary Artery Disease	---	---	9%	15%	13%	---	---	---	---	---	---
Hypertension	---	---	61%*	49%*	49%*	---	---	---	---	---	---
Smoking†	---	---	18%*	29%*	24%	---	---	74%	72%	---	---
Tumor Characteristics											
Subtype											
Adenocarcinoma	71.8%	76.4%	96%	90%	94%	0%	0%	68%	74%	96%	98%
SCC	28.2%	23.6%	5%	10%	6%	100%	100%	32%	27%	4%	2%
Stage											
0	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%
I/II	39%	40%	36%	37%	36%	60%	26%	36%	33%	35%	40%
III/IV	60.9%	60%	64%	63%	64%	36%‡	74%‡	64%‡	67%‡	64%	60%
Tumor location											
Upper/Middle	23.6%	23.6%	1.8%*	11.7%*	5.5%	12%	36.8%	---	---	4%	3%
Lower	76.4%	76.4%	98.2%	88.3%	94.5%	88%	63.2%	---	---	96%	97%
Radiation Treatment Characteristics											
Technique	---	IMRT	---	3DCRT	IMRT	Passive Scatter	XRT	---	IMRT	---	IMRT
Median total dose (Gy)	50.4	50.4	50.4§	50.4§	50.4§	60	60	50.4	50.4	50.4	50.4
No. fractions	28	28	---	---	---	---	---	---	---	28	28
Additional Treatments											
Prior to Radiation Treatment											

	Effectiveness and Safety								Safety only	
	Fang 2018		Lin 2017			Makishima 2015		Xi 2017		Shiraishi 2018	
	PBT (n=110)	Photon RT (p=110)	PBT (n=111)	Photon RT 1 (n=214)	Photon RT 2 (n=255)	PBT (n=25)	Photon RT (n=19)	PBT (n=132)	Photon RT (n=211)	PBT (n=136)	Photon RT (n=136)
Biopsy	0%	0%	---	---	---	---	---	100%	100%	---	---
Surgical Resection	0%	0%	100%	100%	100%	---	---	8%	13%	100%	100%
Induction Chemotherapy	27.3%	28.2%	39%*	4%*	35%*	---	---	29%	28%	35%	37%
Concurrent/Adjuvant											
Chemotherapy	100%		100%	100%	100%	100%	100%	Yes**	Yes**	100%	100%
Study Design	**Retrospective Propensity Score Matched Comparative Cohort**		**Retrospective Comparative Cohort**			**Retrospective Comparative Cohort**		**Retrospective Comparative Cohort**		**Retrospective Propensity Score Matched Comparative Cohort**	
F/U, months (% followed)	**55 (NR)**		**F/U (NR) (%NR)**			**24 (75%)**	**20 (68%)**	**48.4 (NR)**	**65.1 (NR)**	**F/U (NR) (%NR)**	
Risk of bias	**Moderately High**		**Moderately High**			**Moderately High**		**Moderately High**		**Moderately High**	

3DCRT = Three-dimensional conformal radiotherapy; F/U = follow-up; Gy = Grey; IMRT = Intensity modulated radiation therapy; KQ = Key Question; NR = Not reported; PBT = Proton Beam Therapy; RT = radiation therapy; SCC = squamous cell carcinoma; XRT = X-ray chemoradiotherapy

*Indicative of a statistically significant different between groups.

† Defined as: Smoking at diagnosis (Lin 2017); History of Smoking (Xi 2017)

‡All patients had stage III disease. No stage IV patients included in group.

§Mean heart doses were 13.2 vs. 28.4 vs. 22.4 Gy for PBT, 3DCRT, and IMRT, respectively; mean lung doses were 6.1 vs. 10.5 vs. 9.5 Gy. Mean doses to the heart and lungs were significantly different ($p<0.0001$) between all groups.

**Authors report that patients generally received concurrent chemotherapy, but no data is provided.

Results

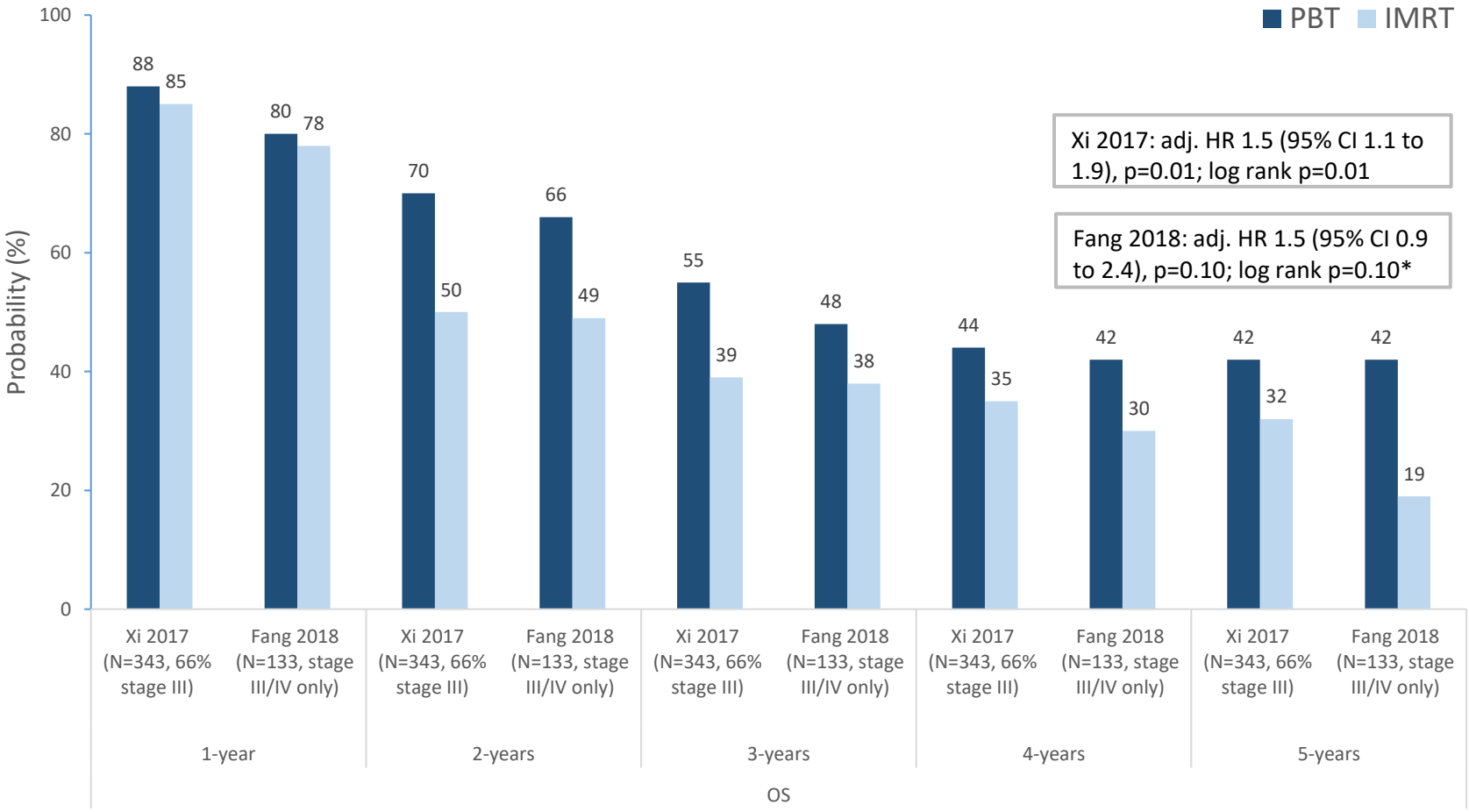
Key Question 1 (Effectiveness, curative intent)

Across the four comparative cohort studies that provided data on effectiveness,[77,164,173,317] sample sizes ranged from 44 to 580; of the three that reported demographics, patients were predominately male (range, 80% to 86%) and of older age (majority ≥65 years), Table 22. Comorbidities were poorly reported. In three studies, the tumor subtype was primarily adenocarcinoma (72% to 93%) and in the fourth, all tumors were squamous cell carcinoma.[173] Across all studies, most patients had stage III to IV disease. The technique used for PBT was reported by only one study (passive scatter technique).[173] Comparators included intensity modulated radiation therapy (IMRT) in three studies,[77,164,317] three-dimensional conformal radiotherapy (3D-CRT) in one,[164] and X-ray radiation therapy (XRT) in one.[173] Median total radiation dose for both treatment arms was similar across studies (range, 50.4 to 60 Gy). All patients were receiving concurrent and/or adjuvant chemotherapy. In one study, all patients had surgical resection.[164] Of note, one study was a propensity-score matched analysis.[77]

Survival outcomes

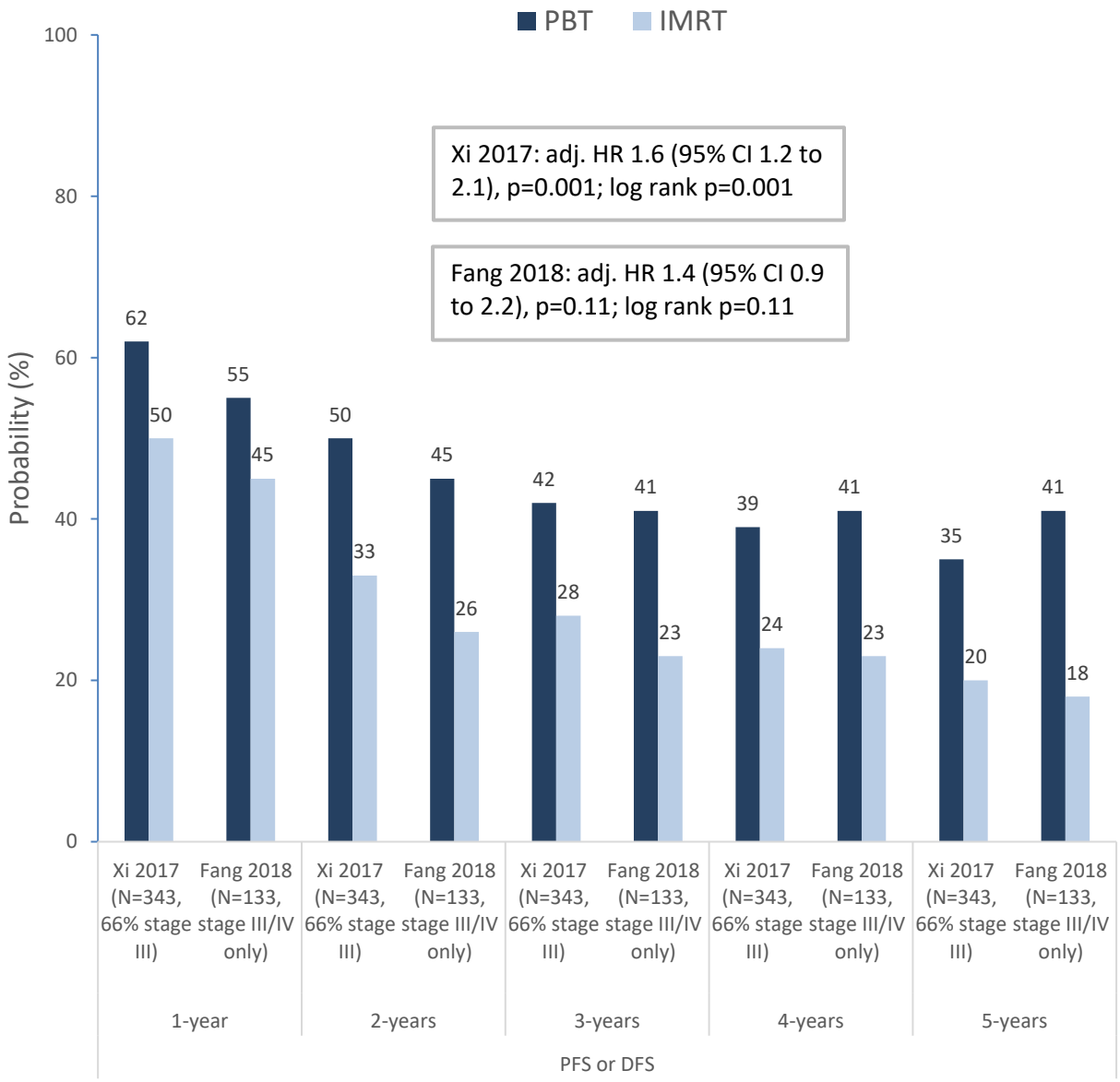
Two comparative studies reported both overall survival (OS) and progression-free (PFS) or disease-free survival (DFS).[77,317] The probability of both OS and PFS/DFS was greater following PBT versus IMRT across 1 to 5 years; however, only one of the studies found the differences between groups statistically significant (for OS: adjusted HR 1.5, 95% CI 1.1 to 1.9; for PFS: adjusted HR 1.6, 95% CI 1.2 to 2.1),[317] Figures 9 and 10. The study by Fang et al. reported OS and PFS only in patients with more advanced disease (57% stage III; 3% stage IV) whereas Xi et al., the study which reported a statistical difference, included patients of all stages (34% stage I/II; 66% stage III). The latter study conducted a subgroup analysis and found no statistical differences in 5-year OS or PFS between the PBT and IMRT groups for patients with stage I or II disease whereas probabilities were statistically higher for patients with stage III disease (OS: 35% vs. 25%, p=0.04; PFS: 34% vs. 13%, p=0.005). The difference in the proportion of patients with early versus advanced disease may partially explain the difference in statistical findings.

Figure 9. Probability of OS from Retrospective Cohort Studies Comparing Definitive Chemoradiotherapy using PBT versus IMRT for Curative Intent in Adults with Esophageal Cancer.



adj. = adjusted; CI= confidence interval; IMRT = intensity modulated proton therapy; NS = not statistically significant; OS = overall survival; PBT = proton beam therapy.
*all patients HR 0.82 (0.56 to 1.20) p=0.3

Figure 10. Probability of PFS/DFS from Retrospective Cohort Studies Comparing Definitive Chemoradiotherapy using PBT versus IMRT for Curative Intent in Adults with Esophageal Cancer.



adj. = adjusted; CI = confidence interval; DFS = disease-free survival; IMRT = intensity modulated proton therapy; NS = not statistically significant; PBT = proton beam therapy; PFS = progression-free survival.

The other two comparative studies reported mortality (as opposed to OS) with fewer deaths in the PBT group compared with the photon groups, though none of the differences were statistically significant (Table 23).[164,173] The mortality rates were very different between the two studies, likely due to the difference in follow-up periods (median of 22 months and 3 months, respectively) and tumor types (squamous cell and adenocarcinoma, respectively).

Table 23. Morality from Retrospective Comparative Cohort Studies Comparing PBT versus Photons (IMRT, 3D-CRT or XRT) for Curative Intent in Adults with Esophageal Cancer

			Morality, % (n/N)		
Author, Year, N, Design, RoB	**Tumor type**	**Timing**	**PBT**	**Photon (various)**	**Effect size (95% CI) P-value***
Mortality					
Makishima (2015) N=44, Retrospective Comparative Cohort Moderately High RoB Definitive Chemoradiotherapy	SCC (100%)	NR (median f/u 22.3 mos.)	20% (5/25)	XRT: 31.6% (6/19)	RR 0.63 (0.23 to 1.77)†
Lin (2017), N=580, Retrospective Comparative Cohort Moderately High RoB Trimodal Therapy (Chemotherapy, Radiation and Surgery)	AC (92%) or SCC (8%)	1 mo. post-op	0% (0/111)	Any photon: 1.5% (7/469) • 3DCRT: 1.9% (4/214) • IMRT: 1.2% (3/255)	p=0.425
		2 mos. post-op	0.9% (1/111)	Any photon: 2.6% (12/469) • 3DCRT: 2.3% (5/214) • IMRT: 2.7% (7/255)	P=0.590
		3 mos. post-op‡	0.9% (1/111)	Any photon: 4.3% (20/469) • 3DCRT: 4.2% (9/214) • IMRT: 4.3% (11/255)	p=0.264‡

3DCRT = Three-dimensional conformal radiotherapy; AC: adenocarcinoma; CI = confidence interval; IMRT = intensity modulated proton therapy; NR = not reported; PBT = proton beam therapy; RoB = Risk of Bias; RR = risk ration; SCC = squamous cell carcinoma; XRT = X-ray chemoradiotherapy.
*p-value when reported is Chi-squared.
†Calculated by AAI. Authors did not report statistical significance.
‡According to authors, the differences at 3 months, though not statistically significant, the differences were clinically meaningful between PBT vs. photon groups.

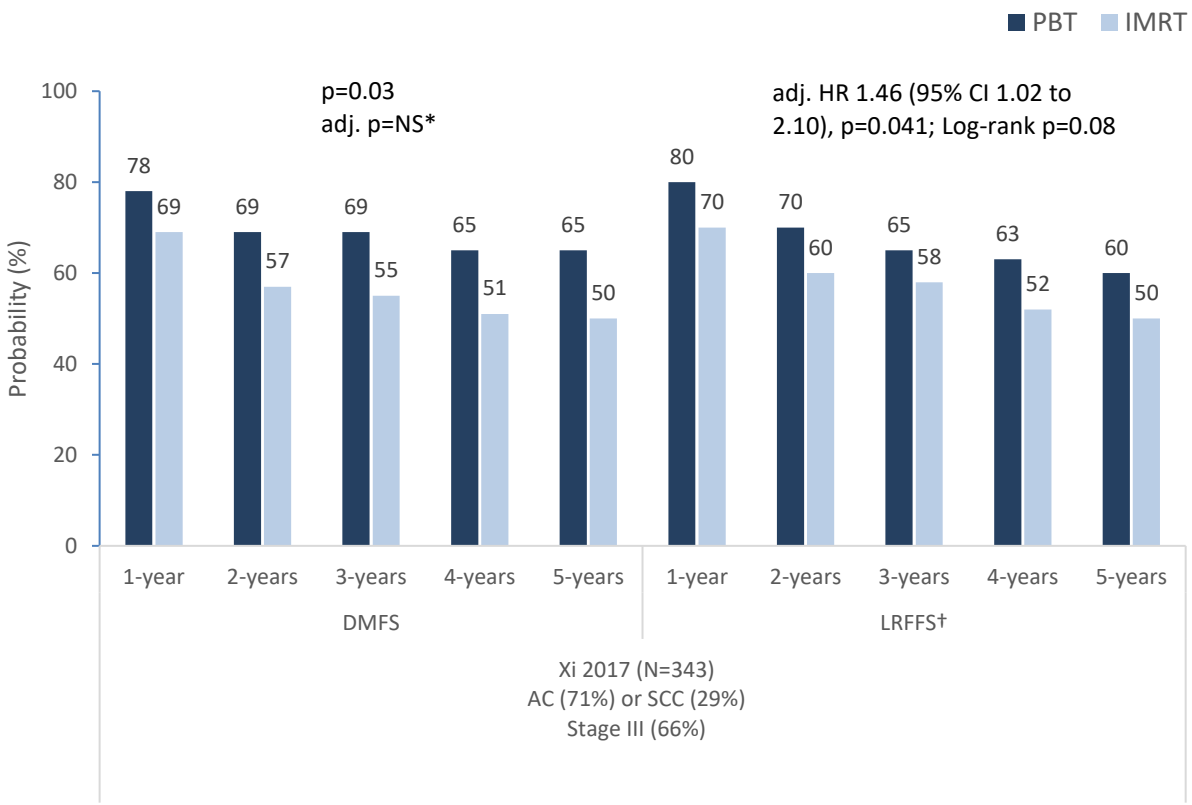
Across two small case-series (N= 40 and 47),[122,269] overall survival following PBT for esophageal cancer was 75.1% at 2 years (1 study)[122] and 59.2% and 70.4% at 3 years (2 studies). Treatment-related death was reported by both studies with rates of 0% and 4.3% (2/47); of note, all patients received concomitant chemotherapy. The probability of 3-year PFS was 56.3% in one study and of 2-year cause-specific survival was 77% in the other (Main Appendix F, Table F13).

Other outcomes

One comparative study[317] reported statistically greater probabilities of both distant metastasis-free survival (DMFS) and locoregional failure-free survival (LRFFS) after treatment with PBT, though only the latter remained statistically significant on multivariate analysis (adjusted HR 1.46, 95% CI 1.02 to 2.10), Figure 11. A subgroup analysis based on clinical stage found no statistical differences in 5-year DMFS or LRFFS for patients with stage I or II disease; probabilities were non-statistically higher for LRFFS (LRFFS: 63% vs. 43%, p=0.051) and similar for DMFS (data not report, p=0.191) in patients with stage III disease. This same study also reported the proportion of patients with locoregional recurrence who went on to receive salvage surgery with similar frequencies between groups (33% vs. 34%).

One case series (N=40) reported a distant progression rate of 7.5%.[122]

Figure 11. Probability of DMFS and LRFFS from Retrospective Cohort Studies Comparing Definitive Chemoradiotherapy using PBT versus IMRT for Curative Intent in Adults with Esophageal Cancer.



AC: adenocarcinoma; adj. = adjusted; CI = confidence interval; DMFS = Distant metastasis-free survival; IMRT = intensity modulated proton therapy; LRFFS = Locoregional failure-free survival; NS = not statistically significant; PBT = proton beam therapy; SCC = squamous cell carcinoma.
*DMFS not related to survival on mulivariate analysis.
†Fang et al. (not depicted in figure) reported only that treatment modality (IMRT or PBT) was not significantly associated with locoregional recurrence free survival

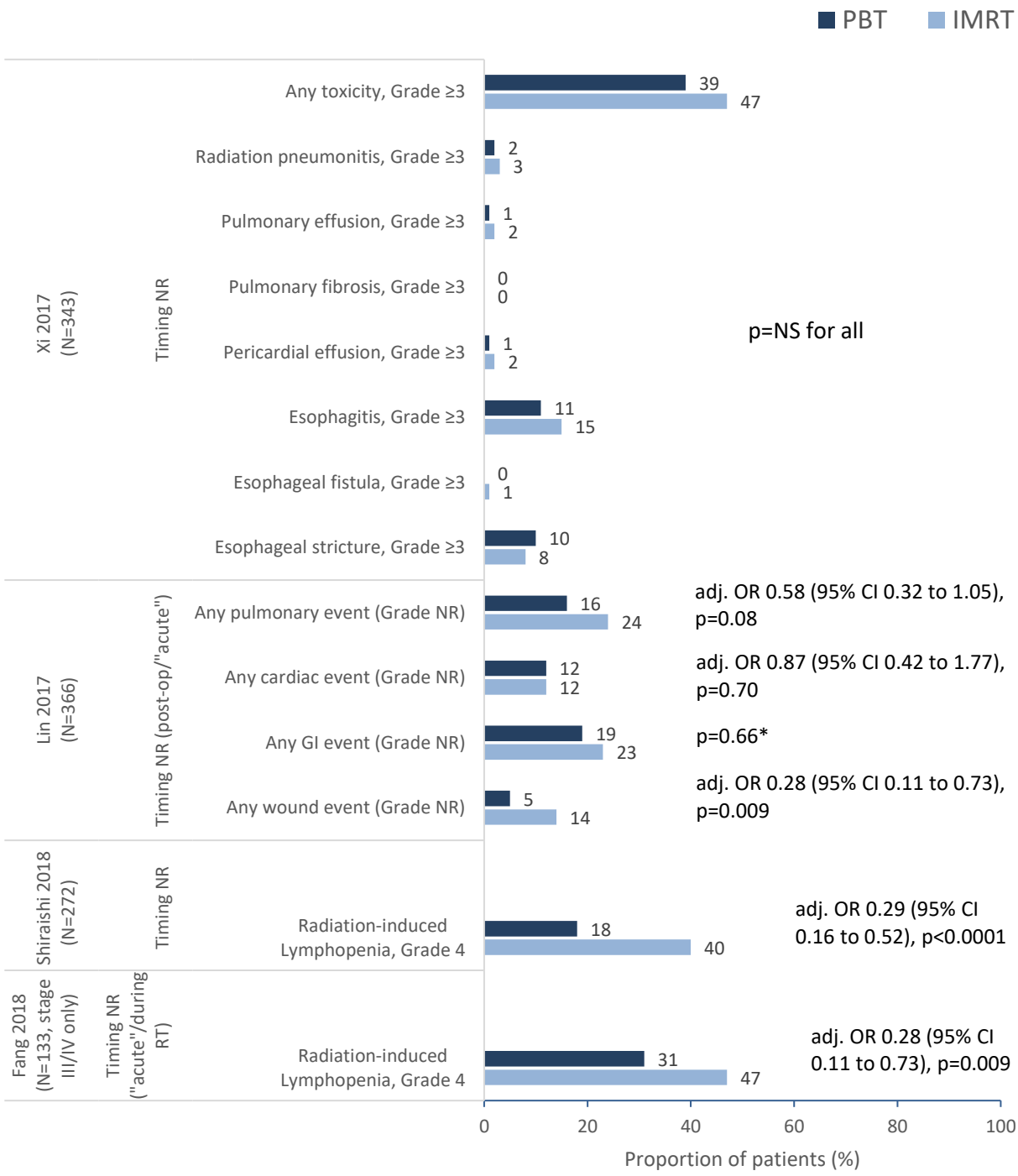
Key Question 3 (Safety)

Comparative studies

A total of five comparative cohort studies were identified that reported safety outcomes following PBT for curative intent. All four studies that provided data on effectiveness under Key Question 1 also report on safety; one additional retrospective case-matched comparative cohort (N=272)[255] was identified that reported only safety results, specifically radiation-induced lymphopenia following PBT versus IMRT. The patient population and treatment characteristics were similar to the other studies (see Table 22) as was the study quality (i.e., poor, moderately-high risk of bias).

In general, regardless of comparator, PBT resulted in fewer complications and reduced toxicity; however the differences between groups were not always statistically significant and clinical significance is unclear. Two studies that conducted propensity-score matched analyses (similar author groups from same treatment center) and compared PBT with IMRT reported a statistically significant reduction in radiation-induced grade 4 lymphopenia following PBT (adjusted ORs, 0.29 and 0.28),[77,255] Figure 12. Two studies, one that compared PBT with both IMRT and 3DCRT[164] and another that compared PBT with XRT,[173] found that PBT resulted in statistically fewer pulmonary and cardiac adverse events (grade NR) compared with 3DCRT (adjusted ORs 0.34 for both) and XRT ($p<0.001$ for both) but not with IMRT (Figures 12 and 13). No other statistical differences were seen between groups for any other safety measure across the studies, including grade ≥3 radiation pneumonitis (2 studies)[173,317] and radiation esophagitis (1 study).[317]

Figure 12. Safety Outcomes from Retrospective Cohort Studies Comparing PBT versus IMRT* for Curative Intent in Adults with Esophageal Cancer.

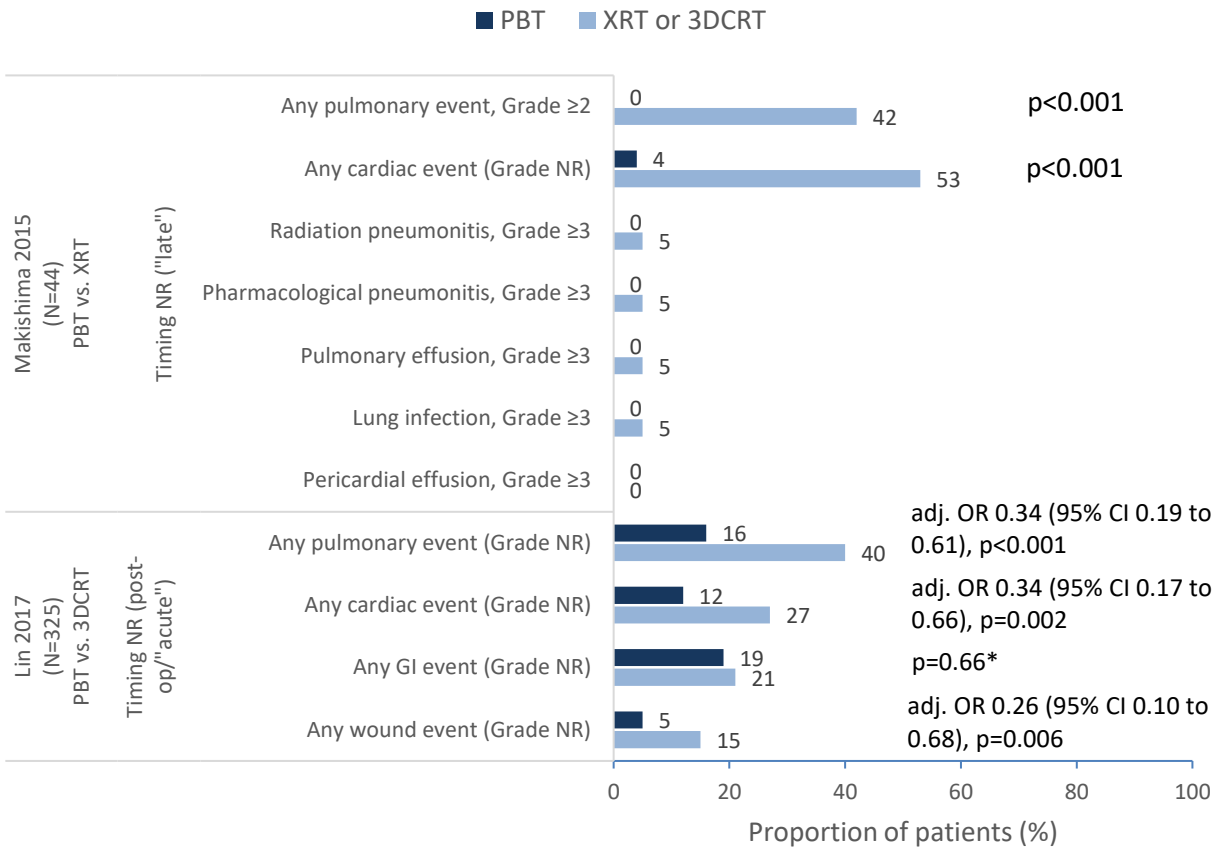


adj. = adjusted; CI = confidence interval; GI = gastrointestinal; NR = not reported; NS = not statistically significant; OR = odds ratio

* **Definitive Chemoradiotherapy**: Fang 2018, Xi 2017.
Trimodal Therapy (Chemotherapy, Radiation, Surgery): Shiriashi 2018, Lin 2017.

Final

Figure 13. Safety Outcomes from Retrospective Cohort Studies Comparing PBT versus 3D-CRT or XRT* for Curative Intent in Adults with Esophageal Cancer.



3D-CRT = Three-dimensional conformal radiotherapy; adj. = adjusted; CI = confidence interval; GI = gastrointestinal; NR = not reported; OR = odds ratio; PBT = proton beam therapy.

* **Definitive Chemoradiotherapy**: Makishima 2015.
Trimodal Therapy (Chemotherapy, Radiation, Surgery): Lin 2017.

Case series

In addition, two case series (N= 40 and 47)[122,269] were identified that reported toxicity following PBT (with concurrent chemotherapy) in adult patients with esophageal cancer (Main Appendix F, Table F15). In the acute period, the frequency of various hematological toxicities (grade 3 or 4) ranged from 25% to 55% and non-hematological toxicities (grade 3 or 4) ranged from 13% to 28%; for the latter, the most common event was esophagitis in both studies (11% and 22%). Regarding late toxicities, grade 3 events ranged from 5% to 9% (again primarily affecting the esophagus in both studies, 5% and 6%); no late grade 4 events were seen in either series. The authors of these studies do not distinguish between toxicities specifically attributed to radiation therapy/PBT versus concurrent treatments (e.g., chemotherapy); it is unclear to what degree PBT was associated with the above events.

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety) and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.5 Gastrointestinal Tumors

Key Points

- One small retrospective cohort study that compared PBT with hyper-fractionated accelerated radiotherapy (HART) for curative intent in adult patients with locally advanced and unresectable pancreatic adenocarcinoma reported no statistically significant differences between groups in the probability of 1- to 3-year OS, disease control/local progression or metastases or in the frequency of grade ≥3 radiation-related hematological or nonhematological toxicities which were rare; clinical importance of differences is unclear (Insufficient SOE).
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of included studies

One small (N=25),[168] retrospective, comparative cohort study that compared PBT with photon radiation therapy for curative intent in adult patients with locally advanced and unresectable pancreatic adenocarcinoma was identified (Table 24). Patients (48.2% male) were a median of 64 years of age with a mean Karnofsky performance status score of 86. The tumors were located in the head (84%) and the body/tail (16%) of the pancreas. Conventional three-dimensional (3-D) conformal PBT (median total dose 50 to 67.5 Gy) was compared with hyper-fractionated accelerated radiotherapy (HART) (median total dose 56 Gy). All patients received more than two cycles of induction chemotherapy prior to starting RT as well as adjuvant chemotherapy beginning one month after RT completion. For the reasons stated in Section 4.1, this study was considered to be of poor quality and did not control for possible confounding.

In addition, two case series of PBT for curative intent in adult patients with pancreatic cancer (adenocarcinoma) were identified.[112,142] All case series were considered to be at high risk of bias.

No studies evaluating PBT for salvage therapy were identified that met inclusion criteria.

Table 24. Pancreatic Cancer in Adults: Study Characteristics and Demographics for a Retrospective Cohort Comparing PBT versus HART for Curative Intent

	Maemura 2017	
Characteristics	**PBT (n=10)**	**Photon RT (n=15)**
Patient demographics		
Males, % (n)	50%	47%
Age, years; median (range)	64.5 (46 to 73)	64.2 (43 to 83)
Mean Karnofsky performance status ± SD	88 ± 4.2	85 ± 6.3
Tumor characteristics		
Subtype	Pancreatic Cancer (Adenocarcinoma)	
Location: Head/Body and Tail	80%/20%	87%/13%
Tumor size (cm, mean ± SD)	2.9 ± 0.88	3.3 ± 0.83
Unresectable factor: SMA or CA	80%	67%
CEA (ng/mL, mean ± SD)*	5.2 ± 3.8	4.8 ± 4.9
CA19-9 (U/mL, mean ± SD)*	279 ± 511.4	215 ± 291
Radiation Treatment		
Technique	3D Conformal Spot Scanning	Hyper-fractionated accelerated RT (HART)
Median total dose (Gy)	50 to 67.5	56
Concurrent Treatment		
Adjunctive chemotherapy	100%	100%
Study Design	**Retrospective Comparative Cohort**	
Follow-up, months (% followed)	**NR (% NR)**	
Risk of bias	**Moderately High**	

CA = celiac axis; CA19-9 = carbohydrate antigen 19-9; CEA = carcinoembryonic antigen; PBT = proton beam therapy; RT = radiation therapy; SD = standard deviation; SMA: superior mesenteric artery.
*Tumor markers.

Results

Key Question 1 (Effectiveness, curative intent)

Survival outcomes

In the retrospective cohort study,[168] the probability of OS did not differ statistically between the PBT and HART groups at any timepoint measured, respectively: 80% vs. 87% (1-year), 45% vs. 33% (2-years), and 23% vs. 27% (3-years). Sample size likely played a role in these findings.

Across two small case-series (N=37 and 48),[112,142] the probability of 1-year overall survival following PBT was 65% and 76%; at 2, 3 and 4 years probabilities were 42%, 23% and 23%, respectively, in one study.[112] The probability of 1-year progression-free survival was 45% and 65% across both studies and 24%, 18% and 10% at 2, 3, and 4 years, respectively, in one study.[112] (Main Appendix F, Table F16)

Other outcomes and secondary outcomes

Comparative studies

In the comparative cohort study,[168] with the exception of mean tumor reduction rate (1.6% for PBT vs. 29.9% for HART, p=0.02), no statistically significant differences were seen between treatment groups in any other outcome measured (Table 25). Sample size likely played a role in these findings.

Table 25. Other Primary and Secondary Outcomes from the Retrospective Cohort Study Comparing PBT versus HART for Curative Intent in Patients with Pancreatic Cancer

Outcome*		PBT (n=10)	HART (n=15)
Median overall survival	NA	23.4 months	22.3 months
Median time to progression	NA	15.4 months	15.4 months
Mean (±SD) tumor reduction rate, %*	NA	1.6% ± 35.7%	29.9% ± 22.1%
Other primary outcomes, % (n/N)			
Disease Failure†	Local Progression	40% (4/10)	60% (9/15)
	Metastasis	30% (3/10)	20% (3/15)
	Lung	10% (1/10)	0% (0/15)
	Liver	30% (3/10)	7% (1/15)
	Peritoneum	10% (1/10)	13% (2/15)
Secondary outcomes, % (n/N)			
CEA Response	>50% decrease	40% (4/10)	53% (8/15)
	<50% decrease	20% (2/10)	13% (2/15)
	Increase	20% (2/10)	33% (5/15)
CA19-9 Response	>50% decrease	50% (5/10)	27% (4/15)
	<50% decrease	40% (4/10)	60% (9/15)
	Increase	10% (1/10)	13% (2/15)

CA19-9: cancer antigen 19-9; CEA: carcinoembryonic antigen; CI: confidence interval; HART: Hyper-fractionated accelerated radiotherapy; NS: not statistically significant; NR: Not reported; OS: Overall Survival; SD: standard deviation; PBT: Proton Beam Therapy

*With the exception of mean tumor reduction rate (p=0.02), differences between groups were not statistically significant.

†Two patients in the PBT group exhibited simultaneous progression of local and metastatic lesions.

Case series

Locoregional failure was reported in 16.2% of resected patients (37 out of 48) in one case series[112] while 49% and 19% of patients showed local and regional progression, respectively, in the other series in which only two patients (5%) received subsequent surgery.[142] In both case series, 70% of patients developed distant metastasis over 2 to 3 years of follow-up (70% [26/37] and 73% [35/48]).

Key Question 3 (Safety)

Unless designated below as radiation-related, authors do not distinguish between toxicities specifically attributed to radiation therapy/PBT versus concurrent treatments (e.g., chemotherapy); it is unclear to what degree PBT was associated those events.

Comparative studies

In the small retrospective cohort study,[168] PBT generally resulted in reduced radiation-related toxicity compared with HART although none of the differences were statistically significant. Sample size may have played a factor in these findings. Other than one case of grade 3 ulcer, no other grade 3 non-hematological or hematological toxicities were seen following PBT (Table 26). No grade 4 events occurred in either group.

Table 26. Radiation-related Toxicity from the Retrospective Cohort Study Comparing PBT versus HART for Curative Intent in Patients with Pancreatic Cancer.

Outcome	Grade	PBT, % (n/N)	HART, % (n/N)	Effect size p-value
RT-related Hematological Toxicities*				
Leukopenia	Grade 2	10% (1/10)	13% (2/15)	NR
	Grade 3	0% (0/10)	20% (3/15)	NR
Neutropenia	Grade 2	0% (0/10)	0% (0/15)	NR
	Grade 3	0% (0/10)	0% (0/15)	NR
Anemia	Grade 2	0% (0/10)	0% (0/15)	NR
	Grade 3	0% (0/10)	0% (0/15)	NR
Thrombocytopenia	Grade 2	10% (1/10)	20% (3/15)	NR
	Grade 3	0% (0/10)	7% (1/15)	NR
RT-related Non-Hematological Toxicities*				
Malaise	Grade 2	0% (0/10)	0% (0/15)	NR
	Grade 3	0% (0/10)	0% (0/15)	NR
Nausea	Grade 2	0% (0/10)	7% (1/15)	NR
	Grade 3	0% (0/10)	0% (0/15)	NR
Anorexia	Grade 2	0% (0/10)	20% (3/15)	NR
	Grade 3	0% (0/10)	0% (0/15)	NR
Ulcer	Grade 2	10% (1/10)	0% (0/15)	NR
	Grade 3	10% (1/10)	0% (0/15)	NR

CI: confidence interval; HART: Hyper-fractionated accelerated radiotherapy; NR: not reported; RT: radiation therapy
*No grade 4 toxicities occurred in either group.

Case series

Two small case series (N=37 and 35)[112,142] reported grade 3 chemoradiation-related toxicity (hematological and non-hematological) in 0% and 6% of patients; in the latter study, there was one case each of colitis and chest wall pain.[112] No grade 4 or 5 toxicities were seen in either study and no late toxicities (e.g., gastrointestinal bleeding, duodenal ulcer) were reported in one study.[142] Appendix F.

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety) and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.6 Head and Neck (including Skull-base)

Key Points

- Across three retrospective cohort studies, the probabilities of 1- to 3-year OS and PFS (one case-matched study, primary oropharyngeal cancer), the incidence of all-cause mortality over a median 24 months (one small study, primary nasopharyngeal cancer), and 1-year OS (one small study, primary or metastatic salivary gland cancer) were not statistically different between PBT and IMRT groups. Clinical significance of differences is unknown. (Low SOE for primary oropharyngeal and nasopharyngeal cancer; Insufficient SOE for primary or metastatic salivary gland cancer).
- Across three retrospective comparative studies evaluating different tumor types (primary oropharyngeal; primary nasopharyngeal; and primary or metastatic salivary gland cancer), there were no statistically significant differences in the frequency of grade ≥3 acute or late toxicities or the incidence of ED visits/unplanned hospitalizations (1 study) following PBT versus IMRT (Low SOE based on largest, best quality study). A third retrospective comparative study in oropharyngeal cancer reported no statistical difference in the incidence of osteoradionecrosis after 6 months between PBT and IMRT (Insufficient SOE).
- Across five retrospective comparative cohorts evaluating different tumor types (2 primary oropharyngeal; 1 each of primary nasopharyngeal; primary nasopharyngeal or paranasal sinus; and primary or metastatic salivary gland cancer), gastrostomy tube dependence tended to be lower with PBT, however adjusted estimates from the largest study were not statistically significant, while smaller studies reported statistically significant differences. For the smallest study, the large confidence interval suggests instability of the effect estimate. Clinical significance of differences is unclear. It is unclear what role differences in study populations (including tumor characteristics, etc.) and possible residual confounding may play in these findings.
- One good quality cost-effectiveness analysis (QHES 90/100) took both societal and payer perspectives and concluded that, compared with IMRT, PBT was not cost-effective for patients with stage III-IV oropharyngeal squamous cell carcinoma using either perspective. However, at extremes of PBT superiority, it becomes cost-effective for younger human papilloma virus (HPV)-positive patients.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy (i.e., no comparative studies) or differential effectiveness and safety in this population.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of included studies

Eight retrospective comparative cohort studies that compared PBT with alternative therapies for curative intent in adult patients with head and neck cancers (to include the skull base) were identified[33,111,178,238,251,259,260,326] (Table 27 and

Table 28); two studies[178,326] reported on safety only. In addition, 23 case series evaluating a variety of cancer types involving the head and neck were identified, 18 evaluating PBT for curative intent[59,64,67,78,81,91,93,177,193,199,267,272,281,301,308,309,324,325] and 5 for salvage therapy.[103,104,179,224,239] Some case series included populations undergoing PBT for both curative and salvage intent; these studies were categorized under Key Question 1 or Key Question 2 based on what the majority of patients received. Consistent with the prior report, studies where the majority of the population had skull-based or cervically-located chordomas and chondrosarcomas (one comparative study[259] and eight case series [64,67,78,81,177,267,308,309]) are reported here along with non-skull-base head and neck malignancies, although the results are reported separately. For the reasons described previously (Section 4.1), all comparative cohort studies are considered moderately high risk of bias; however, four of the eight studies included here[33,111,178,251] did control for confounding. All case series are high risk of bias.

In addition, one cost-effectiveness analysis (CEA)[253] which compared PBT with IMRT (both accompanied by chemotherapy) in 65 year-old patients with stage III-IV oropharyngeal squamous cell carcinoma that met inclusion criteria was identified.

Table 27. Study and patient characteristics from comparative studies reporting on effectiveness and safety: Head and Neck (including Skull-base) cancers

Effectiveness and Safety								
	Blanchard 2016		Romesser 2016		Holliday 2015		Simon 2018	
	PBT (n=50)	Photon RT (n=100)	PBT (n=18)	Photon RT (n=23)	PBT (n=10)	Photon RT (n=20)	Surgery + PBT (n=23)§	Surgery alone (n=24)
Patient Characteristics								
Males, % (n)	54%	33%	---	---	70%	70%	57%	41%
Age, years; median (range)	61 (37 to 84)	55.5 (34 to 78)	60.4	60.9	45 (IQR, 18 to 55	54 (IQR, 39 to 59)	Mean: 42 (12 to 69)	Mean: 52 (10 to 85)
Smoking†	42%	38%	---	---	---	---	---	---
Diplopia	---	---	---	---	---	---	57%	29%
Headache	---	---	---	---	---	---	35%	17%
Nasal Obstruction	---	---	---	---	---	---	4%	29%
Tumor Characteristics								
Subtype	Oropharyngeal Cancer		Major salivary gland cancer or cutaneous SCC metastasis to salivary gland‡		Nasopharyngeal Cancer		Skull base Chondrosarcoma	
Tumor location								
Anterior skull-base	---	---	---	---	---	---	4%*	50%*
Petroclival	---	---	---	---	---	---	96%*	50%*
Base of tongue	46%	46%	---	---	---	---	---	---
Tonsil	54%	54%	---	---	---	---	---	---
Parotid gland	---	---	78%	91%	---	---	---	---
Submandibular Gland	---	---	22%	10%	---	---	---	---
Stage								
I	2%		---	---	0%	10%	---	---
II	0%		---	---	---	---	97.9%	
III	18%		---	---	---	---	---	---

Effectiveness and Safety								
	Blanchard 2016		Romesser 2016		Holliday 2015		Simon 2018	
	PBT (n=50)	Photon RT (n=100)	PBT (n=18)	Photon RT (n=23)	PBT (n=10)	Photon RT (n=20)	Surgery + PBT (n=23)§	Surgery alone (n=24)
II/III	---		---	---	90%	75%	---	---
IVA	74%		---	---	---	---	---	---
IVB	6%		---	---	---	---	---	---
Unknown	---		---	---	10%	15%	---	---
Radiation Treatment Characteristics								
Technique	Intensity Modulated spot scanning	IMRT	Uniform Scanning- beam	IMRT	IMPT	IMRT	---	---
Median total dose (Gy)	small volume disease: 66 advanced disease: 70 elective regions: 54 to 63	small volume disease: 66 advanced disease: 70 elective regions: 54 to 63	66	66	70	70	70	N/A
Additional Treatments								
Prior to Radiation								
Induction Chemo	40%	44%	---	---	80%	75%	---	---
Neck Dissection	6%	11%	---	---	---	---	---	---
Neck Nodal Irradiation	---	---	50%	26%	---	---	---	---
Gross Total Resection	---	---	---	---	---	---	13%*	54%*
Partial Resection	---	---	---	---	---	---	87%*	46%*
Concurrent/Adjuvant Treatment								
Concurrent Chemo	64%	64%	22%	30%	100%	90%	---	---
Adjuvant Chemo	---	---	---	---	10%	0%	---	---
Neck Dissection	12%	15%	---	---	---	---	---	---
Study Design	**Retrospective Matched Pairs Comparative Cohort**		**Retrospective Comparative Cohort**		**Retrospective Matched Pairs Comparative Cohort**		**Retrospective Comparative Cohort**	

Effectiveness and Safety								
	Blanchard 2016		Romesser 2016		Holliday 2015		Simon 2018	
	PBT (n=50)	Photon RT (n=100)	PBT (n=18)	Photon RT (n=23)	PBT (n=10)	Photon RT (n=20)	Surgery + PBT (n=23)§	Surgery alone (n=24)
F/U, months (% followed)	**29 (NR)**	**33 (NR)**	**16.1 (NR)***	**4.7 (NR)***	**21.6 (NR)**	**25.8 (NR)**	**91 months (95.7%)**	
Risk of bias	**Moderately High**		**Moderately High**		**Moderately High**		**Moderately High**	

Chemo = chemotherapy; F/U = follow-up; Gy = Gray; IMRT = intensity modulated radiation therapy; PBT = proton beam therapy; SCC = squamous cell carcinoma; SD = standard deviation

*Indicates a statistically significant difference between groups

†Defined as having smoked >10 packs per year

‡In the PBT and Photon groups, respectively, 44% vs. 57% had perineural involvement and 11% vs. 9% had unresectable disease.

§ 4 patients received combined photon/proton therapy

Table 28. Study and patient characteristics from comparative studies reporting on effectiveness only or safety only: Head and Neck (including Skull-base) cancers

	Effectiveness only				Safety Only			
	Sharma 2018		Sio 2016		Zhang 2017		McDonald 2016	
	PBT (n=31)	IMRT (n=33)	PBT (n=35)	Photon RT (n=46)	PBT (n=50)	Photon RT (n=534)	PBT (n=14)	Photon RT (n=26)
Patient Characteristics								
Males, % (n)	87%	82%	86%	91%	86.5%	84%	78.6%	53.8%
Age, years; median (range)	Mean: 60	Mean: 58	Mean ± SD: 59.1 ± 10.2	Mean ± SD: 58.2 ± 9.9	≤ 60: 56.4% >60: 43.6%	≤ 60: 44% >60: 56%	46.7 (16 to 71	54.1 (22 to 77)
Comorbidities								
HPV positive	---	---	74%*	13%*	---	---	---	---
Tumor Characteristics								
Subtype								
Oropharyngeal SCC	100%		100%		100%		---	---
SCC (location not specified)	---		---		---		21.4%	50%
Poorly differentiated carcinoma	---		---		---		0%	19.2%
Sinonasal undifferentiated	---		---		---		35.7%	15.4%
Esthesioneuroblastoma	---		---		---		35.7%	3.8%
Other**	---		---		---		7.1%	18.5%
Tumor location								
Nasopharynx	---	---	---	---	---	---	14.3%*	57.7%*
Nasal/paranasal	---	---	---	---	---	---	85.7%	42.3%
Base of tongue	35%	39%	57%	50%	48.7%	42%	---	---
Tonsil	65%	61%	31%	50%	51.3%§	58%§	---	---
Parotid gland	---	---	0%	0%	0%	0%	---	---
Submandibular Gland	---	---	0%	0%	0%	0%	---	---
Other	---	---	11%	0%	0%	0%	---	---

	Effectiveness only				Safety Only			
	Sharma 2018		Sio 2016		Zhang 2017		McDonald 2016	
	PBT (n=31)	IMRT (n=33)	PBT (n=35)	Photon RT (n=46)	PBT (n=50)	Photon RT (n=534)	PBT (n=14)	Photon RT (n=26)
Stage								
I	---	---	3%	2%	---	---	---	---
II	---	---	3%	4%	---	---	---	---
III	---	---	26%	15%	---	---	---	---
I – III	13%	15%	---	---	---	---	---	---
IVA	87%	85%	69%	78%	---	---	---	---
IVB	---	---	0%	0%	---	---	---	---
Radiation Treatment Characteristics								
Technique	PBS	Volumetric Modulated Arc Therapy	Intensity Modulated Scanning-beam	IMRT	Intensity Modulated Scanning- beam	IMRT	3D Conformal	IMRT [14 patients also had concurrent PBT]
Median total dose (Gy)	61.7	62.6	70	70	Total: 66-70† Mean mandibular dose: 25.6*	Total: 66-70† Mean mandibular dose: 41.2*	71.4 (63 to 75.6)	71.8 (66 to 76.4)
Additional Treatments								
Prior to Radiation								
Induction Chemo	59%	62%	74.3%	23.9%	40.6%	40%	21.4%	0%
Neck Dissection	---	---	---	---	---	---	7.1%	0%
Neck Nodal Irradiation	---	---	---	---	---	---	---	---
Concurrent/Adjuvant Treatment								
Chemotherapy	---	---	100%	100%	67.4%	64%	64.2%	88.5%
Neck Dissection	---	---	---	---	---	---	7.1%	3.8%
Gastrostomy Tube	---	---	---	---	---	---	14.3%	84.6%
Study Design	**Prospective Comparative Cohort**		**Retrospective Comparative Cohort**		**Retrospective Comparative Cohort**		**Retrospective Matched Pairs Comparative Cohort**	
F/U, months (% followed)	**NR (NR)**	**NR (NR)**	**NR (NR)**	**NR (NR)**	**34.6 (NR)**	**33.8 (NR)**	**21.6 (NR)**	**25.8 (NR)**

	Effectiveness only				Safety Only			
	Sharma 2018		Sio 2016		Zhang 2017		McDonald 2016	
	PBT (n=31)	IMRT (n=33)	PBT (n=35)	Photon RT (n=46)	PBT (n=50)	Photon RT (n=534)	PBT (n=14)	Photon RT (n=26)
Risk of bias	**Moderately High**		**Moderately High**		**Moderately High**		**Moderately High**	

Chemo = chemotherapy; F/U = follow-up; Gy = Gray; IMRT = intensity modulated radiation therapy; PBT = proton beam therapy; SCC = squamous cell carcinoma; SD = standard deviation

*Indicates a statistically significant difference between groups

†For patients receiving concurrent chemoradiation the prescribed dose to the tumor was 70 Gy; for patients who received only radiotherapy, the prescribed dose was 66

§Proportion of patients with tumor located in the tonsil or another location

**To include Lymphoepithelioma, High grade mucoepidermoid, and neuroendocrine carcinoma

Results

4.3.6.1 Head and neck cancers of the paranasal sinuses, nasal cavity, oral cavity, tongue, salivary glands, larynx, or pharynx

Key Question 1 (Effectiveness, curative intent)

Comparative studies of non-skull-base head and neck cancers

Across the six comparative cohort studies that provided data on effectiveness (Table 27 and

Table 28),[33,111,178,239,260] five included patients with primary non-skull-base head and neck cancers: oropharyngeal cancer (3 studies)[33,251,260],) nasopharyngeal cancer (1 study)[111], and nasopharyngeal or nasal/paranasal sinus cancers (1 study)[178]. The sixth study included patients being treated for either primary or metastatic salivary gland cancer effecting the parotid gland (primarily) and the submandibular gland.[238]

In the three studies evaluating primary oropharyngeal cancer (N = 64 and 150),[33,251,260] patient age was similar (57 to 60 years) but the proportion of males varied between studies (40% to 89%), Table 27 and

Table 28. In two studies, the tissue type was squamous cell carcinoma (SCC)[251,260](not reported by the third study). Comorbidities were poorly reported. The majority of patients had stage III to IVA disease. Two studies compared intensity modulated PBT versus IMRT with similar total radiation doses (median 70 Gy).[33,260] Induction chemotherapy was common in both studies; however, in one study, significantly more patients received induction chemotherapy in the PBT group (74% vs. 24%).[260] In one study, 15% of patients underwent neck dissection either prior-to or following RT.[33] The third study compared adjuvant pencil beam scanning PBT with IMRT via volumetric modulated arc therapy (VMAT) following tranoral robotic surgery and selective neck dissection; median total radiation doses were 61.7 Gy(RBE) and 62.6 Gy, respectively.[251] All patients in one study and the majority in the other two studies (61% and 64%) were receiving concurrent and/or adjuvant chemotherapy.

Two small studies (N=30 and 40) evaluated patients with primary nasopharyngeal (n=30)[111] or a mix of nasopharyngeal and nasal/paranasal cancers (N=40)[178]. The study that included only nasopharyngeal cancer case-matched patients who had received intensity modulated PBT versus IMRT (median RT dose 70 Gy(RBE)/Gy in both groups).[111] Median patient age was 49 years and the majority were male (70%) with primarily grade II/III (80%) disease; most patients underwent induction (77%) and concurrent (90%) chemotherapy. Comorbidities were poorly reported. The second study included patients (median age 52 years) with a mix of nasal/paranasal (58%) and nasopharyngeal (42%) cancers treated with PBT versus IMRT[178]; of note, 14 of the 26 patients in the latter group also had concurrent PBT. There were several differences between the two groups at baseline including primary tumor site, tumor histology (e.g., SCC), T and N stage, smoking history, and sequencing of chemotherapy (i.e., induction, concurrent and/or adjuvant) (Table 27 and

Table 28). The median primary tumor dose was similar between groups (PBT 71.4 Gy (RBE) vs. IMRT 71.8 Gy); however the median neck dose to a positive node was greater in the PBT group, 72.9 Gy(RBE) versus 68.3 Gy.

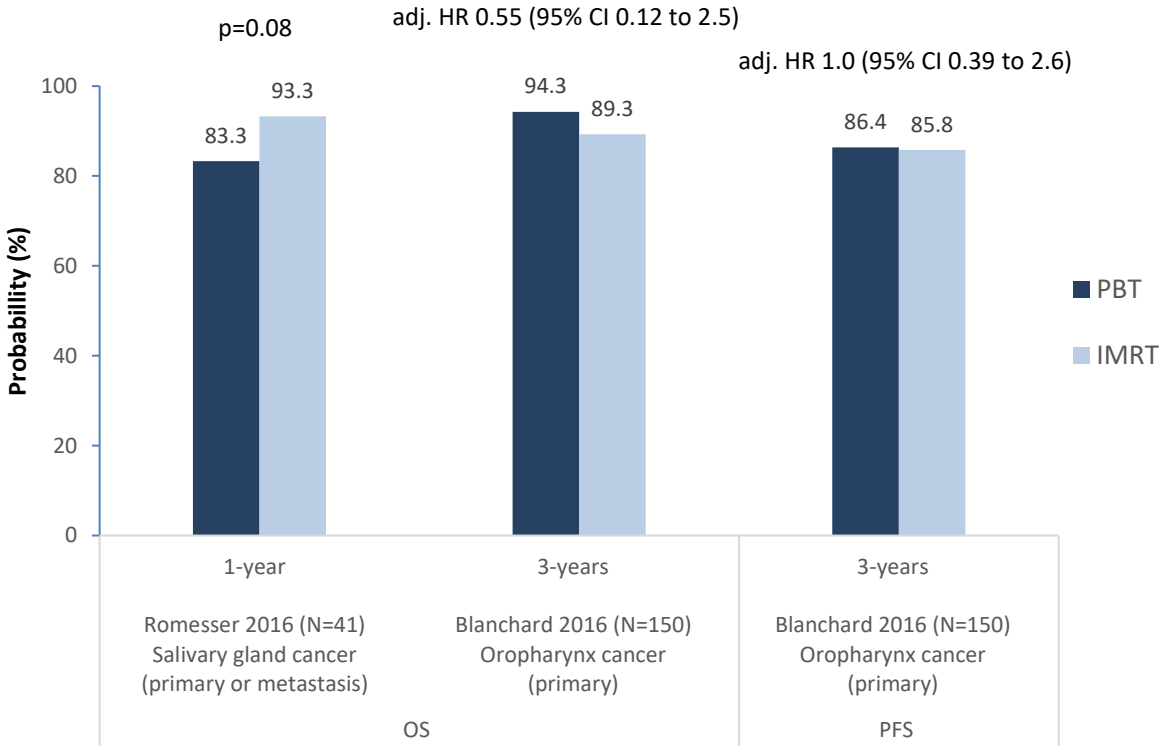
The sixth study in either primary or metastatic salivary gland cancer was small (N=41)[238] and included patients with median age of 61 years; sex distribution and tumor stage were not reported (Table 27). Patients received a median radiation dose of 66 Gy delivered via uniform scanning-beam PBT or IMRT. Twice as many patients in the PBT group had undergone neck nodal irradiation prior to study entry (50% vs. 26%) and only 26% of all patients received concurrent chemotherapy.

Survival outcomes

Comparative studies of non-skull-base head and neck cancers

Two comparative cohort studies reported survival outcomes (Figure 14). One study[33] evaluating patients with primary oropharynx cancer reported no statistically significant differences between RT groups in the probability of 3-year OS or PFS while the second study,[238] in primary or metastatic salivary gland cancer, reported a lower probability of 1-year OS following PBT versus IMRT (83% vs. 93%), though the difference did not reach statistical significance (p=0.08), possibly due to the small sample size. A third, small matched-pairs cohort evaluating patients with grade II or III nasopharyngeal cancer reported one case each of all-cause mortality in the intensity-modulated PBT (10%) and the IMRT (5%) groups.[111]

Figure 14. Probability of OS and PFS from Retrospective Cohort Studies Comparing PBT versus IMRT for Curative Intent in Adults with Head and Neck Cancers.



adj. = adjusted; CI = confidence interval; HR = hazard ratio; IMRT = intensity modulated radiation therapy; OS = overall survival; PBT = proton beam therapy; PFS = progression-free survival.

Case series of non-skull-base head and neck cancers

Seven case series reported OS, all in populations receiving PBT for curative intent.[59,93,193,199,272,324,325] .

The range of probabilities for OS across various timeframes and types of cancer [including sinonasal, nasal, and paranasal (including one study of mucosal melanoma[324]), oropharyngeal, and tongue cancer] are as follows:

- 1-year: 88% to 95% (3 studies of sinonasal, nasal or paranasal cancer, N=32 to 112)[59,324,325]
- 2-year: 60% to 94.5% (4 studies, N=32 to 112)[59,93,324,325]; excluding the study in mucosal melanoma: 80% to 94.5%
- 3-year: 46% to 95% (6 studies, N=32 to 112)[59,93,193,272,324,325]; excluding the study in mucosal melanoma: 61% to 95%
- 5-year: 64% to 81% (2 studies, N=42 to 112)[199,308,325]

Studies in oropharyngeal cancer showed higher probabilities of OS following PBT when compared with studies evaluating sinonasal, nasal, or paranasal cancers; OS was poor following treatment for mucosal melanoma in one study.[324] The one study in 33 patients with stage III-IV tongue cancer reported a 3-year OS of 87%.[272]

Six case series reported PFS,[93,193,199,272,324,325] all in populations treated with curative intent. The range of probabilities of PFS across various time frames and cancer types is as follows:

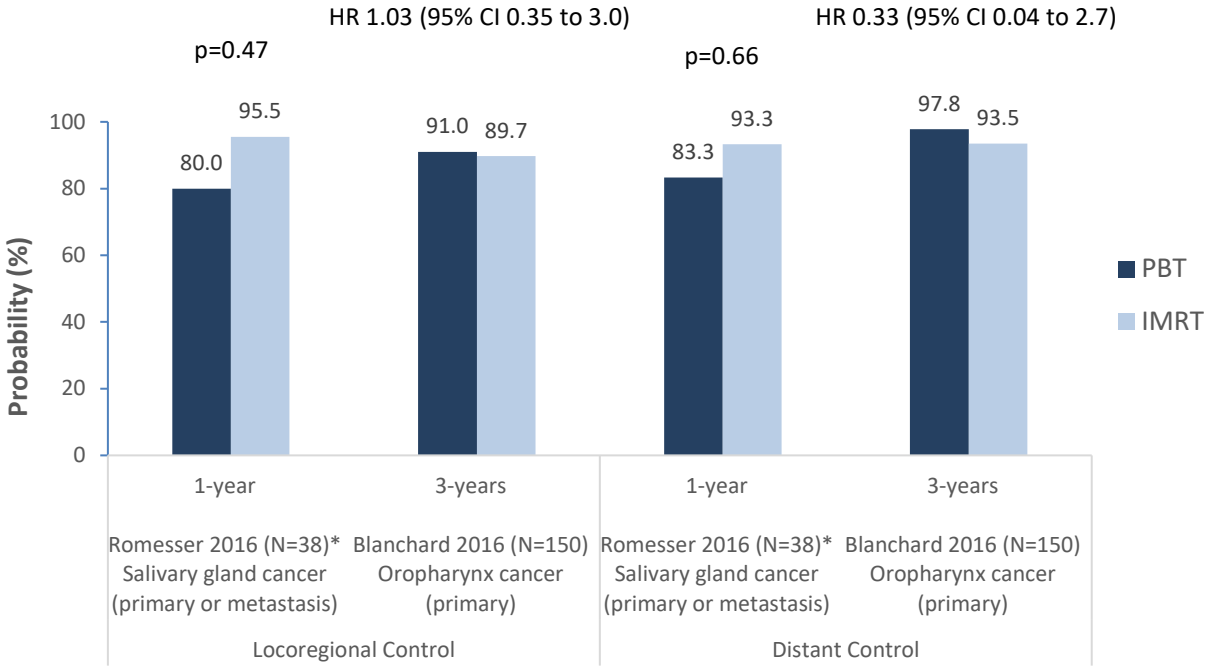
- 2-year: 36% (1 study of mucosal melanoma of the nasal sinuses)[324] and 89% (1 study of oropharyngeal cancer)[93]
- 3-year: 36% to 56% (in 3 studies of patients with sinonasal and paranasal cancers primarily, including mucosal melanoma [36%])[193,324,325]; 74% to 89% (in 2 studies of tongue cancer and oropharyngeal cancer, respectively)[93,272];
- 5-year: 45% and 49% [2 studies of sinonasal, nasal, or paranasal cancer (to include olfactory neuroblastoma) (N=42 and 112)][199,325]

Other outcomes

Comparative studies of non-skull-base head and neck cancers

The three comparative studies that provided data for primary outcomes, also reported data on tumor control. The probability of locoregional control and distant control were reported by two of the studies, one in primary oropharyngeal cancer and one in primary or metastatic salivary gland cancer, with no significant differences between groups in either outcomes in both studies[33,238] (Figure 15). The latter showed a tendency for less locoregional (80% vs. 96%) and distant control (83% vs. 93%) in the PBT group compared with the IMRT group and the small sample size may have contributed to the non-statistically significant findings.[238] The third small study in patients with primary nasopharyngeal cancer reported the proportion of patients who experienced disease failure; for intensity modulated PBT versus IMRT, respectively, local failure was observed in no patient versus one patient (5%) and distant metastases were seen in one patient in each group, 10% vs. 5%.[111]

Figure 15. Probability of Local and Distant Control from Retrospective Cohort Studies Comparing PBT versus IMRT for Curative Intent in Adults with Non-Skull-base Head and Neck Cancers.



CI = confidence interval; HR = hazard ratio; IMRT = intensity modulated radiation therapy; OS = overall survival; PBT = proton beam therapy; PFS = progression-free survival.
*Excludes 1 patient from PBT group and 2 patients from IMRT group who had distant metastases prior to RT

Two comparative cohorts reported data related to patient quality of life (QOL)[251,260]; see Abstraction Appendix G for details related to these outcomes. One study[260] reported no differences between PBT and IMRT in either the short- or longer-term for both change in mean scores and the proportion of patients with moderate to severe symptoms according to the MD Anderson Symptom Inventory-Head and Neck (MDASI-HN) survey. The second studythat compared adjuvant pencil beam scanning PBT versus IMRT via VMAT in patients treated with transoral robotic surgery for primary oropharyngeal squamous cell cancer[251] reported QOL using three different validated instruments. Over the longer term (6 and 12 months), compared with IMRT, PBT patients showed significant improvement in general and severe xerostomia, reported a significantly lower frequency of postoperative dental problems, and showed significantly better results in terms of role function. In all these instances, the differences were considered to be clinically meaningful (i.e., 10-point difference between the two groups). .

Case series of non-skull-base head and neck cancers

The 1-year probability of local control was reported by two studies of sinonasal cancer,[59,324] both in populations undergoing PBT for curative intent. The probability ranged from 76% in the study of mucosal melanoma[324] to 92% in the study that specifically excluded melanoma.

One study treated patients (N=84) with sinonasal cancers and reported distant metastasis free survival probabilities of 88%, 82%, and 73% at 1, 2, and 3 years, respectively.[59]

Key Question 2 (Effectiveness, salvage therapy)

No comparative studies that met inclusion criteria were identified that evaluated PBT for salvage treatment (including treatment for recurrent disease) of non-skull-base tumors. A total of four case series were included, [103,104,224,239] two of which enrolled patents with oral cancers (primarily tongue and gingiva) and had substantial overlap in patients populations[103,104]; the remaining two studies included patients with a variety of head and neck cancers, primarily squamous cell oropharyngeal, sinonasal and nasopharyngeal carcinoma (Abstraction Appendix G).[224,239]

Across the two small studies (N=34 and 46) evaluating PBT for recurrent oral cancer,[103,104] the probability of 1- and 2-year OS, respectively, ranged from 62% to 65% and from 42% to 46%; for local control, probabilities ranged from 77% to 81% and from 60% to 70%.

Across the two studies (N = 60 and 92) evaluating PBT for recurrent oropharyngeal, sinonasal and nasopharyngeal cancers (primarily),[224,239] the probabilities of 1-year OS ranged from 65% to 81%; 2-year OS was 69% as reported by one series.[224] PFS was reported by one study with 1- and 2-year probabilities of 60% and 48%, respectively.[224] The probability of distant metastasis-free survival across two studies was 75% to 84% at 1 year and 64% to 66% at 2 years.[224,239]

Key Question 3 (Safety)

Comparative studies of non-skull-base head and neck cancers

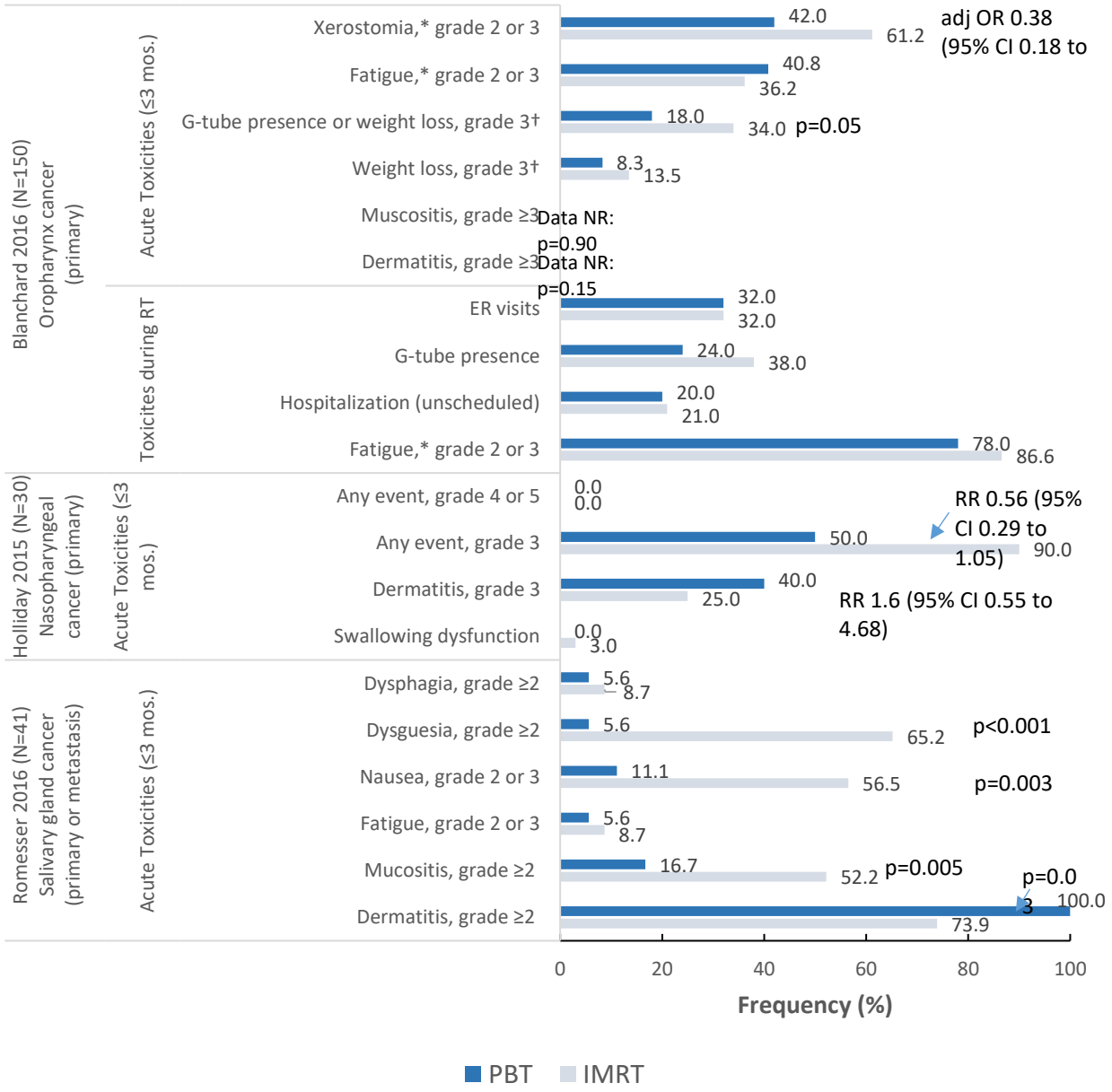
A total of six comparative cohort studies were identified that reported safety outcomes. Four of the six studies that provided data on effectiveness in also report on safety[33,111,238,251]; two additional studies were identified that reported only safety results. .. One study (N=584)[326] evaluated the incidence of radiation-induced osteonecrosis following intensity modulated PBT versus IMRT in patients with primary oropharyngeal cancer. The second, small study (N=40)[178] reported the frequency of gastrostomy tube dependence and opioid pain requirement following 3D conformal PBT versus IMRT for treatment of nasopharynx and paranasal sinus cancers; unlike the other studies, 54% of patients in the IMRT group also received concurrent PBT. Patient demographics and treatment characteristics (other than the aforementioned) of these two studies were similar to the other studies, as was study quality (i.e., poor quality, moderately high risk of bias) (see Table 27 and

Table 28); however, one of the studies did control for confounding[178].

Across all studies and tumor types, PBT generally resulted in fewer complications and reduced toxicity over both the acute and late term compared with IMRT and grade 4 or 5 events were rare; however the differences between groups were not always statistically significant and clinical significance is unclear.

Acute toxicity and adverse events were reported by four studies[33,111,178,238], all evaluating different types of head and neck cancers. On multivariate analysis, one study in patients with primary oropharyngeal cancer[33] reported no statistical differences in grade 2 or 3 acute events, with the exception of acute grade ≥2 xerostomia which was significantly reduced in the PBT group (42% vs. 61%; adjusted OR 0.38, 95% CI 0.18 to 0.79), Figure 16., A second small study in patients with primary nasopharyngeal cancer found that PBT reduced the frequency of any grade 3 event by almost half compared with IMRT (50% vs. 90%). The authors report a statistically significant difference between group (p=0.015), however according to our calculations (crude RR 0.56, 95% CI 0.29 to 1.05) the difference fails to reach statistical significance, likely due to the small sample size and residual confounding; it is unclear whether or not the p-value reported by the authors represents an adjusted estimate. There were no statistical differences between groups in the frequency of grade 3 dermatitis or swallowing dysfunction specifically in this same study (Figure 16). Another study, in patients with primary or metastatic salivary gland tumors,[238] found no significant differences between PBT and IMRT in the frequency of any grade ≥3 acute event (when grade 2 events were included, PBT resulted in statistically fewer cases of grade ≥2 mucositis, dysgeusia, and nausea but statistically more cases of grade ≥2 dermatitis compared with IMRT), Figure 16. The fourth study, in patients with nasopharyngeal or paranasal sinus cancer, reported that patients who received PBT significantly reduced their opioid medication use (compared with baseline) versus those who received IMRT (adj. OR 0.09; 95% CI 0.01 to 0.57).[178]

Figure 16. Acute Toxicity and Adverse Events from Retrospective Cohort Studies Comparing PBT versus IMRT for Curative Intent in Adults with Various Non-Skull base Head and Neck Cancers.



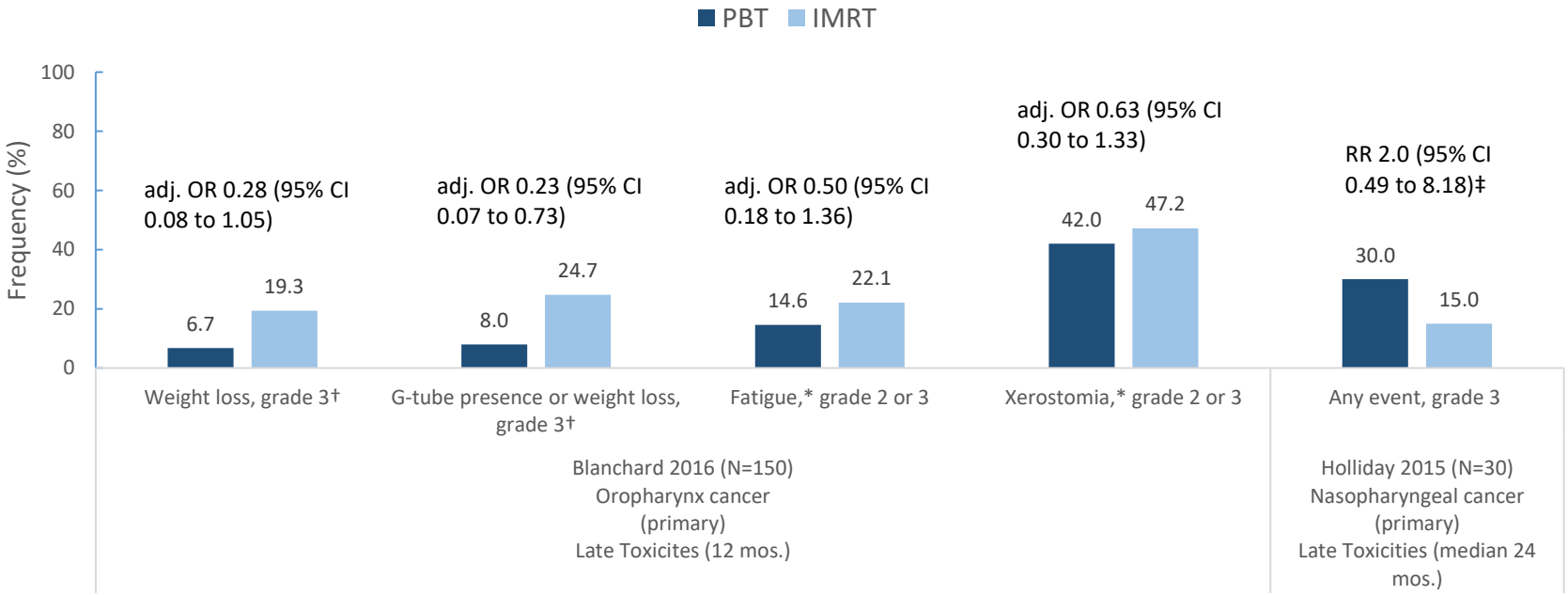
adj. = adjusted; CI = confidence interval; IMRT = intensity modulated radiation therapy; NR = not reported; OR = odds ratio; PBT = proton beam therapy; RT = radiation therapy.
*patient-reported
†>20% weight loss compared to baseline

Late toxicity and adverse events were reported by only two studies (N=150 and 30), one in primary oropharyngeal (N=150)[33] and one in primary nasopharyngeal (N=30)[111] cancer. Only the composite outcome of gastrostomy tube presence or grade 3 weight loss (>20% weight loss compared with baseline) differed statistically between groups and favored PBT at 12 months in the study of

oropharyngeal cancer (of note, neither outcome on its own was statistically significant, though both tended to be lower with PBT); no other statistical differences were note in either study (

Figure 17).

Figure 17. Late Toxicity and Adverse Events from Retrospective Cohort Studies Comparing PBT versus IMRT for Curative Intent in Adults with Primary Oropharyngeal or Nasopharyngeal Cancer.



adj. = adjusted; CI = confidence interval; IMRT = intensity modulated radiation therapy; NR = not reported; OR = odds ratio; PBT = proton beam therapy.

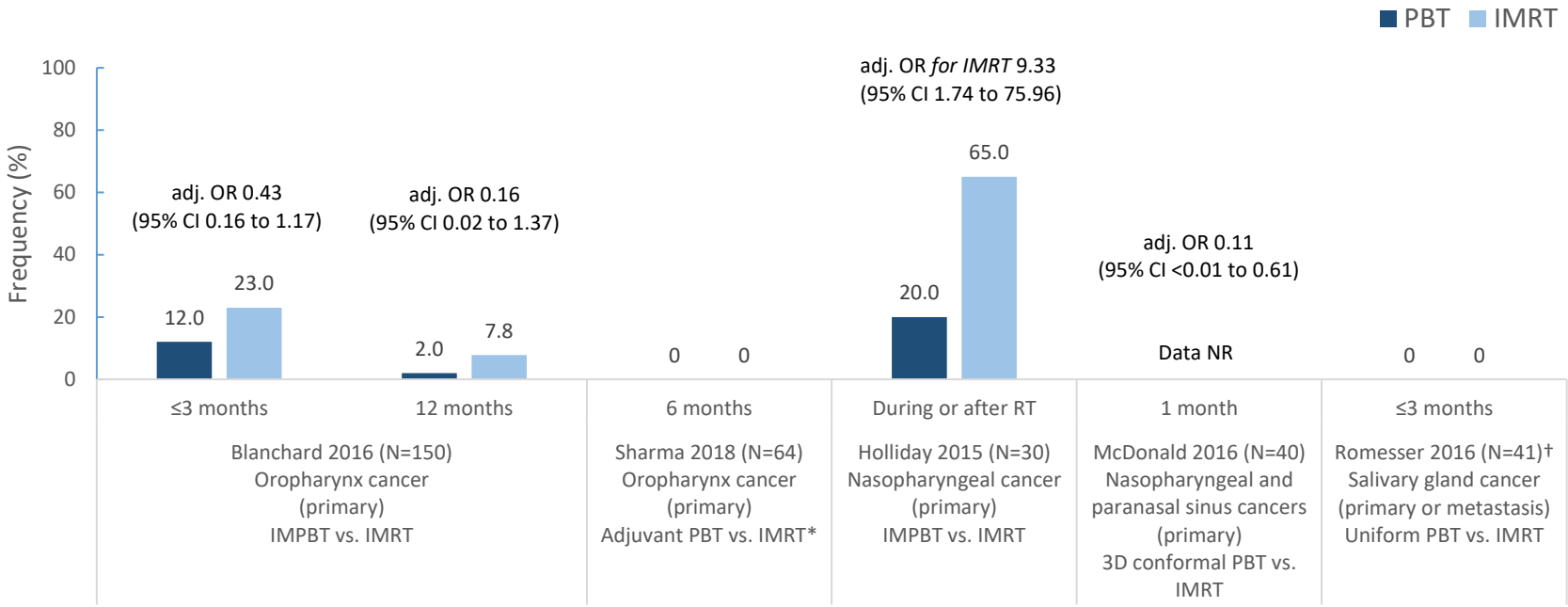
*patient-reported

†Grade 3 weight loss is >20% weight loss compared to baseline. The estimates for G-tube presence only at 12 months in this study can be found in Figure 18 below.

‡RR and 95% CI calculated by AAI.

Gastrostomy tube (G-tube) dependence was reported by five studies (Figure 18).[33,111,178,238,251] Two studies included patients with primary oropharyngeal cancer treated with PBT versus IMRT (as an adjuvant treatment to surgery in one study)[33,251] both of which reported no statistically significant differences between treatment groups in the proportion of patients dependent on a G-tube at up to 12 months of follow-up, though PBT tended to result in less dependence on multivariate analysis in one study.[33] PBT, as compared with IMRT, resulted in statistically significant reductions in the risk of G-tube dependence in the acute period (up to 1 month) as reported by two other studies, one evaluating patients with primary nasopharyngeal cancer[111] (20% vs. 65%; adj. OR for IMRT 9.33; 95% CI 1.74 to 75.96) and the other patients with nasopharyngeal or paranasal sinus cancers[178] (adj. OR 0.11; 95% CI <0.01 to 0.61). The fifth study included patients with either primary or metastatic salivary gland cancer treated with PBT versus IMRT; no patient required a reactive gastrostomy tube or tracheostomy up to 3 months of follow-up in this study.[238]

Figure 18. The Frequency of Gastrostomy Tube Dependence Following PBT Compared with IMRT for Curative Intent in Adults with Various Non-Skull Base Head and Neck Cancers.



3D = three dimensional; adj. = adjusted for confounding factors; CI = confidence interval; IMPBT = intensity-modulated proton beam therapy; IMRT = intensity modulated radiation therapy; NC = not calculable; NR = not reported; OR = odds ratio; PBT = proton beam therapy.
*Adjuvant pencil beam scanning PBT vs. IMRT via volumetric modulated arc therapy (VMAT) following transoral robotic surgery and selective neck dissection.

In a sixth study (primary oropharyngeal cancer),[326] fewer patients who received PBT showed signs of radiation-induced osteonecrosis to include no cases of grade 3 or 4 osteonecrosis, however the differences were not statistically significant (Table 29). The small number of patients for PBT may have precluded identification of rare events. Of note, the mean radiation dose to the mandible was significantly lower in the PBT group (25.6 vs. 41.2 Gy) though the overall dose was similar between groups; the authors state that since radiation dose to the mandible is the main risk factor for osteoradionecrosis this dose reduction could be related to the lower incidence and lesser severity of osteoradionecrosis following PBT.

Table 29. Incidence of Osteoradionecrosis in One Retrospective Cohort Study Comparing PBT versus IMRT for Curative Intent in Patients with Primary Oropharyngeal Cancer

		Grade of Osteoradionecrosis	IMPBT (n=50) % (n)	IMRT (n=534) % (n)	RR (95% CI)†
Zhang 2017 (N=584) Primary Oropharyngeal Cancer	Late toxicities (>6 months)*	Any	2.0% (n=1)	7.7% (n=41)	RR 0.26 (0.04 to 1.85)
		Grade 1	2.0% (n=1)	4.3% (n=23)	RR 0.46 (0.06 to 3.37)
		Grade 2	0%	0.2% (n=1)	NC; p=0.76
		Grade 3	0%	0.9% (n=5)	NC; p=0.49
		Grade 4	0%	2.2% (n=12)	NC; p=0.29

CI = confidence interval; IMPBT = intensity-modulated proton beam therapy; IMRT = intensity modulated radiation therapy; NC = not calculable; RR = risk ratio.
*Earliest occurrence was 6.2 months.
†Crude RR (95% CI) and p-values calculated by AAI. The small number of patients for PBT may preclude identification of rare events.

Case series of non-skull-base head and neck tumors

A total of 14 case series of non-bone head and neck tumors were identified that reported safety outcomes.[59,91,93,103,104,179,193,199,224,239,272,281,324,325]

Overall the frequency of treatment-related deaths ranged from 0% to 3.7% across eight studies of various head and neck cancers (n=34 to 84).[59,103,104,179,224,239,272,281] Acute (≤3 months) treatment-related deaths were rare, ranging from 0% (1 study of tongue cancer)[272] to 1.7% (in two studies of recurrent mixed head and neck[179,224]). Late (>3 months) treatment-related deaths ranged from 0% to 3.7% across six studies of various cancer types.[59,103,104,179,239,281] Treatment-related deaths tended to be higher among recurrent and sinonasal cancers. Of note, patients across all studies received various concurrent or adjuvant therapies besides PBT, therefore it is unclear to what degree PBT specifically related to these deaths.

Six case series reported severe acute toxicities using a definition of ≤3 months, consistent with this report[93,103,179,199,224,239]. Another four case series referred to acute toxicities but did not define a timeframe.[91,193,272,281] One additional case series defined acute toxicities as occurring within a 6 month timeframe. The rates are as follows:

- Any grade: 35.5% overall (1 study of anterior skull based malignancies[91])
- Grade ≥3: 11% study of sinonasal cancers[281]); 3% to 79% for various specific toxicities (1 study of tongue cancers[272])

- Grade 3: 12% to 30% overall, 0% to 9.9% across specific toxicities (1 study of olfactory neuroblastoma,[199] 3 studies of recurrent mixed head & neck diagnoses[179,224,239]); 2% to 58% across specific toxicities (1 study of oropharyngeal cancer[93])
- Grade ≥4: 0% to 2.9% overall (1 study of mixed head & neck diagnoses [193], 1 study of recurrent oral cancers,[103] 2 studies of recurrent mixed head & neck diagnoses[179,239])

Five case series reported late toxicities defined as occurring after 3 months.[103,179,199,224,239] Two case series used alternative late toxicity definitions, either greater than 6 months[91] or greater than 24 months.[272] An additional five did not clarify their definitions of late toxicity timeframes. [59,193,281,324,325] The rates of late toxicities are as follows:

- Any grade: 54.8% overall (1 study of non-bone anterior skull-based malignancies[91])
- Grade ≥3: 2.9% to 24% overall (7 studies; 2 studies in oral or tongue cancers,[103,272] 1 study of olfactory neuroblastoma,[199] 2 studies in sinonasal cancers,[59,281] and 2 studies in mixed head & neck[193,224])

One other case series did not specify toxicities as acute or late[104] and the rates of grade ≥3 hematological and non-hematological varied widely.

The rate of secondary malignancies was 1.2% as reported one case series of sinonasal cancers.[59]

Radiation necrosis (including temporal lobe necrosis, encephalopathy necrosis, bone or soft tissue necrosis, CNS necrosis and brain necrosis) was reported by 10 case series.[59,93,103,104,179,199,224,272,281,308,325]. All were reported as late (>3 months) toxicities, although four case series did not report a definition of timeframe.[59,104,272,308,325] The rates reported were:

- Grade ≥3 Bone or Soft Tissue Necrosis: 0% to 15.2% overall (7 studies; 1 study of recurrent mixed head & neck diagnoses[179], 4 studies of oral cancer,[93,103,104,272] and 2 study in sinonasal cancers[59,325]); 0% to 8.3% when recurrent cancers are excluded[59,93,103,272,325]
- Grade ≥3 CNS Necrosis: 0.5% in one study of sinonasal cancers[59] and 1.2% in skull-base chordomas and chondrosarcomas[308]
- Grade ≥3 Brain Necrosis: 0% to 7.9% overall (5 studies; 1 study in olfactory neuroblastoma, [199] 2 studies in sinonasal cancers,[281,325] 2 studies in skull-base chordomas and chondrosarcomas[308,309]
- Grade 5 Encephalopathy Necrosis: 1% (1 study of skull base chondrosarcomas [81])

Across any grade of various other types of radiation necrosis were reported by three studies [177,224,308] with rates ranging from 5% to 12.4%.

4.3.6.2 Skull-base and cervical bone tumors

Key Question 1 (Effectiveness, curative intent)

Comparative studies of chondrosarcoma

One small (N=47)[259] retrospective cohort study was identified that evaluated surgical resection with or without adjuvant PBT (total dose 70 Gy RBE) for the treatment of grade II skull base chondrosarcoma. Mean patient age was 47 years and just over half (51%) were female (Table 27). The anatomical location of the tumor differed significantly between the treatment groups group (petroclival tumors were present in 96% of patients who received adjuvant PBT compared with only 50% who received surgery

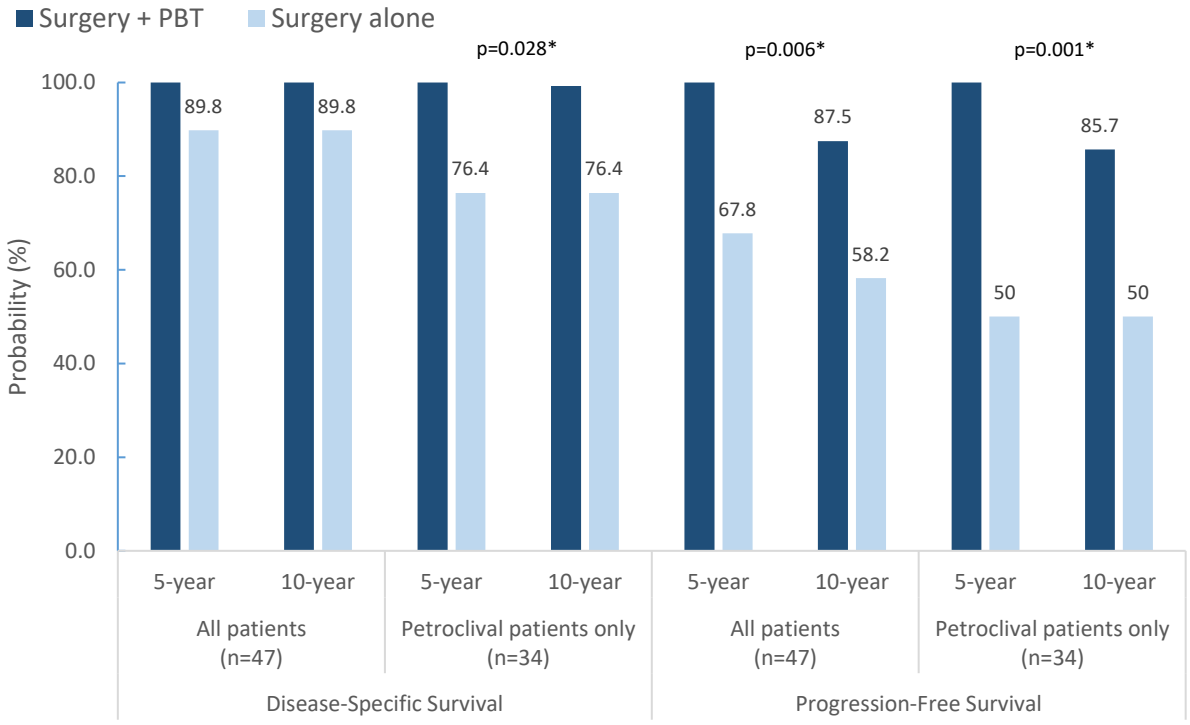
only) as did the extent of surgical resection (partial in 87% vs. 46%, respectively) and abutment of the tumor against the internal carotid artery (74% vs. 42%, respectively)

Survival outcomes

Comparative studies of chondrosarcoma

No statistically significant differences were seen in the probability of 5- or 10-year disease-specific survival (DSS) between patients who did and did not receive adjuvant PBT following surgery in one small retrospective cohort study[259], though PBT patients tended to have improved DSS compared with surgery alone and differences may be clinically meaningful (Figure 19). The probabilities of 5- and 10-year PFS, however, were significantly higher in the group that received adjuvant PBT. When only patients who had petroclival chondrosarcoma were considered, the probabilities of DDS and PFS were significantly higher in patients who received PBT after surgery compared with those who did not (Figure 19).

Figure 19. Disease-specific and progression-free survival following surgery with and without adjuvant PBT from a retrospective comparative study of patients with skull-base chondrosarcoma



PBT = proton beam therapy.
*p-values were provided only for differences that were statistically significant. The difference in disease-specific survival at 5 and 10 years (100% vs. 90%) did not differ statistically between groups.

Case series of skull-base and cervical tumors

Six case series of skull-base and cervical tumors reported OS,[64,78,81,267,308,309] four of which were in populations undergoing curative and salvage therapy. There was substantial overlap in populations between two of the studies.[308,309] The range of probabilities for OS across time frames in these studies is as follows:

- 5-year: 75% to 88.3% (4 studies, N=76 to 222)[64,81,267,308]
- 7- to 10-year: 80% to 93.6% (3 studies, N=159 to 251)[78,308,309]

Two case series in skull-base and cervical tumors reported PFS,[64,78] including one study each receiving PBT for curative intent or mixed curative and salvage therapy. The 5-year probability of PFS ranged between 50% and 93% (2 studies undergoing PBT for curative and salvage therapy[64] and curative intent only[78]).

Other outcomes

Comparative studies of chondrosarcoma

In one small retrospective cohort[259] comparing surgery with and without adjuvant PBT for chondrosarcoma, there were a total of four deaths (8.5%) from any cause (not reported by group) over a median follow-up of 7.6 years; two were disease specific and due to post-operative complications (cerebral abscess) and disease relapse. Nine cases (19%) of local relapse were reported, eight of which occurred in patients who did not received post-operative PBT (33% vs. 4%; RR for PBT 0.13, 95% CI 0.02 to 0.96, p=0.01). Of the nine patients experiencing local relapse, five were treated with secondary PBT. No regional or distant metastases were observed.

Case series of skull-base and cervical tumors

Four case series in skull-base chordomas and chondrosarcomas reported the probability of local control. The 5-year ranges of local control was 71% to 96%.[64,78,81,308] Three of these studies included a mix of curative and salvage PBT; the range across these studies was 71% to 81%[64,81,308]

One case series reported distant metastasis-free survival following a mix of curative and salvage PBT in patients with skull-based chordomas and chondrosarcomas (N=222); at 5 and 7-years the probabilities were 92%.[308]

Key Question 2 (Effectiveness, salvage therapy)

No comparative studies that met inclusion criteria were identified that evaluated PBT for salvage treatment (including treatment for recurrent disease) of skull-base tumors. One case series (N=61) was included that evaluated skull-base (90%) and cervical (8%) tumors (the remaining 2% of patients had oropharyngeal cancer), primarily of the squamous cell subtype (Abstraction Appendix G).[179]

The probability of 1- and 2-year OS, respectively, was 56% and 33% at 2-years; the cumulative incidences of local failure (with death as a competing risk), regional nodal failure, and distant metastases were 20%, 3% and 38%, respectively.

Key Question 3 (Safety)

Comparative studies of chondrosarcoma

Patients who received adjuvant PBT showed a higher risk for any complication, primarily related to hearing loss; sensorineural and severe hearing loss were six- and five-times more frequent compared with patients who received surgery only in one small comparative study of chondrosarcoma (Table

30);however, confidence intervals were very wide. Dizziness was also more frequent in PBT patients. There was no statistical difference between groups in the risk of grade ≥3 toxicities. Temporal lobe necrosis was observed in five (18%) PBT patients. Other outcomes reported, including those specific to PBT and surgery, can be found in Table 30 below.

Table 30. Complications following PBT compared with surgery in one retrospective cohort study evaluating patients treated for skull-base chondrosarcoma

	PBT (N=28)*		Surgery (N=47)†		RR (95% CI)‡
	%	n	%	n	
Any complication	68%	19	26%	12	2.7 (1.5 to 4.6)
Sensorineural hearing loss	39%	11	6%	3	6.2 (1.9 to 20.2)
Severe hearing loss	21%	6	4%	2	5.0 (1.1 to 23.3)
Dizziness	14%	4	0%	0	NC, p=0.008
Conductive hearing loss	11%	3	4%	2	p=0.28
Any grade ≥3 toxicity	25%	7	11%	5	p=0.10
Cranial nerve palsy	11%	3	19%	9	p=0.34
Treatment-related death	0%	0	2%	1	p=0.44
Vision loss	11%	3	-----	-----	-----
Hypopituitarism	18%	5	-----	-----	-----
Temporal lobe necrosis	18%	5	-----	-----	-----
Cerebrospinal fluid leak	-----	-----	13%	6	-----
Meningitis	-----	-----	9%	4	-----
Pulmonary embolism	-----	-----	2%	1	-----

Simon 2018
*Total number of patients having received Proton therapy (23 primary treatment and 5 secondary treatment after local failure).
†Total number of patients in the study since all patient received surgery (either alone, or with PBT).
‡Calculated by AAI only for those differences that were statistically significant.

Case series of skull-base and cervical tumors

A total of 7 case series in skull-base bone tumors were identified that reported safety outcomes.[64,67,78,81,177,308,309]

The rate of grade ≥3 acute toxicities was 0% to 9% across four studies of skull-base chordomas and chondrosarcomas.[64,78,267,309] Two case series used a definition of ≤3 months and two did not report toxicity timeframe.

Late toxicities were reported by six case series,[64,78,81,267,308,309] with an overall rate of 1.3% to 24% for grades ≥3. Four studies used a definition of >3 months[78,81,267,308] and two did not report their timeframe definition.[64,309]

One other case series did not specify toxicities as acute or late[67] but reported a rate of 0% for general toxicities grades ≥3.

The rates of secondary malignancies was 0% as reported by one case series in skull-base chordomas.[67]

Radiation necrosis was reported in three case series,[64,308,309] with an overall rate of 0.3% to 2% (0.3% to 2% brain necrosis, 2% bone and soft-tissue necrosis).

Key Question 4 (Differential Effectiveness and Safety)

No studies that met inclusion criteria were identified.

Key Question 5 (Economic)

One good quality CEA (QHES 90/100)[253] evaluated the cost-effectiveness of PBT accompanied by chemotherapy compared with IMRT (also accompanied by chemotherapy) for a hypothetical cohort of 65 year-old patients with stage III-IV oropharyngeal squamous cell carcinoma using Markov modeling. The source of funding was not reported. Incremental cost-effectiveness ratios (ICER) were reported to reflect the marginal cost of adding one quality adjusted life year (QALY) to a patient's life when using PBT versus IMRT. The primary limitation of this study is that oncologic outcomes were assumed to be same for IMRT and PBT despite lack of evidence, including long-term evidence.

Key points

One good quality CEA took both societal and payer perspectives and concluded that, compared with IMRT, PBT was not cost-effective for patients with stage III-IV oropharyngeal squamous cell carcinoma using either perspective.[253] However, at extremes of PBT superiority, it becomes cost-effective for younger human papilloma virus (HPV)-positive patients.

- ICER for societal perspective: $390,000/QALY for HPV positive patients (range not reported), $695,000/QALY for HPV negative patients (range not reported). ICER for payer perspective: $288/QALY for HPV positive patients, $516,000/QALY for HPV negative patients.
- Probability of PBT cost effectiveness for 55 year old patient: 0.4% (payer perspective) and 2% (societal perspective) at willingness-to-pay (WTP) $100,000/QALY; 25% (payer perspective) and 2% (societal perspective) at WTP $150,000/QLAY
- Limitations:
    - Oncologic outcomes were assumed to be same for IMRT and PBT despite limited evidence
    - Improved side effect profile of PBT was assumed from a single case series
    - Implications of costs from toxicities not described; societal costs assumed to be same for both treatment modalities
    - A lifetime time horizon was used; no comparative data on long-term outcomes was reported.
    - Where multiple toxicities were present, disutilities were added potentially over-estimating the disutility for combined toxicities and thus under-estimating QALYs from IMRT
    - Societal costs were assumed to be the same for both treatments

Detailed results

Study characteristics and framework

One good quality CEA evaluated the cost-effectiveness of PBT accompanied by chemotherapy compared with IMRT (also accompanied by chemotherapy) for 65 year-old patients with stage III-IV oropharyngeal squamous cell carcinoma[253] (Table 31). The costing year was 2016. A lifetime horizon was used. The study adopted both payer and societal perspectives but did not include costs associated with toxicities. Clinical data specific to PBT were from one small registry study comparing PBT with IMRT and longitudinal case series. Costs included treatment, chemotherapy, gastrostomy costs, recurrence costs, and dental costs for osteonecrosis. The payer perspective costs for PBT and IMRT were $45,457 and $23,137, respectively. The societal perspective costs for PBT and IMRT were $56,659 and $27,192, respectively. Probabilities for relapse and death for disease progression were derived from a phase III clinical trial conducted by the Radiation Therapy Oncology Group[204] and a subgroup analysis of the participants from the same trial[75], respectively. Data on toxicities came from observational studies of patients with oropharyngeal cancer including some included in this report.

Base Case Results

Using the payer perspective, PBT and IMRT, respectively, were found to cost $107,649 and $87,485 for HPV-positive patients and $191,769 and $171,129 for HPV-negative patients. From the societal perspective, PBT and IMRT, respectively, were found to cost $118,852 and $91,541 for HPG-positive patients and $202,972 and $175,185. PBT resulted in 12.96 QALYs for HPV-positive patients and 8.45 QALYs for HPV-negative patients. IMRT resulted in 12.89 QALYs for HPV-positive patients and 8.41 QALYs for HPV-negative patients. The ICER for the payer perspective was $288,000/QALY for HPV-positive patients and $516,000/QALY for HPV-negative patients. The societal perspective ICER was $390,000/QALY for HPV-positive patients and $695,000/QALY for HPV-negative patients.

Sensitivity Analyses

Both one-way sensitivity and probabilistic sensitivity analyses (PSA) were conducted.

In one-way sensitivity analyses, ICERs were uniformly above $100,000/QALY for both perspectives, even under assumptions that strongly favored the effectiveness of PBT to reduce percutaneous gastrostomy tube (PEG) dependence or improve long-term xerostemia. PSA similarly suggested that PBT is not cost effective. The probability that PBT was cost-effective was 0% in both perspectives with a WTP of $100,000/QALY. Using a WTP threshold of $150,000/QALY, the probability that PBT was cost-effective was 0.4% from a payer perspective and 0% from a societal perspective. For 55 year-old patients, PBT was cost-effective 0.4% of the time using a payer perspective and 0% of the time using a societal perspective with a WTP of $100,000/QALY. With a WTP of $150,000/QALY, PBT was cost-effective for 55 year-old patients 25% of the time from the payer perspective and 2% of the time from the societal perspective.

Conclusions and Limitations

The authors concluded that PBT is not cost-effective using a WTP threshold of $100,000/QALY. The only scenario in which PBT was cost-effective was using the payer perspective among younger, HPV-positive

patients with the assumption that PBT leads to profound improvements in dysphagia and xerostemia. PBT was not cost-effective for HPV-negative individuals in any of the scenarios examined.

The primary limitation of this study is that oncologic outcomes were assumed to be same for IMRT and PBT despite lack of evidence. The authors' findings could be substantially impacted if this assumption does not hold true. Future research is needed to explore this. In addition, the authors assumed that the side effect profile of PBT was preferable to that of IMRT based on minimal observational evidence. This assumption favors PBT, thus the resulting ICERs were conservative in nature. Societal costs were assumed to be the same for both treatment modalities. If the improved side effect profile of PBT is associated with substantial cost savings, this model did not capture the effect. Finally, where multiple toxicities were present, disutilities were combined. For example, the disutility for an individual with both dysgeusia and xerostomia would be 0.118: 0.059 (the disutility for dysgeusia) + 0.059 (the disutility for xerostomia). This potentially over-estimated the disutility for combined toxicities (i.e. combined utility would be lower versus what may be expected for a given single toxicity), thus under-estimating QALYs from IMRT. The QHES score for this study was 90/100 points.

Table 31. Summary of the economic study comparing PBT with IMRT in patients with oropharyngeal squamous cell carcinoma

	Sher 2018
Population	65 year old patients with stage III-IVB oropharyngeal squamous cell carcinoma
Intervention(s)	PBT (timing unclear, accompanied by chemotherapy)
Comparator(s)	IMRT
Country	USA
Funding	NR
Study design	CUA
Perspective	Payer (Medicare) and societal
Time horizon	Lifetime
Analytic model	Markov model with 6 health states
Effectiveness outcome	QALYs
Effectiveness outcome components	Chemoradiotherapy, percutaneous gastrostomy tube, dysgeusia, xerostomia, distant metastasis, death
Source for effectiveness data	Prior literature (randomized phase III trial, PBT registries, case series)
Costing year	2016
Currency	USD
Discounting	3%
Components of cost data	Treatment cost, chemotherapy, gastrostomy costs, locoregional recurrence, distant recurrence, dental cost for osteonecrosis
Cost sources	Medicare payment schedule, case series, construction costs used by the UT Southwestern Medical Center Department of Radiation Oncology
Sensitivity analysis	One-way; PSA: Markov model run over 50,000 trials, triangular distributions of toxicity outcomes
QHES	90

	Sher 2018
Results:	
Cost / QALY of intervention	**Societal perspective** HPV positive: $118,852/12.96 = $9,171/QALY HPV negative: $202,972/8.45 = $24,020/QALY **Payer perspective** HPV positive: $107,649/12.96 = $8,306/QALY HPV negative: $191,769/8.45 = $22,695/QALY
Cost / QALY of comparator(s)	**Societal perspective** HPV positive: $91,541/12.89 = $7,102/QALY HPV negative: $175,185/8.41 = $20,831/QALY **Payer perspective** HPV positive: $87,485/12.89 = $6,787/QALY HPV negative: $171,129/8.41 = $20,348/QALY
ICER	**Societal perspective** HPV positive: $390,000/QALY HPV negative: $695,000/QALY **Payer perspective** HPV positive: $288,000/QALY HPV negative: $516,000/QALY
One-way SA	Even under assumptions that favored efficacy of PBT to reduce PEG dependence and improve long-term xerostemia, ICERs uniformly above $100,000/QALY (range $101,000/QALY to $1 mil/QALY)
Other SA	Ranging relative benefit of PBT from 0% to 50% in xerostemia, gastrostomy use, and dysgeusia: probability PBT cost-effective was 0% (both perspectives) at WTP of $100,000/QALY and 0.4% (payer) and 0% (societal) at WTP $150,000/QALY PBT cost effective for 55 year-old patients at WTP $100,000/QALY in 0.4% for payer and 2% for societal; at WTP $150,000/QALY 25% (payer) and 2% (societal) were cost-effective
Author's Conclusion	PBT is not cost-effective using either societal or payer perspective; at extremes of PBT superiority it becomes cost-effective for younger HPV-positive patients
Limitations	• Oncologic outcomes assumed to be same for IMRT and PBT despite limited evidence • Lifetime time horizon, however no long-term comparative data available • Improved side effect profile of PBT assumed from minimal 1 case series

	Sher 2018
	• Societal costs assumed to be same for both treatment modalities • Disutilities for toxicities assumed to be additive, potentially under-estimating QALYs from IMRT

CMS = Centers for Medicare and Medicaid Services; CUA = cost-utility analysis; Gy = Gray (unit of absorbed dose); HPV = human papilloma virus; ICER = incremental cost-effectiveness ratio; IMRT = intensity-modulated radiation therapy; PBT = proton beam therapy; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life-year; QHES = Quality of Health Economic Studies; QOL = quality of life; SA = sensitivity analysis; SBRT = stereotactic body radiation therapy; WTP = willingness-to-pay.

4.3.7 Liver Tumors

Key Points

- No statistical differences were seen between PBT and transarterial chemoembolization (TACE) for the probabilities of 2-year OS, PFS, and local control in one small RCT of adult patients with unresectable hepatocellular carcinoma (HCC) treated with curative intent, though PFS and local control tended to be greater following PBT (Moderate SOE).
- OS was statistically higher following PBT versus intensity-modulated radiation therapy (IMRT) in one retrospective cohort study of adult patients with unresectable HCC but there was no difference in local and regional control between groups (Low SOE).
- Acute toxicity and serious complications were not well described in the RCT. Fewer patients who received PBT compared with TACE were hospitalized for a complications within 30 days of treatment, translating into fewer total days hospitalized for complications (Moderate SOE). In the retrospective cohort study, compared with IMRT, PBT was associated with a lower risk of nonclassic radiation-induced liver disease (RILD) (Low SOE) and death due to liver failure (Insufficient SOE).
- One poor quality cost-utility analysis (QHES 51/100) from Taiwan compared PBT with stereotactic body radiation therapy (SBRT) for a hypothetical cohort of patients with advanced, inoperable hepatocellular carcinoma using Markov modeling from a payer perspective and concluded that PBT is cost-effective for high risk patients at a willingness-to-pay threshold (WTP) of New Taiwan Dollars $2,157,024 per quality-life years (QALY) gained.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy (i.e., no comparative studies) or differential effectiveness and safety in this population.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of included studies

A total of 14 studies evaluating PBT for the treatment of liver tumors that met inclusion criteria were identified: one RCT,[42] one retrospective comparative study[244], and seven case series[79,90,114,183,187,213,318] of PBT for curative intent and five case series of PBT for salvage therapy.[80,113,140,143,322] The RCT was considered moderately low risk of bias. For the reasons described previously (Section 4.1), all comparative cohort studies are considered moderately high risk of bias; however, the study[244] included here did control for confounding. All case series were considered to be at high risk of bias.

In addition, one cost-utility analysis (CUA) compared PBT with stereotactic body radiation therapy (SBRT) for patients with advanced, inoperable hepatocellular carcinoma was identified that met inclusion.[159]

Results

Key Question 1 (Efficacy/Effectiveness, curative intent)

One small RCT (N=69)[42] that compared passive scatter PBT with transarterial chemoembolization (TACE) for curative intent in adult patients with unresectable hepatocellular carcinoma (HCC) was identified (Table 32). Patients were primarily male (71%) with a mean age of 60 years; almost all patients had cirrhosis of the liver (95%), primarily due to hepatitis C. Multiple tumors were present in just over half of the patients (55%) with a maximum tumor size of 3.2 cm. For the PBT group, median total radiation dose was 70.2 Gy(RBE); 82% of patients had a single round of PBT and 18% had up to three treatments. In the TACE group, 58% of patients underwent single treatment while 42% had up to four treatments (for persistent disease). There was no mention of patients in the PBT group receiving either induction or concurrent/adjunctive chemotherapy and all eligible patients had untreated HCC. This study was considered to be at moderately low risk of bias (i.e., moderate quality). Methodological shortcomings included failure to report allocation concealment methods and whether or not outcomes assessment was blinded. Of note, this was an interim analysis of an ongoing clinical trial.

One retrospective comparative study (N=133)[244] was also identified which compared passive scatter PBT versus intensity modulated radiation therapy (IMRT) for primarily curative intent (83% vs. 17% for recurrence) in adult patients with unresectable HCC. Patients were mostly male (76%) with a median age of 68 years. Several baseline characteristics differed statistically between the groups; those receiving PBT had a higher incidence of underlying cirrhosis, but had better Child-Pugh and albumin-bilirubin scores, compared with IMRT (Table 32). Most patients had not undergone any previous therapy (62% overall; 76% for PBT vs. 55% for IMRT, p=0.10). The median total radiation doses were identical between PBT [67 Gy(RBE), IQR 60 to 70] and IMRT (67 Gy, IQR 67 to 82). Median follow-up was 14 months. Of note, 32 of the 49 patients receiving PBT in this study were also included in the case series by Hong et al. 2016 (below).

Table 32. Liver Cancer in Adults: Study Characteristics and Demographics for Comparative Studies Comparing PBT versus TACE and Photon RT for Curative Intent

	Bush 2016		Sanford 2019	
	PBT (n=33)	TACE (n=36)	PBT (n=49)	Photon RT (n=84)
Patient Characteristics				
Males, %	76%	67%	80%	73%
Age, years; mean ± SD	61.4 ± NR	58.9 ± NR	Median (IQR): 65 (60 to 74)	Median (IQR): 69 (61 to 79)
Comorbidities				
Cirrhosis	97%	94.4%	96%*	77%*
Hepatitis C	67%	69%	49%	29%
Hepatitis B	3%	3%	12%	5%
Tumor thrombus	---	---	27%	35%
Tumor characteristics				
Subtype	Hepatocellular Carcinoma		Hepatocellular Carcinoma	
Mean maximal tumor size (range), cm	3.2 (1.8 to 6.5)	3.2 (2.0 to 6.5)	---	---
Multiple Tumors	54.5%	55.6%	---	---
Treatment Characteristics				
Technique	Passive Scatter	Transfemoral arterial approach	3D passively scattered	IMRT
Median total radiation dose (IQR)	70.2 CGE (NR)	NA	67 Gy (60 to 70)	67 Gy (67 to 82)
Chemotherapeutic drugs	---	(1) Ethiodol, carboplatin 50-100 mg, doxorubicin 20-50 mg (+/- mitomycin 10 mg); (2) 75-150 mg doxorubicin on 100-300 µm LC bead microspheres†	---	---
No. fractions (radiation)	15	---	---	---
No. of treatments	Single treatment: 82% Up to 3 treatments: 18%	Single Treatment: 58% Up to 4 treatments: 42%†	---	---
Treatment/procedures prior to Radiation				
Biopsy	27%	33%	---	---
Ablation	---	---	10%	10%
Chemoembolization	---	---	6%	14%

	Bush 2016		Sanford 2019	
	PBT (n=33)	TACE (n=36)	PBT (n=49)	Photon RT (n=84)
Selective internal RT	---	---	0%	2%
Chemotherapy	---	---	0%	8%
Resection	---	---	2%	1%
Multiple Treatments	---	---	6%	10%
Study Design	**Randomized Controlled Trial**		**Retrospective Comparative Cohort**	
Follow-up, months (% followed)	**28 (98%)**		**14 (NR)**	
Risk of bias	**Moderately Low**		**Moderately High**	

CGE = Cobalt Grey Equivalent; Hep = hepatitis; NR = Not reported; PBT = Proton Beam Therapy; SD = Standard Deviation; TACE = Transarterial chemoembolization.
*Indicates a statistically significant difference between the groups
†For persistent disease.
†Intially patients were treated with a conventional chemoembolization protocol (#1); however a nationwide shortage of Ethiodol required a change in the drug delivery protocol mid-study (#2).

Additionally, six case series[79,90,114,183,187,213] evaluated effectiveness of PBT for curative intent were identified; six studies reported outcomes for patients with hepatocellular carcinoma (HCC) (N range, 22 to 250)[79,90,114,183,187,213] and two reported outcomes for patients with intrahepatic cholangiocarcinoma (ICC) (N = 21 and 39).[90,114] Of note, there is substantial overlap in patient population in the latter two case series which stratify treatment results by HCC and ICC.[90,114] None of the case series mention whether or not patients received concomitant or adjuvant chemotherapy; across three studies 48% to 90% of patients had undergone previous treatment,[183,187,213] in two studies none of the patients had prior radiation therapy[90,114] and in one study all patients were previously untreated.[79]

Survival outcomes

As reported by one small RCT,[42] the probability of 2-year overall survival (OS)was 59% for the whole population and did not differ statistically between those who received PBT versus TACE (data not provided). For those who went on to receive a liver transplant post-treatment (12 PBT and 10 TACE), 2-year OS was 82%, again with no statistical difference between groups (data not provided). The probability of 2-year PFS was greater following PBT compared with TACE (48% vs. 31%), however the difference failed to reach statistical significance (p=0.06). Sample size may have played a role in this finding.

In the retrospective cohort study[244], the probability of 2-year OS was statistically higher following PBT compared with IMRT: 59.1% versus 28.6% (adj. HR 0.47; 95% CI 0.27 to 0.82).

Across the case series, the probabilities of both OS and PFS were generally greater following PBT for HCC compared with ICC. For patients with HCC, the probability of 1- and 2-year OS, respectively, ranged from 77% to 86% (4 studies)[90,114,183,187] and from 56% to 88% (4 studies)[90,114,183,213]; for ICC, corresponding probabilities ranged from 60% to 70% and from 34% to 47% (2 studies).[90,114] OS at 5-years ranged from

46% to 51% across three studies of HCC.[79,187,213] The probability of PFS was reported at 1 and 2 years by two case series, one in patients with HCC (70% and 60%, respectively)[183] and the other in ICC (41% and 26%, respectively)[114]; at 5 years, PFS was 17% in one case series of HCC,[79] (Main Appendix F, Tables F26 and F27).

Other outcomes and secondary outcomes

In the RCT, the probability of 2-year local control was greater following PBT compared with TACE (88% vs. 45%), however the difference failed to reach statistical significance (p=0.06).[42] Sample size may have played a role in this finding.

In the retrospective cohort study[244], the probability of 2-year local control was high and was similar between PBT (93%) and IMRT (90%); the HR for the cumulative incidence of *local failure* was 0.74 (95% CI 0.18 to 3.01). The cumulative incidence of locoregional recurrence was somewhat greater in the PBT group (53% vs. 42% with IMRT) however, the difference was not statistically significant on multivariate analysis: adjusted HR 0.98 (95% CI 0.54 to 1.75).

Across four case series (N range, 39 to 250), the probability of local control was as follows: 98% at 1 year (1 study of HCC),[187] 94% to 95% at 2 years (2 studies, 1 HCC and 1 ICC),[114,183] and 83% to 85% at 5 years (2 studies of HCC).[79,187] Probabilities were similar between ICC and HCC subtypes (Main Appendix F, Tables F26 and F27).

Key Question 2 (Efficacy/Effectiveness, Salvage therapy)

No comparative studies were identified that met inclusion criteria. Five case series evaluating PBT for salvage therapy provided data on effectiveness.[80,113,140,143,322] Three studies included patients with HCC[140,143,322] and two were in patients with liver metastases; in both studies the primary tumor sites were the colorectum (38% and 43%) and the pancreas (14% and 15%).[80,113] In four studies, PBT was performed for both salvage and curative intent with the majority of patients receiving salvage therapy for recurrence (56%-76%).[80,140,143,322] There is possible overlap in two of the patient populations.[140,143]

Survival outcomes

Two case series (N=41, 71) in patients with HCC reported the probabilities of OS, PFS, and relapse-free survival (RFS), respectively: at 2 years, rates were 51%, 88% and 25% in one study (all patients had tumor vascular thrombosis)[140] and at 3 years, rates were 74%, 90% and 27% in the other.[143]

Across the two studies of metastatic liver tumors, the probability of OS at 2 years ranged from 36% to 46%;[80,113] one study each reported OS at the following time points: 1 and 3 years (66% and 21%)[113] and 5 years (25%)[80] PFS was reported by one of these studies with 1 and 3 year probabilities of 25% and 9%.[113]

Other outcomes

One series in patients with HCC reported tumor response rates only at 1 and 3 months following PBT; at 1 month, 19% of patients showed progressive disease (2% were infield) which increased to 31% (3% infield) at 3 months.[322]

Distant metastases were reported in 16% and 42% of patients undergoing PBT for HCC over median follow-up periods ranging from 15.2 to 31.3 months in two case series.[140,143]

Across the studies evaluating salvage PBT for metastatic liver tumors, the probability of local control at 1 and 3 years was 72% and 61% (1 study)[113] and at 2 and 5 years was 66% and 53% (1 study).[80]

Key Question 3 (Safety)

Comparative studies

In the RCT,[42] data related to acute toxicity and serious complications were not provided. The authors only state that acute toxicity was generally limited to fatigue and radiation skin reaction in the PBT group and abdominal pain and nausea in the TACE group, which were experienced by most patients and that serious complications from PBT were uncommon. Statistically fewer patients who received PBT were hospitalized for a complication within 30 days of treatment (6% vs. 42%, p<0.001). Similarly, the total number of days hospitalized with 30 days of treatment was statistically lower following PBT compared with TACE: 24 days (0.73 day per patient) (all for complications) versus 166 days (4.6 days per patients) (113 days for complications), p<0.001. The authors do not specifically state whether or not patients in the PBT group had received either induction or concurrent or adjuvant chemotherapy but they indicate that all eligible patients had untreated HCC.

In the retrospective comparative cohort study[244], four PBT patients compared with 17 IMRT patients developed nonclassic radiation-induced liver disease (RILD) 3 month post-treatment which translated to a statistically lower incidence of RILD following PBT, odds ratio (OR) 0.26 (95% CI 0.08 to 0.86), a difference that persisted in multivariate analysis (data not provided). RILD was defined as worsening of Child-Pugh score by ≥2 points compared with baseline and was calculated in 100 (of 133) patients for whom data was available; denominators for this subset of patients by treatment group were not provided. Authors report that the development of RILD at 3 months was associated with significantly worse OS (HR 3.83; 95% CI 2.12 to 6.92). Among patients who died without disease progression, death as a result of liver failure was almost half as likely in patients who received PBT (53%, 8/15) versus IMRT (91%, 19/21), RR 0.59 (95% CI 0.36 to 0.97).

Case series

Toxicity was reported by six case series evaluating PBT for curative intent (N range, 37 to 250)[79,114,183,187,213,318] one of which included patients with both HCC (n= 44) and ICC (n=39)[114]; all other studies included only HCC. Two studies reported the frequency of acute toxicities grade ≥3 which ranged from 0% (0/83)[213] to 5% (2/40)[183]. Only one case series reported late toxicities, with no grade ≥3 events reported; there two (5%) grade 2 late events (GI bleed and rib fracture).[183] Three other case-series were unclear about the timing of toxicity. Two studies reported that grade ≥3 radiation-related toxicity occurred in 5% and 11% of patients (median follow-up periods 11 and 20 months)[114,318] and a third reported that hematologic abnormalities were the only toxicities grade ≥3 (no radiation dermatitis grade ≥3) but did not provide data.[79] Treatment-related toxicity resulting in liver failure and death was rare as reported by two case series: 0% (0/83)[213] and 2% (4/250).[187]

Across all five case series evaluating salvage PBT, grade ≥3 toxicity was rare (0% to 2%). Only one patient was reported to experience an acute grade ≥3 toxicity across three studies in patients with HCC (range 0% to 1%, N=41 to 101)[140,143,322]; no late toxicities of grade ≥3 (to include radiation-induced liver

disease or treatment-related death) were seen in two of these studies.[140,143] For the two case series evaluating metastatic liver tumors, grade ≥3 toxicity occurred 0% (0/89) and 1.6% (2/133) of patients.[80,113]

Key Question 4 (Differential Efficacy/Effectiveness and Safety)

No studies that met inclusion criteria were identified.

Key Question 5 (Economic)

One poor quality CUA[159] (QHES 51/100) from Taiwan compared PBT with stereotactic body radiation therapy (SBRT) for a hypothetical cohort of patients with advanced, inoperable hepatocellular carcinoma using Markov modeling from a payer perspective. The source of funding was not reported. Incremental cost-effectiveness ratios (ICER) were reported to reflect the marginal cost of adding one quality adjusted life year (QALY) to a patient's life when using PBT versus SBRT. The primary limitations of this study include lack of detail regarding cost components, use of clinical data from case series and limited sensitivity analysis and unclear applicability to the U.S. healthcare system. In addition it appears that tumor characteristics differed between treatment groups; it is unclear if this may have impacted modeling.

Key points

One poor quality CUA concluded that PBT is cost-effective for high risk patients with inoperable advanced large hepatocellular carcinoma from a payer perspective[159] at a WTP of New Taiwan Dollars (NT) $2,157,024 per QALY gained.

- ICER: (NT) $213,354/QALY (range not reported)
- Sensitivity analyses: ICER was sensitive to health status. PBT was cost-effective at a WTP of NT$2,157,024 in 97% of simulations.
- Limitations:
    - Clinical parameters were derived from separate case series of PBT and stereotactic body radiation therapy (SBRT); study selection process not transparent; source or basis of utilities was not well described
    - Intervention and comparator populations not comparable: important differences in patient populations including tumor size, Child-Pugh class and other factors were noted
    - One-way sensitivity analysis was not clearly presented; limited evaluation of assumptions was done; thus robustness of model is not clear
    - Components and basis for some medical costs not detailed
    - May not be applicable to U.S.

Detailed results

Study characteristics and framework

One poor quality CUA (QHES 51/100) evaluated the cost-effectiveness of PBT compared to SBRT for patients with inoperable advanced large hepatocellular carcinoma[159] (

Table 33). The study was conducted in Taiwan and the costing year was 2016. The time horizon was 5 years, which is likely reasonable given the 5-year survival rate for advanced HCC is low. Clinical data for PBT were from a phase II clinical trial evaluating the efficacy of PBT for patients with hepatocellular carcinoma.[138] Clinical data for SBRT were from sequential phase I and phase II trials including 102 patients with hepatocellular carcinoma.[40] The study adopted a payer perspective. Patients who received PBT were 70 years old and 67% were male. Patients who received SBRT were 69.4 years old and 78.4% were male. Costs included PBT or SBRT treatment, laboratory tests, and treatment for toxicity (specified as any grade 3/4 adverse event). The costs for PBT and SBRT were NT$300,000 and NT$213,660, respectively.

Base Case Results

Using the provider perspective, PBT was found to cost NT$557,907 more than SBRT and resulted in an additional 2.61 QALYs. The ICER was NT$213,354/QALY.

Sensitivity Analyses

Both one-way sensitivity and PSA were conducted.

In one-way sensitivity analyses, the model was highly sensitive to health state utilities for both stable and disease progression states, as well as direct medical costs (range of ICERs not reported). In a PSA, PBT was cost-effective at a WTP of NT$2,157,024 in 97% of simulations while SBRT was cost-effective in 4% of simulations.

Conclusions and Limitations

The authors concluded that PBT is cost-effective for patients with inoperable advanced hepatocellular carcinoma at the WTP threshold of Taiwan (NT$2,157,024 per QALY gained). In addition, the ICER could be considerably lower among patients with higher risk of severe toxicity from SBRT.

The primary limitation of this study is that the intervention and comparator populations are not comparable. There are important differences in the patient populations including tumor size, Child-Pugh class, and other factors. Thus, the results of the study may reflect the differences in the study populations, not a true evaluation of the treatment modalities. In addition, the components of the medical costs were not described in detail and deaths due to cancer were not included in the model. While the study used a 5 year time horizon, the studies they based the model on had only 1-2 years of follow-up data. Limited evaluation of assumptions was done, thus the robustness of the model is not clear. Finally, the results of this study may not be applicable to the United States. The QHES score for this study was 51/100 points.

Table 33. Summary of the economic study comparing PBT with SBRT in patients with advanced, inoperable hepatocellular carcinoma

	Leung 2017
Population	Inoperable advanced, large hepatocellular carcinoma **PBT study**: Age 70, 67% male; Child-Pugh Class A 67%; tumor size 45mm; Hepatitis C 87% **SBRT study:** Age 69.4, 78.4% male; Child-Pugh Class A 100%; tumor size 72mm; Hepatitis C 28%
Intervention(s)	PBT (timing unclear, appears to be primary treatment)
Comparator(s)	Stereotactic body radiation therapy (SBRT)
Country	Taiwan
Funding	NR
Study design	CUA
Perspective	Payer (Bureau of National Health Insurance)
Time horizon	5 years
Analytic model	Markov model with 3 health states
Effectiveness outcome	QALYs
Effectiveness components	Stable disease, disease progression and death
Source for effectiveness data	Prior Phase I/ II trial of SBRT, separate phase II study of PBT and expert opinion (identified through systematic literature review); source of utilities unclear
Costing year	2016
Currency	NT (New Taiwan dollars, no exchange rate given)
Discounting	3%
Components of cost data	Treatment cost, laboratory tests, treatment for toxicity
Cost sources	Bureau of National Health Insurance (Taiwan) database
Sensitivity analysis	One-way PSA: Monte Carlo simulations using 10,000 iterations; varied all parameters over range of ±30%; lognormal distributions for costs; beta distributions for probabilities, utilities, and toxicity
QHES	51
Results:	
Cost /QALY of intervention, comparator	NR
ICER	NT$557,907/2.61 = NT$213,354/QALY
One-way SA	Very sensitive to utilities and direct costs in both states (range of ICERs not reported)
Other SA	Results from Monte Carlo simulations: using threshold of NT$2,157,024 /QALY, PBT has 97% chance of being cost-effective and SBRT has 4% chance
Author's Conclusion	PBT is cost-effective for inoperable advanced HCC at a WTP threshold for Taiwan

	Leung 2017
Limitations	• Clinical data from separate case series of PBT and SBRT; study selection process for clinical outcomes is not transparent (no reporting of systematic review results); basis of utilities not described • Intervention and comparator populations not comparable: differences in patient populations including tumor size, Child-Pugh class and other factors. It is unclear how this may impact analsyses • Components and basis for some medical costs not detailed • Did not include non-cancer deaths • One-way sensitivity analysis not clearly presented; limited evaluation of assumptions, robustness of model is not clear • May not be applicable to US

CMS: Centers for Medicare and Medicaid Services; CUA: cost-utility analysis; Gy: Gray (unit of absorbed dose); ICER: incremental cost-effectiveness ratio; IMRT: intensity-modulated radiation therapy; NT: New Taiwan dollar; PBT: proton beam therapy; PSA: probabilistic sensitivity analysis; QALY: quality-adjusted life-year; QHES: Quality of Health Economic Studies; QOL: quality of life; SA: sensitivity analysis; SBRT: stereotactic body radiation therapy; WTP: willingness-to-pay

4.3.8 Lung Cancer

Key Points

- In one fair-quality RCT, no statistically significant differences were seen between PBT versus IMRT in the probability of OS at any timepoint up to 5 years or in the cumulative incidence of local failure in patients with non-small cell lung cancer being treated with curative intent (Moderate SOE). Findings from four retrospective comparative cohort studies were consistent with those of the RCT.
- For safety, no statistical differences were seen between PBT and IMRT in the frequency of grade ≥3 radiation pneumonitis at any timepoint up to 5 years in the fair-quality RCT (Moderate SOE). There was insufficient evidence from two retrospective cohort studies regarding grade ≥3 toxicities (radiation pneumonitis, radiation esophagitis, and radiation dermatitis) which did not differ statistically between PBT and IMRT; clinical importance of differences in unknown.
- The one comparative study of salvage PBT did not report survival or safety data; no studies that met inclusion criteria were identified that provided data on differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of Included Studies

A total of 19 studies were identified that evaluated PBT for the treatment of lung cancer: 17 studies (1 RCT, 5 retrospective comparative cohorts [to include the nonrandomized group from the RCT], and 11 case series) evaluated PBT for curative intent[44,45,101,108,123,136,156,161,174,190,205,206,212,233,242,282] and two studies (1 prospective comparative cohort, 1 case series) evaluated PBT for salvage therapy.[46,303]

Results

Key Question 1 (Efficacy/Effectiveness, curative intent)

One RCT[161] and four retrospective comparative cohort studies[108,206,233,282] that compared PBT with photon radiation therapies for curative intent in adult patients with non-small cell lung cancer (NSCLC) were identified (Table 34). In addition, 10 case series of PBT for curative intent were identified.[44,45,101,123,136,156,174,205,212,242]

The RCT evaluated the efficacy of passive scattering PBT versus IMRT in 173 patients with locally advanced, inoperable NSCLC.[161] Median patient age was 66 years and 57% were male; almost all patients (93%) had a history of smoking. Primary tumor subtypes were adenocarcinoma (52%) and squamous cell carcinoma (35%); the majority were stage III (44% IIIA; 36% IIIB). Mean radiation doses to lung and esophagus were similar in both groups; mean dose to the heart was statistically lower in the PBT group (6.9 vs. 10.2 Gy[RBE]). All patients received concurrent chemotherapy; 68% had also received induction chemotherapy. This report focused on data from the intention-to-treat (ITT) population; information regarding the per-protocol population can be found in Abstraction Appendix I. Also, a small subset of patients (n=39) from this trial who were unable to be randomized are described below with the nonrandomized cohort studies. This trial was considered to be moderately low risk of bias (i.e., moderate quality) due to failure to report allocation concealment methods.

Five comparative cohort studies (to include the nonrandomized cohort from the RCT) provided data on effectiveness (N = 39 to 1850)[108,161,206,233,282]; one study was a propensity-score matched analysis of patients from the National Cancer registry (N=1850 out of 243,822 patients).[108] Across studies, males comprised 46% to 57% of the populations and the median age ranged from 64 to 68 years, Table 34. Comorbidities were poorly reported. Primary tumor subtypes were adenocarcinoma (range, 31% to 74%) and squamous cell carcinoma (range, 28% to 59%), though the proportion of each varied across the studies. Four of the five studies reported disease stage and in general, most patients had stage III or IV NSCLC. The technique used for PBT was passive scatter in three studies[161,206,282] and either double scatter (81%) or intensity modulated pencil beam scanning (19%) in a fourth[233]; one database study did not indicate types of PBT used.[108] IMRT was the most common comparator (4 studies); one of these studies also compared PBT with 3D-CRT[282] and the database study included various types of photon RT (IMRT, 3D-CRT, "photon", and external beam not otherwise specified). Median total radiation dose for both treatment arms varied across studies (range, 50.4 to 74 Gy). All patients were receiving RT post-operatively in one study[233] and 13% of patients in the database study[108] had prior lung surgery. All patients were receiving concurrent and/or adjuvant chemotherapy.

For the reasons stated in the section 4.1, all comparative cohort studies were considered to be moderately high risk of bias (i.e., poor quality); however three[108,282,303] did control for confounding. All case series were considered to be at high risk of bias.

Table 34. Lung Cancer in Adults: Study Characteristics and Demographics for Studies Comparing PBT versus Photon RT for Curative Intent

	Effectiveness Only					Efficacy and Safety		Effictiveness and Safety					
	Higgins 2017*		Tucker 2016			**Liao 2018**		Liao 2018*		Niedzielski 2017		Remick 2017	
	PBT (n=348)	Photon RT (n=243,474)	PBT (n=45)	Photon RT (n=193)	Photon RT (n=230)	PBT [ITT] (n=72)	Photon RT [ITT] (n=101)	PBT [cohort] (n=13)	Photon RT [cohort] (n=26)	PBT (n=49)	Photon RT (n=85)	PBT (n=27)	Photon RT (n=34)
Patient Characteristics													
Males, % (n)	56.8%		56.4%			61%	53%	46.2%	50%	61.2%	52.9%	52%	41%
Age, years; median (range)	68 (NR)		64 (34–87)			66 (37–78)	66 (33– 85)	66 (42–76)	65 (39–79)	67 (38– 76)	65 (43– 85)	65 (38–77)	63 (38–80
Smoking													
Ever	---		---			94%	91%	92.3%	88.5%	---	---	74%	76%
Formerly	---		---			---	---	---	---	42.9%	68.2%	---	---
Currently	---		---			---	---	---	---	53.1%	21.2%	---	---
Tumor Characteristics													
Subtype													
Adenocarcinoma	30.6%		34.8%			50%	53%	30.8%	73.1%	51%	58.8%	67%	79%
SCC	37.6%		36.8%			36%	35%	53.8%	15.4%	36.7%	34.1%	7%	20%
NSCLC unspecified	---		28.4%			13%	7%	15.4%	3.8%	---	---	---	---
Large Cell	---		---			1%	2%	0%	0%	4.1%	3.5%	4%	---
Other	31.8%		---			0%	3%	0%	7.7%	8.2%	3.5%	22%	---
Stage													
0/I	14.9%		---			---	---	---	---	---	---	---	---
II to III	59.8%		---			---	---	---	---	---	---	---	---
II	---		---			11%	7%	---	---	---	---	---	---
IIA	---		---			---	---	---	---	4.1%	3.5%	---	---
IIB	---		---			---	---	---	---	12.2%	3.5%	---	---
IIIA	---		44.4%			38%	48%	30.8%	26.9%	40.8%	45.9%	---	---
IIIB	---		55.6%			42%	31%	61.5%	46.2%	40.8%	42.4%	---	---
IV	25.3%		---			7%	5%	0%	15.4%	2.1%	4.7%	---	---
Recurrent Disease	---		---			3%	10%	7.7%	11.5%	---	---	---	---

	Effectiveness Only					Efficacy and Safety		Effictiveness and Safety					
	Higgins 2017*		Tucker 2016			**Liao 2018**		Liao 2018*		Niedzielski 2017		Remick 2017	
Radiation Treatment Characteristics													
Technique	---	Various†	Passive scatter	3DCRT	IMRT	Passive Scatter	IMRT	Passive Scatter	IMRT	Passive Scatter	IMRT	Double scatter: 81% Intensity-modulated PBS: 19%	IMRT
Median total dose (Gy)	60	59.4	63	63	63	Lung: 17.2 Esophagus: 23.8 Heart: 6.9‡	Lung: 16.7 Esophagus: 27.4 Heart:10.2‡	Lung: 20.5 Esophagus: 34.7 Heart: 13.9	Lung: 20.4 Esophagus: 35.0 Heart: 14.6	74	74	50.4	54
Additional Treatments													
Prior to Radiation													
Chemotherapy	---		---			67%	68%	27.7%	39.3%	---	---	7%	12%
Chemotherapy													
Concurrent/Adjuvant	Timing NOS: 68.4%		100%/0%			100%/0%	100%/0%	100%/0%	100%/0%	100%/0%	100%/0%	22%/70%	32%/59%
Study Design	Retrospective propensity-score matched Comparative Cohort		Retrospective Comparative Cohort			Bayesian Adaptive RCT		Prospective Comparative Cohort§		Retrospective Comparative Cohort		Retrospective Comparative Cohort	
Follow-up, months (% followed)	39.6 (NR)	24 (NR)	24 (NR)			25.7 (95%)	24.1 (95%)	25.7 (NR)	24.1 (NR)	NR (%NR)	NR (%NR)	23.1 (100%)	27.9 (100%)
Risk of Bias	Moderately High		Moderately High			Moderately Low		Moderately High		Moderately High		Moderately High	

Gy = Gray; IMRT = Intensity Modulated Radiation Therapy; ITT = Intention to Treat; NOS = Not otherwise specified; NR = Not reported; PBS = Pencil Beam Scanning; PBT = Proton Beam Therapy; RCT = Randomized Control Trial; RT = Radiation Therapy; SCC = Squamous Cell Carcinoma

*This study conducted a propensity-score matched analysis (PBT, n=309; Photon, n=1541) using data from the National Cancer Database; demographics were not reported separately for the matched groups.

†To include, External Beam-Not Otherwise Specified (n=44,687), 3DCRT (n=36,406), Other Photons (n=140,035), and IMRT (n=22,346)

‡statistically significant difference.

§For the purposes of this review the non-randomized patients from Liao 2018 will be treated as an observational retrospective comparative cohort

Survival outcomes

Comparative studies

In the RCT,[161] no statistical differences were seen between the PBT and IMRT groups in the probability of OS at any timepoint up to 5 years according to the ITT analysis (Figure 20). Similarly, OS did not differ statistically by treatment type in the per-protocol analysis (Main Appendix I).

Four retrospective comparative cohort studies (to include the nonrandomized subgroup from the RCT) reported the probability of OS.[108,161,233,282] With the exception of the propensity-matched database study,[108] none of the studies reported statistically significant differences between PBT and photon RT over 1, 2 and 3 years (Figure 20). The database study conducted two separate propensity-score matched analyses. According to the 5:1 matching analysis (designed for better statistical power), PBT was associated with a statistically greater probability of OS over 5 years: HR 1.18, 95% CI 1.02 to 1.37. Results from the *a priori* 1:1 matched analysis showed a similar HR (1.16, 95% CI 0.97 to 1.39) but without statistical significance. Compared with the RCT, which tended to show a lower probability of survival following PBT, the observational studies all showed a tendency for higher survival with PBT.

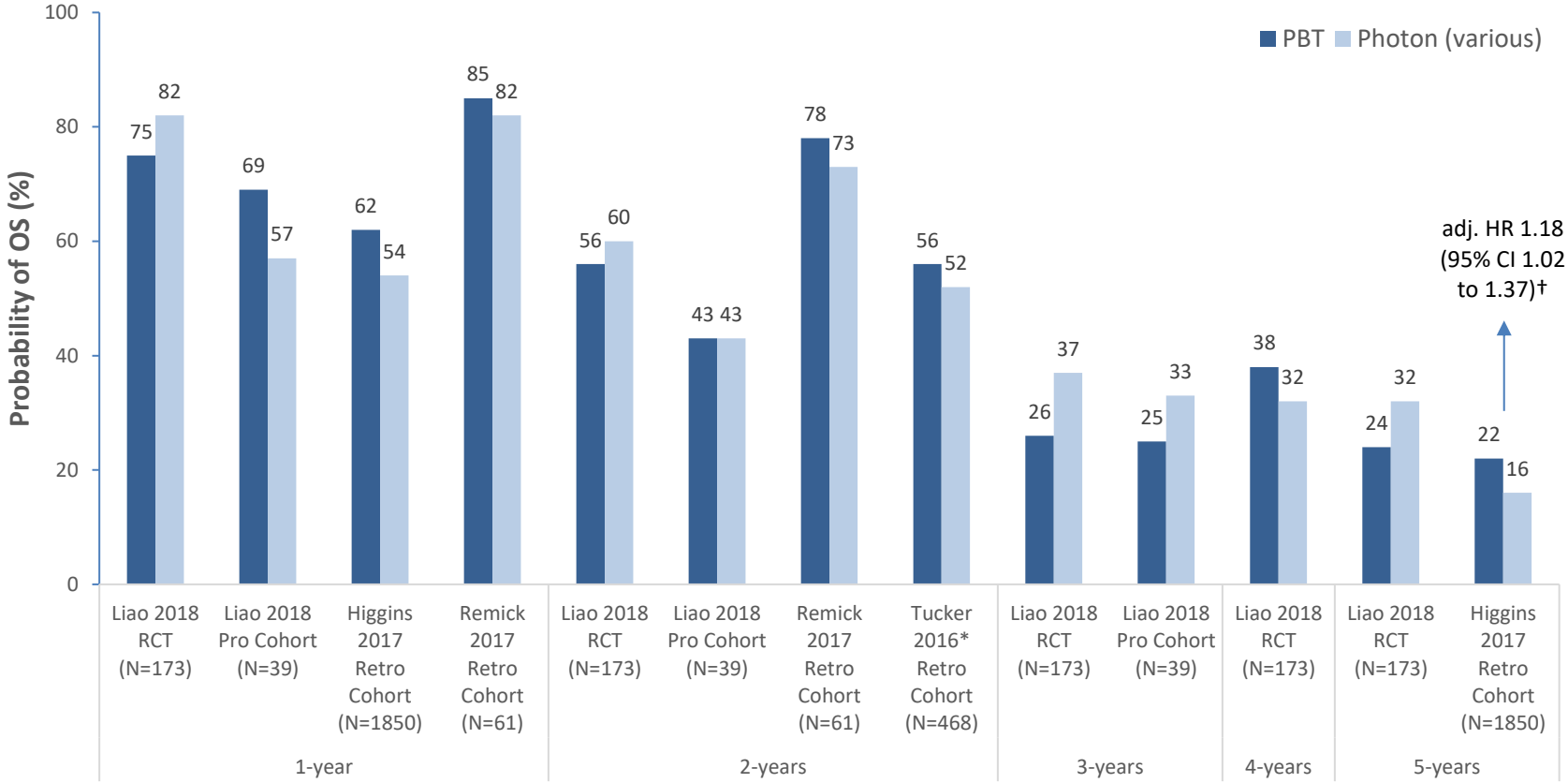
Case-series

The probability of OS was reported by nine case series, eight in NSCLC (N range, 35 to 74)[44,45,101,136,156,174,190,212] and one in limited stage small cell lung cancer (LS-SCLC) (N=30).[242] Across seven of the NSCLC studies reporting OS for all patients, the range of probabilities were as follows: 1-year, 86% to 97% and 2-year, 60% to 74% (2 studies)[45,212]; 3-year, 43% to 88% (6 studies)[45,101,136,156,174,212]; and 5-year, 28% to 66% (3 studies).[44,45,136] The eighth case series (N=506) reported 5-year OS for patients with NSCLC by stage only, probabilities were reported as follows: stage I, 36%; stage II, 34%, stage III, 23%; and stage IC 5%.[190] In the one small series of LS-SCLC, 1- and 2-year OS probabilities were 72% and 58%, respectively.

The probability of PFS was reported by six case series, five in NSCLC (N range, 35 to 74)[44,45,101,136,174] and one in LS-SCLC (N=30).[242] Across the NSCLC studies, the range of probabilities were as follows: 1- and 2-year, 80% and 64%, respectively (1 study)[45]; 3-year, 54% to 76% (4 studies)[45,101,136,174]; and 5-years, 22% to 54% (3 studies).[44,45,136]

For other outcomes reported across the case series, see Main Appendix F.

Figure 20. Probability of OS from One RCT and Four Retrospective Cohort Studies Comparing PBT versus Photon RT for Curative Intent in Adults with Lung cancer.



adj. = adjusted; CI = confidence interval; HR = hazard ratio; OS = overall survival; Pro = prospective; RCT = randomized controlled trial; Retro = retrospective; RT = radiation therapy.

*This study also compared PBT with 3DCRT; results for IMRT are reported here for consistency across studies. OS survival for 3DCRT was 39% (95% CI 32 to 46%), adj. HR 1.08 (95% CI 0.62 to 1.91), p=0.78.

†p=0.026 for the 5:1 propensity matched group [for the entire population, N=243822, 1.21 (1.06 to 1.39), p=0.005; Log-rank p<0.0001; for the 1:1 propensity matched group HR 1.16 (95% CI 0.97 to .39, p=0.12)].

Other outcomes

Comparative studies

In one RCT,[161] the cumulative incidences of both local failure and the composite outcome of local failure and radiation pneumonitis were similar following PBT versus IMRT over 5 years according to the ITT analysis (Table 35). Results according to the per-protocol analyses were similar (Main Appendix I).

Two comparative cohort studies (to include the nonrandomized subgroup from the RCT)[161,233] reported outcomes related to local, regional and distant control with no statistical differences seen between the PBT and IMRT groups in any outcome at any timepoint measured (Table 35).

Table 35. Outcomes related to tumor control in comparative studies evaluating PBT versus IMRT for curative intent in adults with lung cancer.

			Effect estimate (95% CI)	
Study	**Outcome**	**Time**	**PBT**	**IMRT**
Randomized controlled trial				
Liao (2018) N=173 (ITT) Moderately Low RoB	**Local Failure*** (cumulative incidence)	1-year	9%	10%
		2-year	27%	26%
		3-year	37%	37%
		4-year	37%	32%
		5-year	37%	39%
		Log-rank p-value	0.99	
	Combined rate of radiation pneumonitis and local failure* (cumulative incidence)	1-year	19%	19%
		2-year	36%	35%
		3-year	38%	36%
		4-year	38%	36%
		5-year	38%	36%
		Adj. HR (95% CI)	1.02 (0.53 to 1.98), p=0.94; Log-rank p=0.78	
Comparative observational studies				
Liao (2018) N=39 (from above RCT) Prospective cohort Moderately High RoB	**Local Failure** (cumulative incidence)	1-year	6%	3%
		2-year	6%	3%
		3-year	26%	26%
		Log-rank p-value	0.93	
Remick (2017) N=61 Retrospective Cohort Moderately High RoB	**Local-Recurrence-free survival,** probability	1-year	92.3% (82.5% to 100%)	93.3% (84.8% to 100%)
		2-year	93.1% (NR)	85.7% (NR)
	Local Failure, % (n/N)	2 years	11% (3/27)	6% (2/34)
	Regional Failure, % (n/N)	2 years	4% (1/27)	3% (1/34)
	Local and Regional Failure, % (n/N)	2 years	0% (0/27)	3% (1/34)
	Distant Failure (Metastasis), % (n/N)	2 years	41% (11/27)	50% (17/34)
		p-value	NS for all outcomes†	

adj. = adjusted; CI = confidence interval; HR = hazard ratio; IMRT = intensity-modulated radiation therapy; ITT = intention-to-treat analysis; NR = not reported; NS = not statistically significant; PBT = proton beam therapy; RCT = randomized controlled trial; RoB = risk of bias.
*cumulative incidences estimated from figure S3. The combined outcome of radiation pneumonitis and local failure was this trials primary outcome.
†With the exception of local-recurrence-free survival (log-rank p-value 0.82), statistical significance calculated by AAI.

Case series

The probability of 3-year local control was reported by five case series of NSCLC (N range, 35 to 74) and ranged from 82% to 96%.[101,136,156,174,212] In one small case series of LS-SCLC (N=30), 1- and 2-year probabilities of local control were 85% and 69%, respectively.[242]

Key Question 2 (Effectiveness, Salvage therapy)

One prospective comparative cohort[303] and one case series[46] was identified that evaluated PBT for salvage treatment of non-small cell lung cancer (NSCLC).

In the comparative cohort study, 82 patients (mean age 65 years; 50% male; 72% stage III) were treated with passive scatter PBT (n=26) or one of two photon therapies, IMRT (n=34) and 3DCRT (n=22); all patients received concurrent chemotherapy. The indication for treatment differed between the groups; PBT was given to patients with a recurrent tumor after surgery (42%) and/or chemotherapy while photon radiotherapy was given to those with non-operable, primary NSCLC. Because the proton patients were being treated for recurrent tumor this study is included under Key Question 2. Patients in the PBT group received a significantly higher radiation dose compared with the IMRT and 3DCRT groups: median 74.0 versus 63.0 Gy (RBE) (p<0.0001).

Briefly, the case series included 57 patients (median age 65 years; 44% male; 73% stage III/IV) undergoing re-irradiation (previous RT median 19 months prior) using either double-scatter or pencil beam scanning PBT (median dose 66.6 Gy [RBE]); 69% of patient received concurrent chemotherapy.[46]

Survival outcomes

Only the case series reported survival outcomes; the 1- and 2-year probabilities of overall survival were 59% and 43% and for progression-free survival, 58% and 38%, respectively.[46]

Other outcomes

The comparative cohort study examined patient-reported symptom burden according to the MD Anderson Symptom Inventory (MDASI) during the 7-week treatment period and for 5 weeks after the end of treatment with PBT versus photons (3DCRT or IMRT).[303] The authors chose to focus on "systemic symptoms" (i.e., the four most severe symptoms: fatigue, lack of appetite, disturbed sleep, and drowsiness) and "local symptoms" (i.e., pain due to esophagitis). After adjustment for patient and clinical variables, PBT was associated with less severe systemic and local symptoms during treatment compared with photons; however the difference was statistically significant only for local symptoms (i.e., pain). After treatment completion, PBT patients reported a statistically lower decrease in both systemic and local symptoms compared with photons.

In the case series, local/in-field recurrence was reported in 16% of patients, regional recurrence in 9% and distant metastases developed in 11%.[46]

Key Question 3 (Safety)

Curative Intent

Comparative studies

One RCT and two retrospective comparative cohort studies were identified that reported safety outcomes.[161,206,233]

Radiation pneumonitis was reported by two studies, with no statistically significant differences seen following PBT versus IMRT (Table 36). In the RCT,[161] the 5-year cumulative incidence of grade ≥3 radiation pneumonitis was 8.3% versus 5.9%, respectively (ITT analysis); all incidences occurred within 6 months. In one small retrospective comparative cohort study, the rate of acute grade 3 pneumonitis was 3.7% versus 2.9%, respectively.[233]

Acute radiation esophagitis was reported by two retrospective cohort studies, with no statistically significant differences seen following PBT versus IMRT (Table 36). The rate of grade 3 esophagitis, respectively, was 22.4% vs. 17.6% in one study[206] and 3.7% vs. 11.8% in the second, small study.[233]

No cases of acute grade 3 dermatitis following either PBT or IMRT were reported by one small retrospective cohort study.[233] This study reported a variety of other acute toxicities and found no statistically significant difference between groups for any event (Table 36).

Table 36. Toxicity outcomes from comparative studies evaluating PBT versus IMRT for curative intent in adults with lung cancer.

Outcome	Author, Year, N, Study Design	Timing	Grade	PBT	IMRT	Effect size (95% CI) P-value*
Radiation pneumonitis						
Radiation pneumonitis, cumulative incidence, % (n/N)	Liao (2018) N=173 (ITT) RCT Moderately low RoB	6 mos. to 5 years*	Grade ≥3	8.3% (6/72)†	5.9% (6/101)†	p=0.58*
Radiation Pneumonitis, % (n/N)	Remick (2017) N=61 Retro Cohort	Acute (NOS)	Grade 2	3.7% (1/27)	8.8% (3/34)	p=0.43
			Grade 3	3.7% (1/27)	2.9% (1/34)	p=0.87
Esophagitis, % (n/N)						
Esophagitis	Niedzielski (2017)‡ N=134 Retro Cohort	Acute (NOS)	Grade 0	18.4% (9/49)	28.2% (24/85)	p=0.20
			Grade 2	59.2% (29/49)	54.1% (46/85)	p=0.57
			Grade 3	22.4% (11/49)	17.6% (15/85)	OR 1.40 (0.69–2.87)*

Outcome	Author, Year, N, Study Design	Timing	Grade	PBT	IMRT	Effect size (95% CI) P-value*
Radiation pneumonitis						
Radiation Esophagitis: dysphagia and/or odynophagia, % (n/N)	Remick (2017) N=61 Retro Cohort	Acute (NOS)	Grade 2	18.5% (5/27)	29.4% (10/34)	p=0.33
			Grade 3	3.7% (1/27)	11.8% (4/34)	p=0.26
Dermatitis, % (n/N)						
Radiation Dermatitis	Remick (2017) N=61 Retro Cohort	Acute (NOS)	Grade 2	37% (10/27)	11.8% (4/34)	RR 3.1 (1.1–8.9)
			Grade 3	0% (0/27)	0% (0/34)	NA
Other Acute Toxicities, % (n/N)						
Hoarseness	Remick (2017) N=61 Retro Cohort	Acute (NOS)	Grade 2	0% (0/27)	2.9% (1/34)	p=0.37
			Grade 3	0% (0/27)	2.9% (1/34)	p=0.37
Cough			Grade 2	11.1% (3/27)	17.6% (6/34)	p=0.48
			Grade 3	0% (0/27)	0% (0/34)	NA
Dyspnea			Grade 2	18.5% (5/27)	14.7% (5/34)	p=0.69
			Grade 3	0% (0/27)	0% (0/34)	NA
Dyspepsia			Grade 2	11.1% (3/27)	23.5% (8/34)	p=0.21
			Grade 3	0% (0/27)	0% (0/34)	NA
Nausea			Grade 2	0% (0/27)	8.8% (3/34)	p=0.11
			Grade 3	0% (0/27)	2.9% (1/34)	p=0.37
Vomiting			Grade 2	0% (0/27)	2.9% (1/34)	p=0.37
			Grade 3	0% (0/27)	0% (0/34)	NA
Diarrhea			Grade 2	0% (0/27)	5.9% (2/34)	p=0.20
			Grade 3	0% (0/27)	0% (0/34)	NA
Constipation			Grade 2	3.7% (1/27)	14.7% (5/34)	p=0.16
			Grade 3	0% (0/27)	0% (0/34)	NA
Fatigue			Grade 2	22.2% (6/27)	26.5% (9/34)	p=0.70
			Grade 3	0% (0/27)	8.8% (3/34)	NR
Anorexia			Grade 2	22.2% (6/27)	17.6% (6/34)	p=0.66
			Grade 3	0% (0/27)	2.9% (1/34)	p=0.37
Dehydration			Grade 2	0% (0/27)	2.9% (1/34)	p=0.37
			Grade 3	0% (0/27)	2.9% (1/34)	p=0.37

CI = confidence interval; IMRT = intensity-modulated radiation therapy; NR = not reported; NOS = not otherwise specified; RCT = randomized controlled trial; PBT = proton beam therapy; Retro = retrospective.
*Except when indicated with an asterisks p-values or risk ratios were calculated by AAI since they were not provided by the authors.
†All cases occurred at ≤6 months; there were no other incidences after that time.

‡ This study also reported biomarkers for esophageal toxicity and found no statistically significant difference between groups for either measure [the maximum axial expansion of a single slice (MaxExp1); the axial length of the esophagus with at least 30% expansion (LenExp30%)].

Case series

Curative intent

Ten case series[44,45,101,123,136,156,174,205,212,242] (9 in NSCLC and 1 in LS-SCLC) were identified that reported safety outcomes for patients being treated with curative intent. Across the NSCLC studies, treatment-related mortality was low with only one event (1.8%) reported across six studies (N range, 35 to 64).[44,45,101,156,205,212] Grade ≥3 acute and late toxicities were reported by three studies (N range, 50 to 74) which ranged from 0% to 1.8% and from 0% to 17.6%, respectively. There were 11 cases (14.9%) of late grade 4 rib fracture in one of these series[136]; not counting those events the range of grade ≥3 late toxicities was 0% to 2.7% across these same three studies. Another study[44] reported the frequency of broad categories of grade ≥3 toxicities which ranged from 0% (cardiac events) to 22% (hematological events) over the acute term and from 3.1% (gastrointestinal events) to 22% (pulmonary events) over the later term. Four studies (N range, 35 to 55)[45,156,205,212] reported grade ≥3 toxicities but did not provide the timing of events; frequencies ranged from 1.8% to 12.7%. No cases of Grade ≥3 radiation necrosis were reported by one study (N=56).[174] In the one study evaluating LS-SCLC (N=30)[242], there were no treatment-related deaths and four grade ≥3 events, one case each (3.3%) of esophagitis, pneumonitis, anorexia, and pericardial effusion.

Salvage therapy

Only the one case series included for effectiveness reported safety outcomes.[46] Grade ≥3 toxicity occurred in 39% of patients during the acute period and 12% during the late period. Grade 4 toxicity occurred in four cases (3 neutropenia, 1 pericardial effusion) and grade 5 toxicity (i.e., death) in six cases (11%) (bronchopulmonary hemorrhage, neutropenic sepsis, anorexia, pneumonitis, hypoxic respiratory failure/pleural effusion, and tracheoesophageal fistula); the latter were deemed probably (3 cases) or possibly (3 cases) related to PBT.

Key Question 4 (Differential Efficacy/Effectiveness and Safety) and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.9 Lymphoma

Key Points

- There is insufficient evidence from three case series to evaluate the effectiveness and safety of PBT for curative intent in adults (primarily) with Hodgkin or non-Hodgkin lymphoma.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy, differential effectiveness and safety or cost-effectiveness.

Description of included studies

No comparative studies of PBT for the treatment of lymphoma that met inclusion criteria were identified.

Three retrospective case-series (high-risk of bias) that met inclusion criteria were identified that evaluated PBT for lymphoma (Abstraction Appendix J). Two studies by the same author group (with likely overlap in patient population) evaluated patients with Hodgkin lymphoma (HL) treated with chemotherapy and consolidative PBT with curative intent.[115,117] Both series included a mix of adult and pediatric patients however, the larger one (N=138)[115] reported results for the 59 pediatric patients separately (43% of the total population); results for this age group are reported separately in section 4.2.3. Adult patients (N=79) in this study received a median dose of 30.6 Gy(RBE) (range 20 to 45) delivered using primarily passive-scatter and uniform scanning as well as pencil beam scanning. Median follow-up was 2.7 years. The second smaller series (N=40),[117] an analysis of the Proton Collaborative Group Registry, did not report results separated by adults and pediatric patients though the majority (64%) were age 19 or older; the median age was 21 years. Patients in this study received a median dose of 30 Gy(RBE) (range 21–36 [RBE]) delivered using passive-scatter or uniform scanning; no patients were treated with pencil beam scanning. Median follow-up was 1.8 years. The third case series (N=59)[200] included equal proportions of both adults and pediatric patients (49% vs. 51, respectively) with HL (85%) or non-HL (15%) involving the thorax; results were not described separately for the different age groups. All patients had received chemotherapy along with PBT. Eleven patients (19%) were treated for relapsed or refractory disease (seven of these had stem cell transplantation); 66% (n=39) had bulky mediastinal disease. Median radiation dose was 30.6 Gy(RBE) (range 15–45 [RBE]); PBT technique was not reported. Median follow-up was 2 years.

Results

Key Question 1 (Effectiveness, curative intent)

Survival outcomes

Two studies, from the same author, reported on the probability of relapse-free survival (RFS) in patients with HL. In the larger case series (N=138), the 3-year probability of RFS for adults was 96% (97% for favorable early-stage, 93% for unfavorable early-stage, and 96% for advanced-stage disease).[115] In the smaller case series (N=40), the 2-year probability of RFS for a population of both adult (65%) and pediatric (35%) patients was 85%.[117]

Key Question 3 (Safety)

Across all three case series (primarily HL), no grade 3 toxicities were observed in any patient during follow-up (acute or late) and no clinically meaningful pneumonitis was reported,[115,117,200] Abstraction Appendix J.

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety), and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.10 Ocular Tumors

Key Points

- Across two retrospective cohort studies in patient with ocular tumors comparing PBT with brachytherapy or stereotactic radiosurgery (SRS) for curative intent, there were no statistically significant differences in OS at 2 years and mortality at 3 years; at 5-years PBT was associated with a statistically higher risk of mortality with PBT vs. brachytherapy in the larger, higher quality study (Low SOE).
- PBT was associated with a statistically lower frequency of local recurrence over 10 years compared with brachytherapy in one retrospective comparative cohort study (Low SOE). A second, poorer quality study comparing PBT versus stereotactic radiosurgery found no difference between groups in local recurrence at 3 years, however the strength of evidence was insufficient.
- With the exception of optic neuropathy which was statistically lower following PBT versus SRS in one study, no other statistical differences were seen in the frequency of adverse events (radiation retinopathy, enucleation, rubeosis of the iris, neovascular glaucoma, rubeotic glaucoma) over 3 years between PBT versus brachytherapy or SRS across two retrospective comparative cohort studies.
- One good quality (QHES 93/100) concluded that, compared to enucleation, PBT was not cost-effective for patients with intraocular melanoma using a WTP of $50,000/QALY based on a payer perspective. However, results ranged from cost-effective ($9,522/QALY) to very expensive ($441,750/QALY) in sensitivity analyses. PBT cost was a significant driver of the ICER.
- No studies meeting inclusion criteria were identified that evaluated salvage therapy (i.e., no comparative studies) or differential effectiveness and safety in this population.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of included studies

Three retrospective comparative cohort studies that compared PBT with other radiation therapies for curative intent in adult patients with primary uveal melanoma (2 studies)[35,258] or choroidal melanoma (1 study)[163] were identified (Table 37). In addition, 22 case series of PBT in adults with various ocular tumor types were identified; 21 evaluated PBT for curative intent[28,134,147,154,176,217,219,225,230,236,243,246,248,249,275-278,305,312,313] and one for salvage [235] therapy.

In addition, one cost-utility analysis (CUA)[192] that compared PBT with enucleation for treatment of intraocular melanoma that met inclusion criteria was identified.

Table 37. Ocular Tumors in Adults: Study Characteristics and Demographics for Studies Comparing PBT versus Photon RT for Curative Intent

	Effectiveness		Effectiveness & Safety			
	Lin 2017		Sikuade 2015		Boker 2018	
Characteristics	PBT (n=226)	Brachytherapy (n= 226)	PBT (n=106)	Stereotactic Radiosurgery (n=85)	Neoadjuvant PBT (+TSR) (n=70)	Adjuvant Brachytherapy (+TSR) (n=70)
Patient demographics						
Males, % (n)	54%	54%	59%	67%	47%	40%
Age, years; mean ± SD	60.6 ± 13.0	61.0 ± 13.5	57*	63*	57 ± 12	50 ± 12
Comorbidities, % (n)						
Charlson-Deyo score 0/1/≥2	82%/17%/1%	82%/15%/3%	---	---	---	---
Retinal Detachment	---	---	---	---	20%	30%
Ciliary Body Infiltration	---	---	---	---	81%	81%
Ciliary body or extraocular extension	8%	9%	---	---	---	---
Tumor characteristics						
Subtype	Choroid Melanoma		Uveal Melanoma		Uveal Melanoma	
Mean basal diameter (mm)	10.6 ± 4.3	9.9 ± 4.5	11.2	9.6	15.9 ± 2.6	15.7 ± 2.6
Mean distance of tumor from optic disc (mm)	---	---	2.9	2.2	9.3 ± 4.5	9.7 ± 4.5
Mean tumor thickness ± SD (mm)	5.5 ± 6.1	6.1 ± 10	4.3	3.9	10.4 ± 1.7	10.3 ± 1.8
Radiation Treatment						
Technique	---	---	---	Leksell Gamma Knife	---	20.0-mm Ru-106 plaque
Median total dose (Gy) (range)	56 (50-70.4)	---	58.4	35	mean 54.5	mean 470 (400-500)
Number of fractions	median 4 (4-7)	---	4	---	15	---
Number of sessions	---	---	---	1	4	---
Adjunctive/Concomitant treatment						
Chemotherapy or surgery	0%†		---	---	---	---
Study Design	**Retrospective propensity-score matched comparative cohort (National Cancer Database)**		**Retrospective comparative cohort**		**Retrospective matched-pairs comparative cohort**	
Follow-up, months (% followed)	**29 (NR)**	**37 (NR)**	**29 (NR)**	**27 (NR)**	**34.4 (NR)**	**39.8 (NR)**
Risk of bias	**Moderately High**		**Moderately High**		**Moderately High**	

KQ = Key Question; Gy = Gray; mm = millimeters; NR = Not Reported; PBT = proton beam therapy; Ru = ruthenium; SD = standard deviation; TSR = Trans-scleral resection.
*Indicates a statistically significant difference between groups.
†Receipt of surgery or chemotherapy was an exclusion criteria.

Results

Key Question 1 (Effectiveness, curative intent)

Across the three comparative studies evaluating PBT for curative intent, males comprised 44% to 63% of the populations and mean ages ranged from 54 to 61 years (Table 33). Only one study reported comorbidities (no major comorbidities in 83% [Charlson Deyo score 0]).[163] In another study,[35] 25% of patients presented with retinal detachment and 81% had ciliary body infiltration (not reported by other studies). The mean maximum basal diameter ranged from 10.3 to 15.8 mm and the mean tumor thickness from 4.1 to 10.4 mm across all studies. In the two studies evaluating uveal melanoma,[35,258] baseline patient and tumor characteristics were somewhat unbalanced between the treatment arms. None of the studies reported PBT technique. Two studies compared PBT (total doses 54.5 and 56 Gy) versus brachytherapy (one in uveal and one in choroidal melanoma)[35,163]; in one study all patients underwent transscleral resection (prior to PBT and following brachytherapy).[35] In the third study, patients received either PBT (58.4 Gy) or stereotactic radiosurgery via gamma knife (35 Gy).[258] One study excluded patients who received chemotherapy or surgery while the other two studies did not indicate whether or not patients had or were receiving chemotherapy or additional therapies. Two of these studies were case-matched analyses. One study conducted a propensity-score matched analysis of data from the National Cancer Database[163] and the other study case-matched pairs of patients treated at a single institution.[35]

Seventeen case series provided data on the effectiveness of PBT for curative intent (N range, 36 to 3088).[134,147,154,217,219,230,236,243,246,248,275-278,305,312,313] Tumor types included melanoma of the uvea, choroid, ciliary body, and iris; one study evaluated patients with uveal metastases [primarily from the breast (49%) and lung (22%)].[134] There is likely heavy overlap in populations across four studies in various subtypes of uveal melanomas[275-278]; determining the extent of overlap was difficult in these studies.

As described previously in section 4.1, all cohort studies were considered moderately high risk of bias (i.e., poor quality); however, two of the studies included here did control for confounding.[35,163] All case series were considered to be at high risk of bias.

Survival outcomes

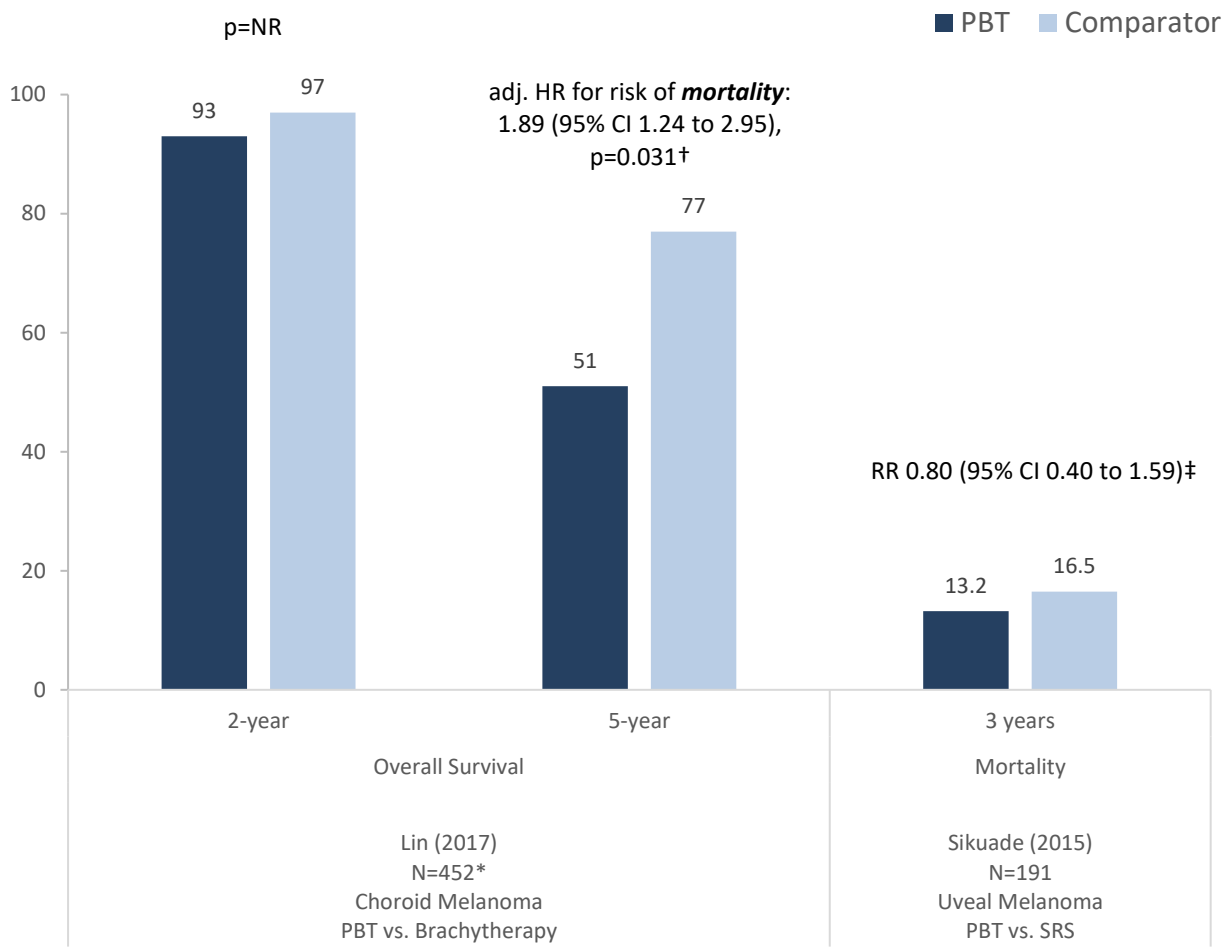
Comparative studies

Only the propensity-score matched database study, conducted in patients with choroidal melanoma, reported survival outcomes (

Figure 21). The probability of OS at 2 years was similar between groups but by 5 years PBT was associated with a statistically lower probability of OS compared brachytherapy: 51% versus 77% (adjusted HR for risk of mortality: 1.89, 95% CI 1.24 to 2.95).[163] Results for the whole, unmatched cohort (N=1224) were similar to that of the matched cohort (Abstraction Appendix M). A second study, comparing PBT with stereotactic radiotherapy in patients with uveal melanoma, found no statistical difference between groups in the risk of mortality over 3 years (OS not reported),

Figure 21.[258]

Figure 21. Probability of OS and Risk of Mortality in Retrospective Cohort Studies of PBT for Curative Intent versus Brachytherapy or Stereotactic Radiosurgery for Adults with Ocular Tumors.



*graph includes the propensity matched-case analysis; results for the whole, unmatched cohort (N=1224) were similar to that of the matched cohort.
†log rank p for OS = 0.008
‡Calculated by AAI. Effect estimate/p-value not provided by authors.

Case series

Five case series reported OS, three in uveal melanomas[28,275,276] and two in choroidal melanoma.[236,305] Across studies of uveal melanomas, 2-, 5-, 10-, and 15-year probabilities of OS were 95% (1 study), 74% to 87% (across 3 studies), 57% to 70% (across 2 studies), and 47% to 58% (across 2 studies), respectively. Across the studies of choroidal melanoma, corresponding probabilities were 91% (1 study), 77% and 94% (2 studies) and 63% (1 study), respectively. Mortality (as opposed to OS) was reported by six case series, four in choroidal melanoma[147,154,219,246] and two in melanoma of the iris.[230,278] Across the studies of choroidal melanoma, all-cause and disease-related mortality ranged from 16% to 48% (3 studies) and from 10% to 20% (4 studies), respectively, across follow-up periods of 30 to 148 months. In the two

studies of iris melanoma, all-cause mortality was similar over 50 months (6% and 7%); no deaths attributed to the disease were reported.

Visual acuity and other outcomes

Comparative studies

Two studies, both in patients with uveal melanoma, reported changes in visual acuity following radiation therapy with differing results, Table 38. PBT was associated with statistically better visual acuity compared with stereotactic radiotherapy at a median follow-up of 3 years in one study (55% vs. 33% of patients had a rating ≥6/60 on the Snellen scale)[258] while in the second study, a case-matched cohort, visual acuity was worse in those who had received PBT compared with brachytherapy (plus transscleral resection in both groups) at every timepoint measured (with the exception of 2 years) although no statistically significant difference was seen after 5 years.[35]

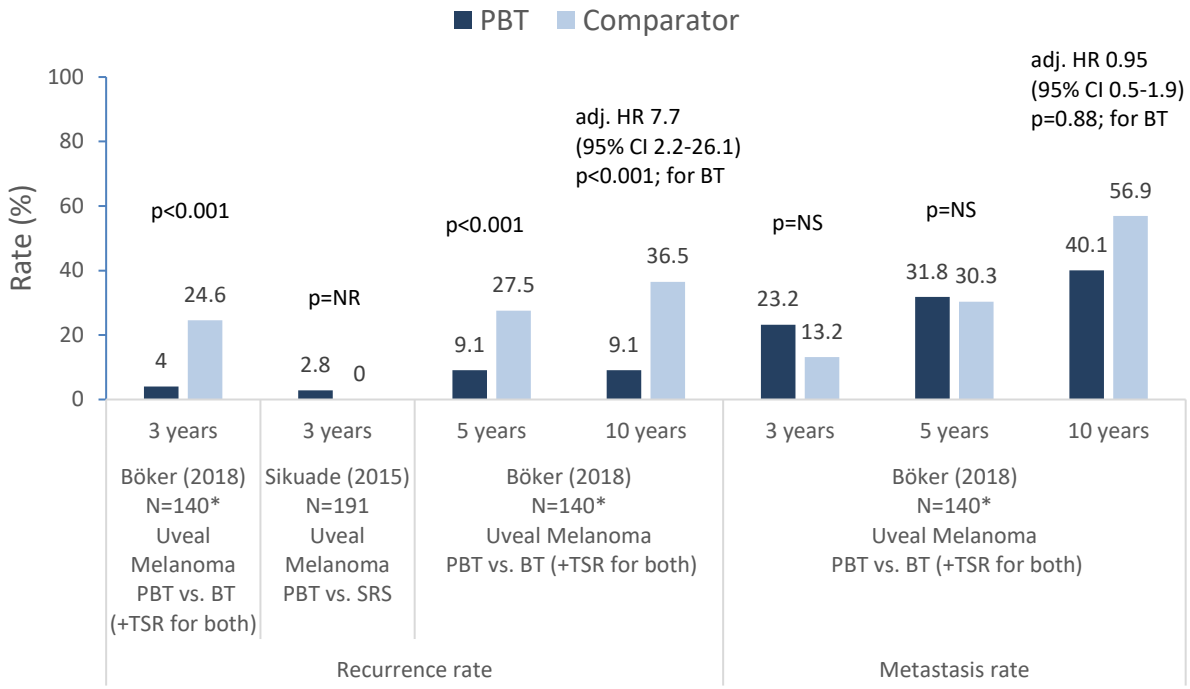
Table 38. Visual Acuity Outcomes from Retrospective Cohort Studies Comparing PBT for Curative Intent with Brachytherapy or Stereotactic Radiosurgery for Primary Ocular Tumors in Adults.

Author, Year, N, Design, Treatment groups	Tumor	Outcome	Time	PBT	Comparator	Effect size (95% CI) P-value
Sikuade (2015) N=191 Retrospective Cohort PBT vs. Stereotactic radiosurgery	Uveal melanoma	**Visual acuity ≥6/60, % (n/N)** [Snellen scale]*	Median f/u 3 years	55% (58/106)	33% (28/85)	RR 1.7 (1.2–2.4)†
		Significant Vision Loss, % (n/N) [loss of ≥3 Snellen lines]*	Median f/u 3 years	45% (48/106)	65% (55/85)	RR 0.7 (0.5–0.9)†
Böker (2018) N=140 Retrospective Case-Matched Cohort PBT vs. Brachytherapy	Large uveal melanoma	**Visual Acuity, median (IQR)** [logMAR; higher score = worse vision]‡	Baseline	0.4 (0.2 to 0.7)	0.3 (0.1 to 0.7)	p=0.03
			1-year	0.8 (0.5 to 1.3)	1.5 (1 to 2)	p<0.001
			2-year	1.2 (0.8 to 1.5)	0.8 (0.4 to 1.2)	p<0.001
			3-year	NR	NR	p=0.007§
			4-year	NR	NR	p=0.036§
			5-year	NR	NR	p=0.011§
			6-year	NR	NR	p=0.074§
			7-year	NR	NR	P=0.412§

CI = confidence interval; f/u = follow-up; IQR = interquartile range; PBT = proton beam therapy; RR = risk ratio.
*A Snellen test consists of a number of rows of letters which get smaller as you read down the chart. Normal visual acuity on this scale is called 6/6, which corresponds to the bottom or second bottom line of the chart. A score of 6/60 (i.e., can only read the top line of the chart) means that a person can see at 6 meters what someone with standard vision could see from 60 meters away.
†Calculated by AAI. Effect estimate/p-value not provided by authors.
‡6/6 Snellen (normal vision) = 0 logMAR; 6/60 Snellen = logMAR 1.0.
§Other than at 1-year, visual acuity was worse in the PBT group compared with the brachytherapy group although no significant difference was noted after 5 years.

These same two studies also reported tumor recurrence rates. In the single center, case-matched analysis, compared with brachytherapy, PBT was associated with statistically lower rates of local recurrence at 3, 5 and 10 years (Figure 20); the rate at 10 years was 9.1% versus 36.5% (adjusted HR for brachytherapy: 7.7, 95% CI 2.2 to 26.1).[35] Of note, all patients in this study also underwent transscleral resection. Conversely, the second study reported no statistical differences between groups at 3 years[258]; the tumor recurred in three (2.8%) PBT patients (all underwent secondary enucleation) compared with none of the patients who received stereotactic radiotherapy (Figure 20). The case-matched analysis also reported the rate of distant metastases with no statistical difference between groups at any timepoint measured (Figure 22).

Figure 22. Tumor Recurrence and Metastasis Rates in Retrospective Cohort Studies of PBT for Curative Intent versus Brachytherapy or Stereotactic Radiosurgery for Adults with Ocular Tumors.



adj. = adjusted; BT = brachytherapy; CI = confidence interval; HR = hazard ratio; NS = not statistically significant; PBT = proton beam therapy; SRS = stereotactic radiosurgery; TSR = trans-scleral resection.
*Matched-case analysis; 95% CIs for rates for PBT vs. Brachytherapy, respectively, in Boker were:
- Recurrence rate: 3-year: 4% (1.2% to 17.8%) vs. 24.6% (15.8% to 37.1%); 5-year: 9.1% (2.9% to 27.3%) vs. 27.5% (17.8% to 41.1%); 10-year: 9.1% (2.8% to 27.3%) (3/70) vs. 36.5% (20.7% to 59.1%) (18/70)
- Metastasis rate: 3-year: 23.2% (5.6% to 37.1%) vs. 13.2% (6.8% to 24.9%); 5-year: 31.8% (20.7% to 46.8%) vs. 30.3% (18.3% to 47.5%); 10-year: 40.1% (26.6% to 58.6%) (19/70) vs. 56.9% (34.9% to 80.8%) (18/70)

Case series

The probability of local control following PBT, as reported by four studies, was relatively high regardless of the type of ocular tumor evaluated (uveal or choroidal): 2-years (99% in one study),[305] 5-years (85% to 96% across 4 studies) [28,246,305,312] and 10-years (85% to 96% across 3 studies).[28,246,305] Similarly, the probability of local recurrence/relapse was relatively low as reported by two studies of uveal

melanomas: 1-year (2% in one study),[217] 3 to 4-years (5% and 6% in both studies),[217,277] and 5- (8%) and 10- (13%) years in one study.[217] In the case series of uveal metastases, the probability of local recurrence was 8% at 1 year.[134] Additionally, five other case series reported low rates of local or regional recurrence regardless of tumor type: 3.2% to 5.7% across follow-up periods ranging from 30 to 77 months.

Five case series reported the probability of metastasis-free survival following PBT. Two case series of uveal melanomas reported the probability of metastasis-free survival as follows: 5-year, 74% to 96%; 10-year, 66% to 70%; 15-year, 55% to 58%.[28,275] Across the three studies evaluating patients with choroidal melanoma, respective probabilities were 90% in one study,[305] 72% to 90% across three studies[236,246,305] and 57% to 82% across two studies.[246,305]

Key Question 2 (Effectiveness, salvage therapy)

No comparative studies were identified for salvage PBT in adults with ocular tumors. One case series (N=48) of salvage PBT for local recurrent choroidal melanoma was included.[235] Mean patient age was 61 years; patient sex was not reported. Previous treatments (and combinations thereof) included brachytherapy, transpupillary thermotherapy, photodynamic therapy, CyberKnife therapy, or PBT and the median interval between primary treatment and PBT as salvage therapy was 17.6 months. Mean follow-up time was 81 months.

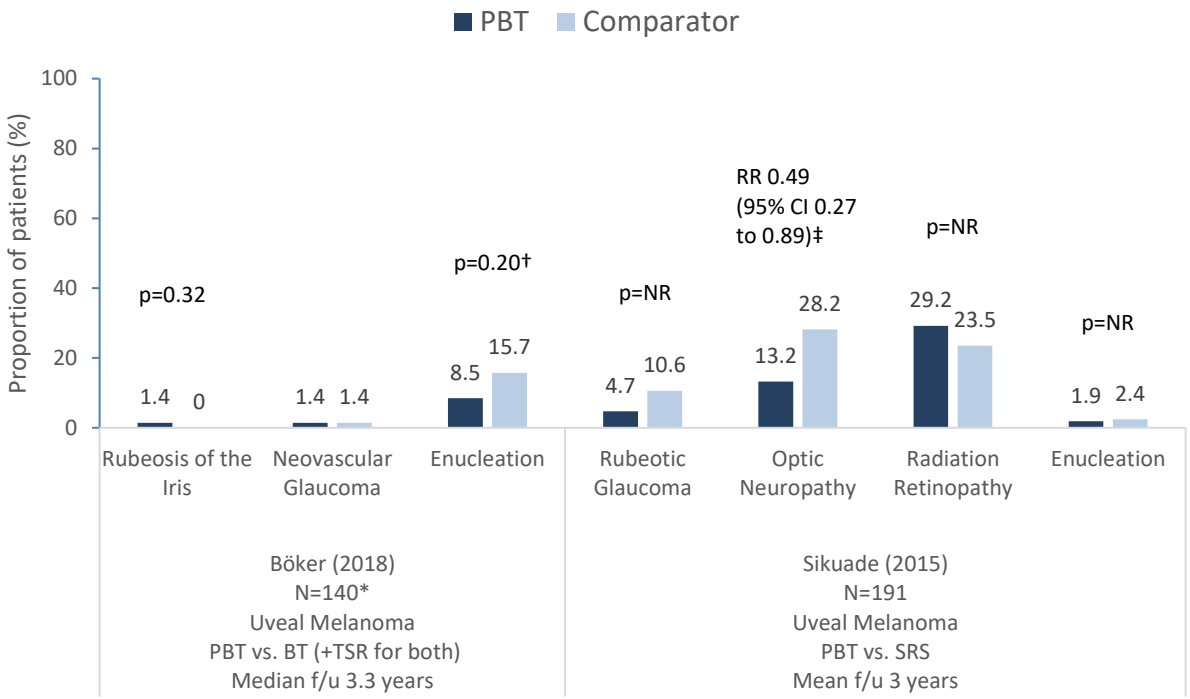
The 5- and 10-year probabilities of OS were 89.1% and 77.4%, respectively.[235] For metastasis-free survival the respective probabilities were 80.7% and 70.1%. Local control was achieved in 92.1% of patients at 10 years post-salvage PBT. Enucleation was indicated in two (4%) of the three patients with local recurrence. Regarding visual acuity, the 5-year probability of vision worse than 20/200 on the Snellen scale was 24%, with the sharpest decrease seen within the first 2 years post-PBT. Two patients (4%) had no light perception.

Key Question 3 (Safety)

Comparative studies

Two retrospective comparative cohort studies reported safety outcomes following curative PBT versus brachytherapy (1 study; all patients also underwent transscleral resection)[35] or stereotactic radiotherapy (1 study)[258] over 3 years of follow-up in patients with primary uveal melanoma. No statistical differences were seen in the incidence of any adverse event in either study, to include need for subsequent enucleation, radiation retinopathy, and glaucoma, Figure 23.

Figure 23. Adverse Events in Retrospective Cohort Studies of PBT for Curative Intent versus Brachytherapy or Stereotactic Radiosurgery for Adults with Ocular Tumors.



BT = brachytherapy; CI = confidence interval; NS = not statistically significant; PBT = proton beam therapy; RR = risk ratio; SRS = stereotactic radiosurgery; TSR = trans-scleral resection.
*matched-case analysis
†proportion for this outcomes is out of eyes (as opposed to patients)
‡Calculated by AAI

Case series

All but one of the included case series evaluating PBT for curative intent reported on safety. For a complete list of all safety outcomes reported please see Tables F51 and F52 in Main Appendix F. The rate of subsequent enucleation ranged from 3% to 16% across seven studies of choroidal melanoma (follow-up range, 30 to 70 months)[147,219,236,246,249,305,313]; from 3% to 9% across three studies of uveal melanoma (follow-up range, 44 to 53 months)[225,276,277]; and from 0% to 6% across three studies of iris melanoma (follow-up range, 50 to 55 months).[230,243,312] One patient (1.3%) in the study of uveal metastases required enucleation over a median of 7.7 months. The incidence of neovascular glaucoma varied widely across the five studies evaluating choroidal melanoma (range 2% to 23% across 30 to 51 months of follow-up)[147,248,249,305,313]; in two studies of uveal melanoma the incidences were similar (18% and 25% across 69 to 84 months of follow-up)[217,275] and there were no cases of neovascular glaucoma in one small case series in patients with iris melanoma followed for 36 months. The frequencies of maculopathy and neuropathy post-PBT were high and varied across studies. For maculopathy, frequencies ranged from 24% to 49% across two studies evaluating choroidal melanoma[147,219] and 7.2% to 34% across two studies of uveal melanomas.[275,276] Similarly, the frequencies of neuropathy ranged from 24% to 55% across three choroidal melanoma studies,[147,248,249] from 7.5% to 47.5% across two uveal melanoma studies,[275,276] and was 4.7% in one iris melanoma study.[278] Scleral necrosis was rare and occurred in 0% to 0.9% of patients across four studies of chorodial or iris melanoma.[230,249,278,312]

The one case series of salvage PBT for recurrent choroidal melanoma reported that 21% of patients required cataract surgery following PBT, 6% (n=3) of which were due to secondary complications from PBT (the other 15% were due to refractive indications only).[235] Vitrectomy was performed in 8% of patients for complications due to PBT; three patients (6%) had vitreous bleeding (at 15, 30 and 74 months) and one (2%) had exudative retinal detachment 0.9 months post-PBT.

Key Question 5 (Economic)

One good quality CUA (QHES 93/100) compared PBT with enucleation for treatment of intraocular melanoma using Markov modeling of a hypothetical cohort (Table 39). Authors received no funding for this study. Incremental cost-effectiveness ratios (ICER) were reported to reflect the marginal cost of adding one quality adjusted life year (QALY) to a patient's life when using PBT versus IMRT. The primary limitations of this study are based on generalization of some model parameters from technologies or populations not specific to the use of PBT in patients with intraocular melanoma, e.g. the use of utilities from a general population of patients with melanoma and generalization of progression risk from brachytherapy to the other treatments and generalization of some costs across treatments.

Key points

One good quality CEA concluded that, compared to enucleation, PBT was not cost-effective for patients with intraocular melanoma using a WTP of $50,000/QALY based on a payer perspective.[192] However, results ranged from cost-effective ($9,522/QALY) to very expensive ($441,750/QALY) in sensitivity analyses. PBT cost was a significant driver of the ICER.

- Base case ICER: $106,100/QALY
- Sensitivity analysis: ICER range $9,522/QALY to $441,750/QALY); the model was not robust to sensitivity analyses.
- Limitations:
  - Relative risk for progression from local recurrence to distant metastasis was derived from study using plaque brachytherapy; the extent to which this applies to other therapies is unclear.
  - Health state utilities not specific to study population; utilities used were from a study of general melanoma
  - Evidence for cost assumptions not well documented (same cost for cancer recurrence for all treatment therapies; cost of radiotherapy could not be identified so cost of enucleation was used in its place; no cost specific to distant metastasis was modeled)
  - Costs for treatment complications were not included.

Detailed results

Study characteristics and framework

One good quality CEA (QHES 93/100) evaluated the cost-effectiveness of PBT compared to enucleation for 59 year-old patients with intraocular melanoma.[192] The costing year was 2011. The time horizon was 5 years, which may be reasonable based on 5-year survival rates for melanoma that has metastasized. The study adopted a provider perspective but did not include costs of treatment complications. Costs included treatment, local recurrence, and end-of-life costs. The costs for PBT and enucleation were $12,438 and $8,678, respectively. Clinical data on probability of tumor recurrence following PBT were

derived from cohort studies of patients with intraocular melanoma at two cancer centers. [60,65]. For enucleation, probabilities came from a study that used a Markov model to estimate the cost-utility of high-dose interferon alpha therapy in stage III cutaneous melanoma[56] and follow-up from a multicenter randomized trial of pre-enucleation compared to enucleation for patients with ocular melanoma[102].

Base Case Results

Using the provider perspective, PBT and enucleation were found to cost $24,894 and $22,772, respectively. PBT resulted in 2.938 QALYs while enucleation resulted in 2.918 QALYs. The ICER was $106,100/QALY.

Sensitivity Analyses

A one-way sensitivity analysis was conducted

In one-way sensitivity analyses, results were very sensitive to 13 model parameters: probability of local recurrence for all 3 therapies, end-of-life costs for disease, treatment costs for all 3 therapies, and post-treatment utility for all 3 therapies. Low values of model parameters resulted in ICERs for PBT versus enucleation ranging from $9,543/QALY to $234,862/QALY. High values of model parameters were associated with greater variability, resulting in ICERs ranging from $9,522/QALY to $441,750/QALY. Across low and high values for model parameters, PBT dominated enucleation in 4 instances (high estimate of probability of distant metastasis for patients treated with enucleation, low estimate of probability of distant metastasis for patients treated with PBT, high estimate of enucleation cost, low estimate of PBT cost) and was dominated by enucleation in 5 instances (high estimate of probability of local recurrence for patients treated with PBT, high estimate of probability of distant metastasis for patients treated with PBT, low estimate of probability of distant metastasis for patients treated with enucleation, high estimate of post-treatment utility for enucleation, low estimate of post-treatment utility for PBT).

Conclusions and Limitations

The authors concluded that PBT is not cost-effective using a WTP threshold of $50,000/QALY. However, results were not robust to sensitivity analyses and showed that decreased payment rates for PBT could result in PBT dominating enucleation.

The main limitation of this study is that some model parameters were derived from suboptimal sources (case series, studies of treatments other than PBT). For example, the relative risk for progression from local recurrence to distant metastasis was derived from a study using plaque brachytherapy and may not apply to PBT. In addition, the health state utilities were drawn from a study of general melanoma, not specific to this study population. The authors also made assumptions about costs that do not appear to be supported by evidence. For example, there was no inclusion of cost specific to distant metastasis and the cost of radiotherapy could not be identified so cost of enucleation was used in its place. All of these assumptions could substantially impact the ICER. PSA was not done. The QHES score for this study was 93/100 points.

Table 39. Summary of the economic study comparing PBT with enucleation in patients with intraocular melanoma

	Moriarty 2015
Population	59 years of age with intraocular melanoma
Intervention(s)	PBT (timing unclear)
Comparator(s)	Enucleation
Country	USA
Funding	None
Study design	CUA
Perspective	Provider
Time horizon	5 years
Analytic model	Markov model with 5 health states
Effectiveness outcome	QALYs
Effectiveness outcome components	QOL post-treatment, QOL w/ local recurrence, QOL with metastasis
Source for effectiveness data	Prior literature; data from systematic reviews
Costing year	2011
Currency	USD
Discounting	3%
Components of cost data	Treatment cost, local recurrence, end-of-life costs (disease or other causes)
Cost sources	Publicly available databases (Medicare reimbursement rates, 2010 Nationwide Inpatient Sample database, Healthcare Cost and Utilization Project); End-of-life costs derived from prior literature (case series)
Sensitivity analysis	One-way
QHES	88
Results:	
Cost / QALY of intervention	$24,894/2.938 = $8,473/QALY
Cost / QALY of comparator(s)	$22,772/2.918 = $7,804/QALY
ICER	$106,100/QALY
One-way SA	Model was sensitive to 13 parameters: probability of local recurrence for all 3 therapies, end-of-life costs for disease, treatment costs for all 3 therapies, and post-treatment utility for all 3 therapies

	Moriarty 2015
	ICER range for low parameter values: $9,543/QALY to $234,683/QALY ICER range for high parameter values: $9,522/QALY to $441,750/QALY
Other SA	Not done
Author's Conclusion	PBT was not cost-effective compared to enucleation using a WTP of $50,000/QALY; Results were not robust to sensitivity analyses and showed that decreased payment rates for PBT could be result in PBT being dominant over enucleation
Limitations	• Relative risk for progression from local recurrence to distant metastasis derived from study using plaque brachytherapy; may not apply to other treatment strategies • No costs of treatment complications • QOL data derived from study of general melanoma (not specific to this population) • Strong assumptions about costs (same costs for recurrence for all treatment therapies; cost of radiotherapy substituted with cost of enucleation; no cost specific to distant metastasis)

CUA = Cost Utility Analysis; ICER = Incremental cost effectiveness ratio; PBT = Proton Beam Therapy; QALY = Quality Adjusted Life Year; QHES = Quality of Health Economic Studies; QOL = Quality of Life; SA = Sensitivity Analysis; USD = United States Dollar

Key Question 4 (Differential Effectiveness and Safety)

No studies that met inclusion criteria were identified.

4.3.11 Prostate Cancer

Key Points

- In one quasi-RCT, there were no statistically significant differences in the probabilities of 5- and 10-year overall survival and biochemical relapse-free survival between the combined photon and PBT boost group and the photon only group (Low SOE).
- The probabilities of acute and late grade 2 gastrointestinal (GI), but not genitourinary (GU), toxicity were significantly lower in patients who received the photons plus PBT boost versus photons only in one quasi-RCT; however, there were no statistically significant differences for grade 3 or 4 toxicities. Across three retrospective cohort studies comparing PBT with IMRT results regarding acute and late GU and GU toxicity differed, with two finding no statistical difference between groups and the third, a large database study, reporting lower cumulative incidences with PBT (to include erectile dysfunction) compared with IMRT; differences between groups were small and clinical significance is unknown (SOE Low for all).
- No studies that met inclusion criteria were identified that provided data on PBT for salvage therapy, differential effectiveness and safety or cost-effectiveness.
- Limited information from case series does not provide sufficient information to evaluate radiation safety or effectiveness of PBT.

Description of included studies

One quasi-RCT[139] and three retrospective comparative cohort studies[69,76,216] that compared PBT with photon radiation therapies (RT) for curative intent in adult men with locally advanced prostate cancer were identified (Table 40); two[76,216] of the three cohort studies reported on safety only and will be described further in Key Question 3 below. In addition, 11 case series (across 12 publications) of PBT for curative intent were identified.[20,39,50,53,110,116,124,172,181,227,270,285]

Table 40. Prostate Tumors in Adults: Study Characteristics and Demographics for Studies Comparing PBT versus Photon RT for Curative Intent

Author (year)	Effectiveness and Safety				Safety Only			
	Dutz 2019		Khmelevsky 2018		Fang 2015		Pan 2018	
Characteristics	PBT (n=29)	Photon RT (n=29)	Photon RT + PBT Boost (n=116)	Photon RT (n=173)	PBT (n=94)	Photon RT (n=94)	PBT (n=693)	Photon RT (n=3465)
Patient demographics								
Males, %	100%	100%	100%	100%	100%	100%	100%	100%
Mean Age ± SD (years)	Median Age (range): 70.4 (49.3 to 83.6)**	Median Age (range): 74.9 (65.9 to 83.8)**	66.9 ± 6.4	69.0 ± 5.8	60 to 69 years: 50%	60 to 69 years: 47%	≤55: 29% 56-60: 39% 61-64: 32%	≤55:29% 56-60: 39% 61-64: 33%
Comorbidities*								
Any	---	---	---	---	---	---	13%	11%
Hypertension	---	---	---	---	46%*	67%*	---	---
Diabetes	27.6%	24.1%	---	---	---	---	---	---
Tumor characteristics, % (n)								
Risk Level								
Low	6.9%	0%	7.0%	3.8%	55%	55%	---	---
Intermediate	75.9%	79.3%	36.0%	46.5%	31%	31%	---	---
High	17.2%	20.7%	57.0%	49.7%	7%	7%	---	---
Radiation Treatment								
Technique	Conventionally Fractionated	IMRT	Standard Conformal PBT Boost: NR	Standard Conformal	Passive Scatter	IMRT	---	IMRT
Mean total dose (Gy)	74	78	71.8 (to prostate)† 44.9 (to small pelvis)	68.6 (to prostate)† 44.8 (to small pelvis)	79.2	---	---	---
Number of fractions	---	---	Photon: 22-23 PBT Boost: NR	Photon: 22	44	---	Median: 39	Median: 42
Additional Treatments Prior to Radiation								
Transurethral Resection	6.9%	3.4%	14%	17%	---	---	---	---
Adenomectomy	---	---	6%	9%	---	---	---	---

Final

Author (year)	Effectiveness and Safety				Safety Only			
	Dutz 2019		Khmelevsky 2018		Fang 2015		Pan 2018	
Characteristics	PBT (n=29)	Photon RT (n=29)	Photon RT + PBT Boost (n=116)	Photon RT (n=173)	PBT (n=94)	Photon RT (n=94)	PBT (n=693)	Photon RT (n=3465)
Cystectomy	---	---	3%	5%	---	---	---	---
ADT	44.8%	44.8%	95%	95%	16%	29%	19%	19%
Anticoagulants	31%	37.9%	---	---	---	---	---	---
Study Design	**Retrospective Propensity score Matched Comparative Cohort**		**Quasi-Randomized Controlled Trial**		**Retrospective Matched Pairs Comparative Cohort‡**		**Retrospective Propensity Matched Comparative Cohort§**	
F/U, months (% followed)	**NR (NR)**		**67.8 (94.1%)**		**29 (100%)**	**47 (100%)**	**23 (NR)**	**23 (NR)**
Risk of bias	**Moderately High**		**Moderately High**		**Moderately High**		**Moderately High**	

ADT = Androgen Deprivation Therapy; F/U = follow-up; GI = Gastrointestinal; GU = genitourinary; Gy = Gray; IMRT = Intensity Modulated Radiation Therapy; KQ = Key Question; NR = not reported; PBT = Proton Beam Therapy; RT = Radiation therapy; SD = Standard Deviation

*Comorbidities were poorly reported by the studies. Khmelevsky et al. 2018 did not report any comorbidities. Fang et al. 2015 also reports the following comorbidities in the PBT and IMRT groups, respectively: Hemorrhoids (14% vs. 10%); Diabetes mellitus (14% vs. 23%); Prior GI disorders (12% vs. 15%); Prior GU disorders (16% vs. 22%).

†Indicates a statistical difference between groups

‡Matched for risk group, age, and prior GU and GU disorders; both exact matching (risk group) and nearest-neighbor matching (age, prior GI/GI disorders).

§Matched for age, residence type, median household income, geographic region, treatment year, employee relation, capitated insurance plan, medical comorbidity, baseline GU/bowel comorbidity, and concurrent ADT; patients were matched using a greedy algorithm and a maximum allowed caliper distance of 0.1.

**Indicates a statistically significant difference between the groups

Results

Key Question 1 (Effectiveness, curative intent)

The quasi-RCT compared patients who underwent standard conformal photon therapy with (n=116) and without (n=173) a PBT boost.[139] The study design was considered quasi-randomized because allocation to the groups was performed according to time of presentation/arrival time for treatment. Mean patient age was 68 years and the majority of patients were considered either high (53%) or moderate (42%) risk for disease progression. Almost all patients (95%) had received 3 to 12 months of androgen deprivation therapy (ADT) prior to starting RT. In the PBT boost group, all patients received the same total mean dose of radiation (photon + proton) to the prostate (71.8 Gy); however, patients were divided up sequentially to receive three variants of proton boost fractionation: 3.0 Gy in 8 daily fractions (n=46), 4.0 Gy in 5 fractions, 3 or 5 fractions/week (n=44), and 5.5 Gy in 3 fractions, 3 fractions/week (n=24). Mean total radiation dose for the photon only group was 68.6 Gy ($p<0.01$ compared with the combined proton/PBT group). Similar proportions of patients in both groups had undergone previous urinary tract surgeries and therefore did not received additional radiation to the small pelvis.

One retrospective, propensity score-matched comparative cohort study (N=58)[69] reported data on the effectiveness of passive scatter PBT compared with intensity modulated radiation therapy (IMRT) for primarily intermediate-risk (78%) prostate cancer. The included patients were from a single institution and matched based on risk group, transurethral resection of the prostate (TURP), prostate volume, diabetes mellitus, and anticoagulation use (all potential confounder for the occurrence of toxicities according to authors). The median age of the entire cohort was 73 years and patients who received PBT were statistically younger than those who received IMRT ($p=0.001$) (Table 36). Almost half (45%) of the patients had been treated with ADT and only 5% had undergone TURP prior to study entry. All patients were treated with conventionally fractionated radiation therapy; the prescribed total radiation dose was significantly lower in the PBT group [74 Gy(RBE) vs. 78 Gy in the IMRT group, $p<0.001$). The median length of follow-up was not reported.

Additionally, nine case series (across 10 publications) provided data on the effectiveness of curative PBT for prostate cancer (N range, 49 to 1375).[20,39,53,110,124,172,181,227,270,285]

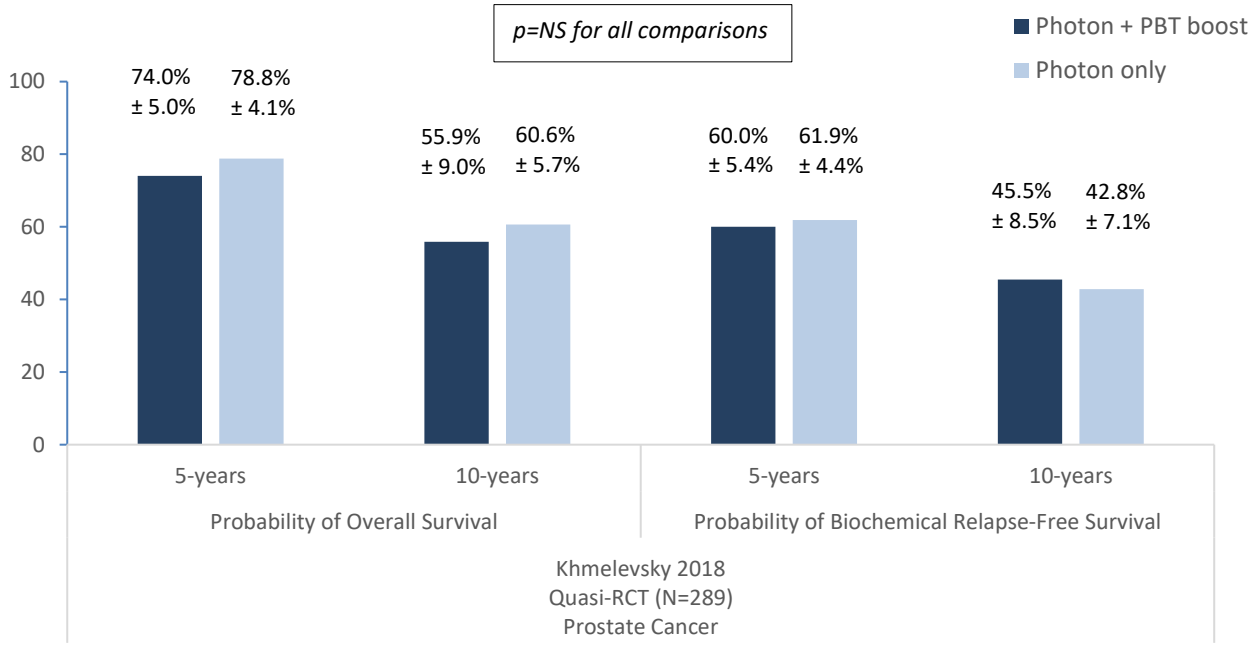
For the reason stated in Section 4.1, the quasi-RCT and retrospective comparative studies (two reported on safety only and are described further below) were considered to be at moderately high risk of bias (i.e., poor quality); however, all four studies controlled for confounding All case series were considered to be at high risk of bias.

Survival outcomes

Comparative studies

As reported by one quasi-RCT,[139] the probability of 5- and 10-year overall survival did not differ statistically between the combined photon and PBT boost group and the photon only group; likewise, the probability of 5- and 10-year biochemical relapse-free survival was similar between groups (Figure 24).

Figure 24. Probability of OS and BRFS in a Quasi-RCT Comparing Photons + PBT Boost with Photons alone for Curative Intent for Adults with Prostate Cancer



± standard deviation
BRFS = Biochemical Relapse-Free Survival; NS = not statistically significant; OS = overall survival; PBT = proton beam therapy; RCT = randomized controlled trial

Case series

Across case series, 5-year overall survival following PBT was 93% to 98% in low risk patients (4 studies),[39,53,124,181,270] 88% to 97% in intermediate risk patients (5 studies),[20,39,53,124,181,270] 86% to 98% in high risk patients (5 studies)[20,39,53,124,181,270] and 90% in very high risk patients in one study[270] In one case series of patients at low or intermediate risk, the probability of 7-year OS was 99%. In another study, 8-year OS for those at low, intermediate, high or very high risk was 94%, 90%, 89% and 86%, respectively.[270] Three case series did not report OS survival but did report the incidence of mortality across all patients[172,227,285]; all-cause mortality ranged from 0% to 1% (2 deaths total) and there were no disease-related deaths reported.

Two case series reported the 5-year probability of progression-free or clinical relapse-free survival.[20,124] In patients at low risk the probability was 100% as reported by one study[124]; across both series, probabilities ranged from 97% to 98% for intermediate risk and 83% to 96% for high risk patients. Two other studies reported low rates (1%) of local or regional recurrence or relapse across all patients.[227,270]

Distant metastasis-free and nodal metastasis free-survival probabilities at 5 years were reported by one case series for patients at low (99% for both), intermediate (99% for both) and high (98% and 96%, respectively) risk of progression.[39,53] Four other studies reported low incidences of distant metastases in all patients (range, 0.5% to 2.9%) over mean follow-up periods ranging from 52 to 70 months.[20,124,227,270]

The probability of freedom from biochemical failure was also reported by a number of the case series and can be found in Main Appendix F Tables F53, F54, and F55.

Quality of Life

One retrospective propensity score-matched cohort study[69] evaluated quality of life according to two European organization for research and treatment of cancer (EORTC) questionnaires: the EORTC QLQ-C30 (general quality of life) and QLQ-PR25 (prostate-cancer specific) questionnaires. The only statistically significant differences seen between patients who received PBT compared with IMRT were for two subscale scores on the EORTC QLQ-C30: constipation at 3 months post-radiation which favored PBT (mean change from baseline: -6.7 ± 13.8 vs. 6.7 ± 22.5, respectively, p=0.03) and the global health status subscale score 12 months post-radiation which favored IMRT (mean change from baseline: -2.8 ± 26 vs. 8.3 ± 15, p=0.04).

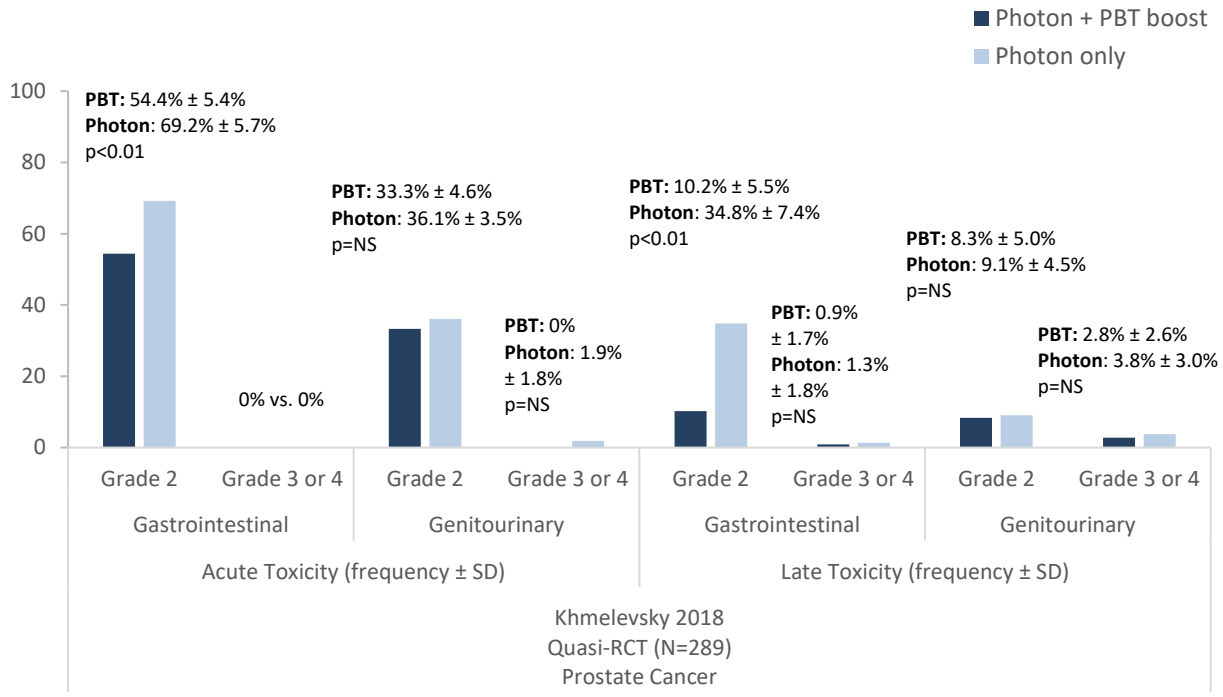
Key Question 3 (Safety)

Comparative studies

In addition to the quasi-RCT and the retrospective cohort study included for effectiveness, two additional retrospective cohort studies, both employing case-matching methods, were identified that reported only safety outcomes following PBT compared with IMRT for prostate cancer. One study case-matched patients from a single institution based on risk group, age, and prior GI and GU disorders[76] and the second study conducted a propensity score-matched analysis using data from the MarketScan Commercial Claims and Ecounters database[216] (Table 40, Section on Effectiveness). Across the studies, patient ages ranged from 55 to 69 years. The proportion of patients with comorbidities differed between the studies and only one study[76] reported patients' risk level, which was primarily low (55%) or intermediate (31%) (opposite of the population included in the quasi-RCT). Passive scatter PBT was employed in one study[76]; PBT technique was unknown in the other study as the database did not distinguish between types of PBT.[216] The mean total radiation doses were not reported. Roughly 20% of both populations had undergone prior androgen deprivation therapy prior to commencing RT (compared with 95% in the quasi-RCT).

In the quasi RCT,[139] PBT boost resulted in a significantly lower frequencies of grade 2 acute (54% vs. 69%) and late (10% vs. 35%) gastrointestinal (GI) toxicity compared with the photon only group (p<0.01); grade 3 or 4 events were rare and occurred with similar frequency between groups (Figure 25). No statistical differences between groups were seen for acute or late genitourinary (GU) toxicity (grade ≥2). The authors note that the different PBT fractionation regimens did not significantly differ in toxicity levels.

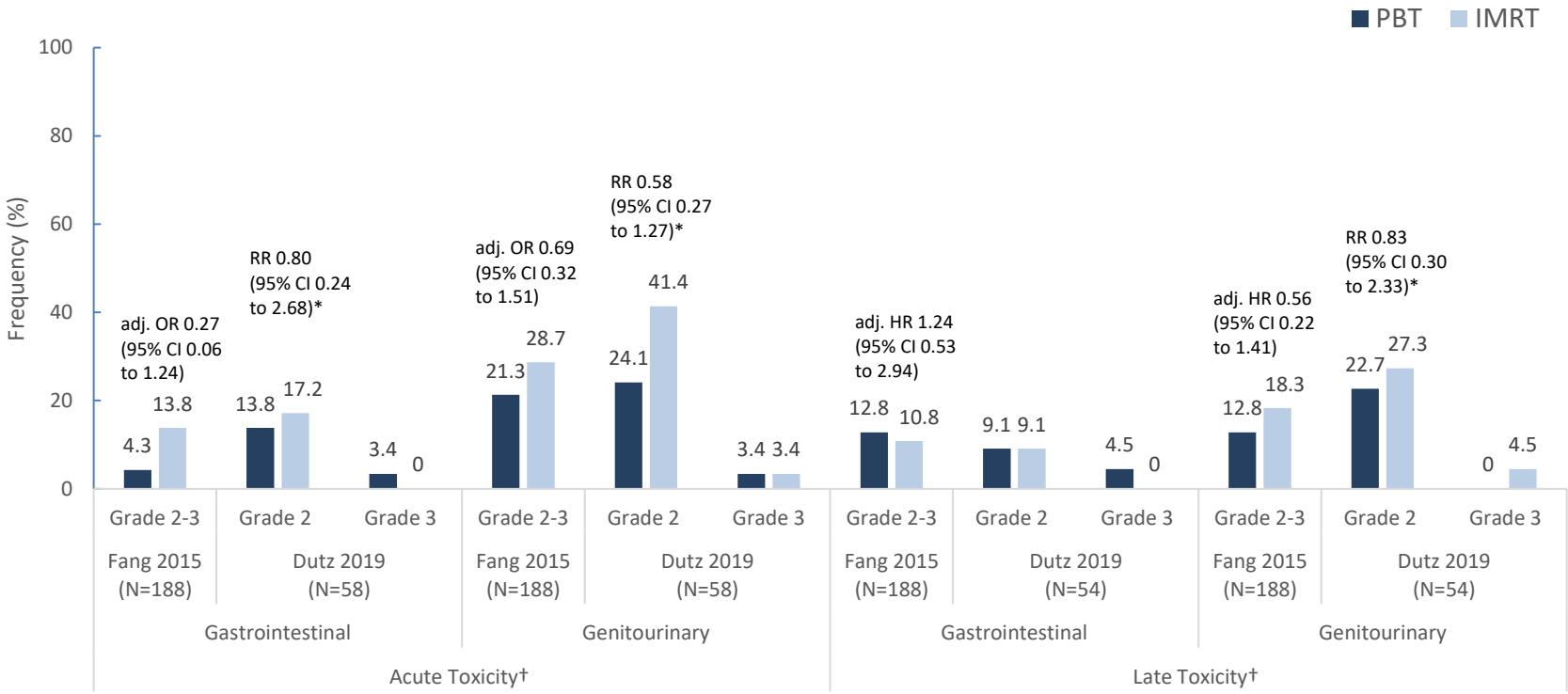
Figure 25. Frequency of Gastrointestinal and Genitourinary Toxicity in a Quasi-RCT Comparing Photons + PBT Boost with Photons alone for Curative Intent for Adults with Prostate Cancer



NS = not statistically significant; PBT = proton beam therapy; RCT = randomized controlled trial; SD = standard deviation.

The two smaller cohort studies (one case-matched)[76] and one propensity-score matched)[69] found that the frequencies of acute and late GI and GU toxicity (grade 2 and/or 3) were statistically similar following treatment with PBT versus IMRT. Figure 26. (For data regarding grades 0 to 1 toxicities, see the table in section 5.2.9. Sample size may have played a role in some of these findings. Conversely, in the large propensity score-matched analysis (N=4,158),[216] the cumulative incidences of any grade of urinary toxicity and erectile dysfunction were statistically reduced while bowel toxicity was statistically increased following PBT compared with IMRT, Figure 27; sample size likely influenced the finding of statistical significance. The latter database study also conducted two sensitivity analyses. In one, they included only procedure codes (excluded diagnosis codes) as a surrogate of toxicity severity and found that the 2-year incidence remained significant only for any urinary toxicity: 1.3% vs. 4.7%, HR 0.24, 95% CI 0.12 to 0.48 (any bowel toxicity: 2.5% vs. 2.3%, HR 1.50, 95% CI 0.91 to 2.47; erectile dysfunction: 2.0% vs. 3.1%, HR 0.63, 95% CI 0.36 to 1.10). In the second analysis, they included only combinations of procedure and diagnosis codes previously validated for five severe radiation-induced pelvic toxicities (cystitis, rectal complications, urethral stricture, ureteral stricture, and urinary/rectal fistula); only the 2-year incidence of urethral stricture differed significantly between the groups: PBT 0% vs. IMRT 1.3% (HR 0.12, 95% CI 0.02 to 0.86).

Figure 26. Frequency of Gastrointestinal and Genitourinary Toxicity across Two Retrospective Cohorts Comparing PBT with IMRT for Curative Intent for Adults with Prostate Cancer

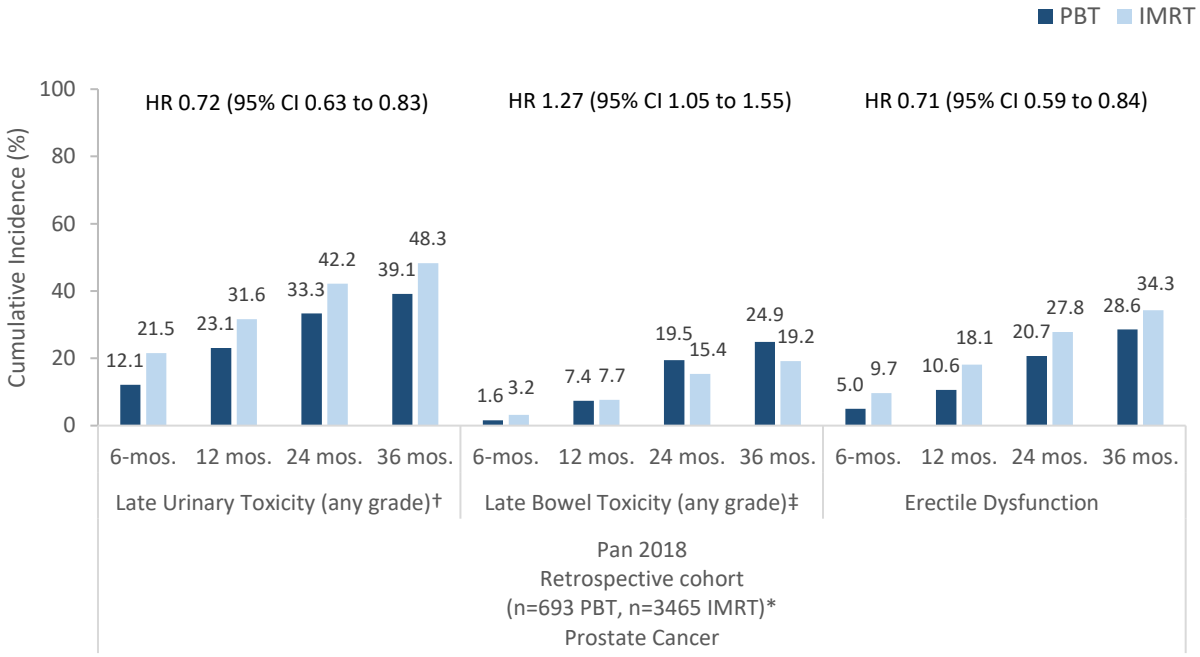


adj. = adjusted; CI = confidence interval; PBT = proton beam therapy; HR = hazard ratio; IMRT = intensity-modulated radiation therapy; mos. = months; OR = odds ratio; RR = risk ratio.

*RR and 95% CI were calculated by Aggregate Analytics, Inc.

†Acute toxicity was defined by both studies as occurring at ≤3 months. Late toxicity was defined by Fang 2015 as occurring at >3 months and by Dutz 2019 as occurring at 12 months.

Figure 27. Frequency of Gastrointestinal and Genitourinary Toxicity in a Retrospective Database Study (Pan et al.) Comparing PBT with IMRT for Curative Intent for Adults with Prostate Cancer



CI = confidence interval; PBT = proton beam therapy; HR = hazard ratio; IMRT = intensity-modulated radiation therapy; mos. = months.
*Propensity score-matched. Number of patients available for analysis at each timepoint is as follows: 6 mos. (PBT: n=693; IMRT: n=3465); 12 mos. (PBT: n=572; IMRT: n=2862); 24 mos. (PBT: n=341; IMRT: n=1718); and 36 mos. (PBT: n=205; IMRT: n=1003)
†Includes: Bleeding/irritation (most common urinary adverse event in both groups), incontinence, obstruction/retention, stricture, and fistula.
‡Includes: Bleeding/proctitis (most common bowel adverse event in both groups), ulcer/stricture/fistula, incontinence, proctectomy/hyperbaric oxygen.

Case series

Severe gastrointestinal (GI) and genitourinary (GU) toxicities were rare as reported by eight case series. In the acute period, there were no grade ≥3 GI toxicities (4 studies)[20,50,227,285] and the rate of grade ≥3 GU toxicities ranged from 0% to 0.9% (5 studies).[20,39,50,53,227,285] Late GI and GU toxicities of grade ≥3 ranged from 0% to 1.2% and from 0% to 4.7%, respectively, across eight studies.[20,39,50,53,124,172,227,270,285]

Key Question 4 (Differential Effectiveness and Safety)

For this key question, RCTs that stratified on baseline patient characteristics and evaluated effect modification were sought. Subgroups of interest included (but were not limited to): age, sex, race, ethnicity, socioeconomic status, payer, and worker's compensation. All RCTs included to evaluate the efficacy or safety of PBT were assessed. No trials meeting the inclusion criteria were identified.

To be consistent with the prior report, comparative studies related to different treatment protocols or different dosing regimens for PBT were included here for context. We identified four such comparative studies, two RCTs and two retrospective cohort studies, all in men with prostate cancer. Three studies compared PBT treatment regimens; two studies, one RCT conducted in the United States (US) and one

cohort study from Japan,[198,286] compared hypofractionated versus standard fractionated image-guided PBT regimens and the third study, a RCT conducted in Korea[98], compared "moderate" hypofractionation versus "extreme" hypofractionation. The fourth study, a retrospective cohort from the US, compared two methods of PBT delivery, passively scattered (PSPT) versus spot scanning proton therapy (SSPT).[227] Of note, loss to follow-up could not be determined for either comparative cohort study and neither controlled for potential confounding. See Abstraction Appendix P for details regarding study and patient characteristics.

PBT Treatment regimens

Hypofractionation versus standard fractionation

In the RCT (N=82), men with primarily stage T1c prostate cancer were randomized to receive either 38 Gy RBE in 5 fractions (hypofractionated PBT) or 79.2 Gy RBE in 44 fractions (standard fractionated PBT).[286] All patients had the same volume definitions, margins, immobilization, and setup; however, dose volume constraints were proportionally scaled down for the hypofractionated arm. Patients were followed for a median of 18 months. In the retrospective cohort study (N=526), men with primarily stage T2 prostate cancer were treated with hypofractionated PBT [60 Gy (RBE) in 20 fractions for low risk and 63 Gy (RBE) in 21 fractions for intermediate and high-risk patients] or standard fractionated PBT [74 Gy (RBE) in 37 fractions and 78 Gy (RBE) in 39 fractions, respectively]; PBT was delivered via passive scattering technique in most cases (94%).[198] Patients were followed for about 6 months.

No differences between groups were seen in any quality of life (QoL) or safety outcome measured in the RCT [American Urological Association Symptom Index, Expanded Prostate Index Composite (EPIC) questionnaire, grade 2 genitourinary (GU) and gastrointestinal (GI) tract toxicities].[286] In the retrospective cohort study, no difference between treatment groups overall was seen in the International Prostate Symptom Score (IPSS).[198] The cumulative incidences of grade 2 GU toxicity and grade 1 dermatitis, but not grade 1 GI toxicity, were statistically lower following hypofractionated PBT (6% vs. 15% and 7% vs. 18%, respectively, p<0.001); when results were analyzed by risk group, the differences remained statistically significant for the intermediate and high risk groups but were similar for those at low risk. Across both studies, no grade ≥3 toxicity occurred and no treatment-related deaths in either arm were reported in the RCT.

"Moderate" versus "extreme" hypofractionation

A total of 82 men with primarily stage T2 prostate cancer who had not received androgen-deprivation therapy (ADT) were randomized to one of five treatment groups (see Table O1 in the Data Abstraction Appendix for details), which were further categorized into the "moderate" hypofractionated (MHF) group (<5 Gy/fraction, n=52) and the "extreme" hypofractionated (EHF) group (≥5 Gy/fraction, n=30).[98] Patients were followed for a median of 7.5 years. The probability of 7-year overall survival (OS) was 97.5% for the entire population (total of three deaths); OS was not compared between the groups. The probability of 7-year biochemical failure-free survival (BCFFS) was significantly lower in the EHF compared with the MHF group (46.2% vs. 76.2%; adjusted HR 3.2, 95% CI 1.5 to 6.9, p=0.003); this was also the case when comparing those at intermediate risk (but not at low or high risk) of progression: 42.9% versus 83.5% in the MHF group (p=0.02). Acute gastrointestinal (GI) or genitourinary (GU) toxicities of grade ≥3 were not observed in either group. There were only two cases (4%) of late grade 3 toxicity (GI) which occurred in the MHF group. Overall, acute GU toxicities (grades 0-2) were more

frequent following MHF (85%) versus EHF (57%) (p=0.009), but late GI and GU toxicities did not differ significantly between groups.

PBT delivery method

One retrospective cohort study evaluated men with previously untreated, localized prostate cancer (primarily stage T1c) treated with PSPT (n=226) or SSPT (n=65).[227] The total prescribed radiation dose was 76 Gy (RBE) delivered in 38 equivalent fractions. Follow-up period was 24 months. No statistically significant differences between groups were seen for QoL (EPIC questionnaire) or for cumulative frequencies, respectively, of grade ≥2 GU (14% vs. 11%) and GI (10% vs. 8%) toxicities or of argon plasma coagulation application for rectal bleeding (4% vs. 2%, respectively). There was one grade 3 GI toxicity which occurred in the PSPT group; no other grade ≥3 toxicity was reported.

Key Question 2 (Effectiveness, salvage therapy) and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.12 Benign and Mixed Tumors

4.3.12.1 Hemangiomas

Key Points

- There is insufficient evidence from two case series to evaluate the effectiveness and safety of PBT for curative intent in adults with hemangiomas.
- No studies meeting inclusion criteria were identified that evaluated salvage PBT, differential effectiveness and safety or cost-effectiveness.

Description on included studies

No comparative studies of PBT for the treatment of hemangiomas that met inclusion criteria were identified.

Two small case-series (N=43 and 55) evaluating curative PBT for the treatment of circumscribed chorodial hemangiomas that met inclusion criteria were identified.[169,323] Mean patient age in both studies was 49 years and the majority were male (64% and 74%). Retinal detachment was present at initial assessment in 44% and 91% of patients; almost all patients in both studies presented with loss of visual acuity. PBT technique was not reported. Total PBT dose was 20 Gy (RBE) in both studies administered in four fractions over four sequential days in one[323] or in eight fractions over 11 days in the other.[169] In both studies, about one-fifth of the patients had undergone prior treatment before PBT (18% and 23%). Mean follow-up periods ranged from 26 to 55 months.

All case series are considered to be high risk of bias.

Results

Key Question 1 (Effectiveness, curative intent)

Primary and other outcomes
Improvement in visual acuity was reported by both case series (Main Appendix F). One study, which reported Snellen equivalents, found that visual acuity improved from 6/15 at baseline to 6/12 at final follow-up (mean 55 months); the proportion of patients with two line improvement at 2, 3, and 4 years was 37%, 44%, and 59%, respectively in this study.[323] The second study reported that visual acuity according to the Snellen scale improved to 20/25 (from 20/63 at baseline) at a mean follow-up of 26 months; 86% of patients had either stabilized or two line improvement.[169] Tumor thickness decreased significantly compared with baseline in both studies. Complete attachment of the retinal detachment was seen in all but one patient (2.3%; 1/43)[169] across both studies.

Key Question 3 (Safety)

One case series (N=50)[323] reported various radiation-related side effects over the course of follow-up, including cataract formation in 20% of patients; optic neuropathy in 8%; retinopathy of stage IV Finger classification (i.e., sight-threatening), vitreous hemorrhage (secondary to radiation retinopathy), and retinal vein occlusion in 4% each. There were no cases of rubeosis iridis. It is unclear from the information provided whether or not patients in this study could have had more than one event. In the second case series, 7% of patients developed a radiation cataract; there were no cases of radiation maculopathy or papillopathy.[169]

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety) and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.12.2 Other Benign

Key Points

- There is insufficient evidence from three case series to evaluate the effectiveness and safety of PBT for other non-cancerous tumors (i.e., meningioma, pituitary adenoma).
- No studies meeting inclusion criteria were identified that evaluated salvage PBT (i.e., no comparative studies), differential effectiveness and safety or cost-effectiveness.

Description of included studies
No comparative studies of PBT for the treatment of benign tumors that met inclusion criteria were identified.

Four case-series (N=61to 170) evaluating PBT for the treatment of benign tumors (including meningioma and pituitary adenoma) that met inclusion criteria were identified (Main Appendix F).[73,197,299,304] One of these series included both malignant (WHO grade 2/3) and benign (WHO grade 1) meningiomas[197]; only data for those patients with benign tumors is described here and information regarding the malignant population can be found in the section on Brain Tumors.

In the three studies of benign meningioma (N=61, 110 and 170)[73,197,299] the majority of the participants were female (70% to 80%)with median ages ranging from 52 to 54.2 years. Indication for treatment was entirely curative in one study (benign meningioma)[299] and a combination of curative (primarily) or salvage in the others.[73,197] PBT techniques included raster-scanning,[73] hypofractionated passive-scattering[299] and pencil beam scanning[197] with a median total doses ranging from 21.9 (in the hypofractionated PBT study) to 54 Gy(RBE), respectively. The majority of patients in two studies (62% and 74%) underwent pre-radiation surgery (either subtotal resection or biopsy).[73,299] Median follow-up periods ranged from 46.8 to 84 months.

The third case series (N=165)[304] included patients (76% female, median age 43 years) with functional pituitary adenoma who received passive-scatter PBT for salvage treatment or for residual tumors, with a median total dose of 20 Gy (RBE) after either prior resection (98.2%) or prior photon irradiation (8.5%). Median follow-up for all patients was 51.6 months.

All case series are considered to be high risk of bias.

Key Question 1 (Effectiveness, curative intent)

Primary and other outcomes

The probability of overall survival at 5 years was high reported by two studies (92.1% and 96.2%)[73,197]; one of these studies also reported OS at 10 years (98.1%) and 15 years (90.7%) post-diagnosis.[73]. The third study reported mortality (as opposed to OS) with an all-cause mortality rate of 13.5% over a median follow-up of 84 months; in three cases (1.7%) the deaths were related to the treated meningioma.[299] Two studies reported the probability of PFS which was 100% at 3 years (1 study), 93% and 97% at 5 years (2 studies) and 85% at 10 years (1 study).[73,299] The 5-year probability of local control was 95.7% as reported by one study.[197]

Key Questions 2 (Effectiveness, salvage therapy)

Primary and other outcomes

One study (N=165) of salvage PBT for recurrent pituitary adenomas met criteria for inclusion. [304] The probability of complete metabolic response was 42% at 3 years and 59% at 5 years, with a median time to complete response of 47 months. Local control was achieved in 98% of patients with available follow-up data

Key Question 3 (Safety)

Three case series of PBT for primary and recurrent disease reported data on safety outcomes. {Vlachogiannis, 2017 #66; El Shafie, 2018 #136; Wattson, 2014 #90;

In the two studies of benign meningioma,{Vlachogiannis, 2017 #66;El Shafie, 2018 #136} toxicity data were reported over either acute (≤6 months) and late (>24 months) time periods or were in general (i.e., time period not specified). Grade ≥3 acute toxicities occurred in two patients (1.8%) and late toxicities in five patients (4.5%) in one study; the latter included three cases (2.7%) of radiation necrosis. In the other case series evaluating meningioma, the frequency of any toxicities (grades not reported) was 9.4% (N=70).

Rates of hormone deficiency that required hormone replacement therapy were reported at 3 years (45%) and 5 years (62%) in the single study of pituitary adenoma.[304] This study reported that 4.2% of patients suffered some toxicity but severity and timing post-PBT was not reported. Osteoradionecrosis of the skull-base was rare, occurring in <1% of patients and was considered to be unrelated to PBT. No cases of secondary malignancy were reported.

Key Question 4 (Differential Effectiveness and Safety) and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

4.3.12.3 Mixed Tumor Types

Key Points

- There is insufficient evidence from three case series to evaluate the effectiveness and safety of PBT for mixed tumor populations.
- No studies meeting inclusion criteria were identified that evaluated salvage PBT, differential effectiveness and safety, or cost-effectiveness in mixed tumor populations.

Description of included studies

No comparative studies of PBT in mixed tumor populations were identified that met inclusion criteria.

Three case-series (N=365, 90, and 56) evaluating PBT for curative intent in the treatment of mixed tumor populations were identified that met criteria for inclusion. [194,207,327]. Conditions in these studies varied widely including mixed brain, spinal and bone cancers as well as tumors of the head and neck, lung, liver, ovarian and more with no particular conditions making up a majority of any study. Age also varied across all three studies, with a median age of 11.2 in a study with mixed adult and pediatric population,[194] to median and mean ages of 66 and 54 years in the other studies.[207,327] Spot-scanning PBT was used in one study[207]; PBT technique was not reported in the other studies. PBT dose varied depending on tumor site and grade. Median total radiation dose ranged between 59.8 and 64.5 Gy (RBE) in the largest case series[327], between 4655 and 5500 cGy in another study[194], and between 20 and 76 Gy(RBE) in the third.[207] Only one case series reported follow-up (12 months).

All case series are considered to be high risk of bias.

Key Question 1 (Effectiveness, curative intent)

Primary and other outcomes

One case series (N=56) reported mortality; over a 12 month follow-up period four patients (7.1%) died, two of whom had disease progression.[207] None of the other case series reported primary outcomes.

Key Question 3 (Safety)

The rates of grade ≥3 acute toxicities ranged from 2% to 10% across two case series (N=90 and 56)[194,207]. One study reported an incidence of grade 4 acute toxicities of 0% (95% CI 0% to 6.38%) over a 12 month follow-up period. Grade ≥3 late toxicities were reported in only one study,[207] which reported a single

case (1.8%) of osteoradionecrosis. The third study (N=375)[327] (n=375) reported only weight loss outcomes; these included average weight lost (0.55 kg), mean body weight decrease (-2.2 kg), and average percent of body weight loss among patients with critical weight loss (8.7%) and among patients without critical weight loss (0.2%).

Key Question 2 (Effectiveness, salvage therapy), Key Question 4 (Differential Effectiveness and Safety) and Key Question 5 (Economic)

No studies that met inclusion criteria were identified.

5 Strength of Evidence (SOE)

Strength of evidence (SOE) tables are provided only for comparative studies. Most were considered to be at moderately high risk of bias. Individual study ratings are found in Main Appendix E. All case series were considered to be at high risk of bias; in the absence of studies comparing patients from the same underlying population (using contemporaneous cohorts of patients assigned to respective treatments), the evidence was considered to be insufficient to draw conclusions regarding effectiveness or safety of PBT particularly with regard to other forms of radiation therapy.

For the following tumor types only evidence from case series was available and therefore SOE was not competed: Pediatric lymphoma, Pediatric ocular tumors, Pediatric soft tissue sarcomas, Pediatric bone tumors, bladder cancer, bone tumors, breast cancer, lymphoma, and benign tumors.

Determination and interpretation of SOE are described in the Methods section. Bodies of evidence consisting of RCTs are initially considered as High strength of evidence. In general, the GRADE and AHRQ methodologies initially consider nonrandomized studies as Low strength of evidence as such studies typically are at higher risk of bias due to lack of randomization and inability of investigators to control for critical confounding factors. Observational studies with few methodologic limitations which control for risk of bias via study conduct or analysis may be initially considered as moderate versus low, particularly for harms and outcomes when such studies may be at lower risk of bias due to confounding.[29]

5.1 Strength of Evidence Summary: Pediatric Tumors

5.1.1 Strength of Evidence Summary for Pediatric Brain, Spinal, and Paraspinal Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
KQ 1 Curative intent								
Survival outcomes								
Probability, overall survival	3 year	Sato 2017† (N=79) Ependymoma	No	Unknown (different tumor types)	No	Yes[3] (-1)	97% (83%-99%) vs. 81% (63%-90%) NR; p=0.08	PBT resulted in similar (3 studies, Bishop, Kopecky, Eaton) or slightly greater (2 studies, Sato, Gunther) OS compared with IMT or CRT however statistical significance was not reached in any study at any time; sample sizes may play a role. ⊕⊕◯◯ LOW
		Bishop 2014 (N=52) Craniopharyngioma					94.1% (NR) vs. 96.8% (NR) NR; p=0.742	
	4 year	Gunther 2015† (N=72) Ependymoma					87.5% (51.6% - 97.3%) vs. 78.8% (60.6% - 89.3%); NR; p=0.21	
	5 year	Kopecky 2017§ (N=783) Medulloblastoma					%NR HR 0.99 (0.41 to 2.4); p=0.98 (PBT vs. CRT)	
	6 year	Sato 2017† (N=79) Ependymoma					88% (NR) vs. 70% (NR)‡ NR	
		Eaton 2016a,b (N=88) Medulloblastoma					82.0% (65.4% - 91.1%) vs. 87.6% (72.7% - 94.7%) adjHR, 2.17 (0.66 to 7.16);	
Probability, Progression free or relapse free survival	3 year	Sato 2017† (N=79) Ependymoma	No	Unknown (different tumor types)	No	Yes[3] (-1)	PFS: 82% (64%-92%) vs. 60% (42%-74%) HR (vs IMRT), 0.42 (0.16-1.10);	At 3 and 6 years, PFS in patients with ependymoma who received PBT tended to

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
	6 year	Eaton 2016a,b (N=88) Medulloblastoma					RFS: 78.8% (63% -89%) vs. 76.5% (60.6% - 86.6%); adjHR 1.31 (0.5 to 3.41);	have longer PFS vs. IMRT, but differences were not statistically significant at 3 years. RFS was similar between groups in patients with medulloblastoma ⊕⊕◯◯ LOW
		Sato 2017† (N=79) Ependymoma					PFS: 82% (NR) vs. 38% (NR) NR	
Other Primary								
Any recurrence or relapse	74.4 mos. vs. 85 mos.	Eaton 2016a (N=88) Medulloblastoma	No	Unknown (different tumor types)	No	Yes[3] (-1)	22.2% (10/45) vs. 23.3% (10/43); NR	Recurrence was similar between groups in patients with medulloblastoma however was significantly less common in patients with ependymoma ⊕⊕◯◯ LOW
	31.2 vs. 58.8 mos.	Sato 2017† (N=79) Ependymoma					17% (7/41) vs. 55% (21/38), RR 0.31 (0.15 to 0.64)	
KQ 3: Safety Outcomes								
Hypothyroidism	56.4 mos. vs. 121.2 mos.	Bielamowicz (N=84) Medulblastoma	No	No	No	Yes[3] (-1)	**Hypothyroidism (any):** 19% vs. 46.3%; adj HR 1.85 (0.8 to 4.2) **Primary hypothyroidism:** 7.3% vs. 20.4%; adj HR 2.1 (0.6 to 7.7) **Central hypothyroidism:** 9.8% vs. 24.0% ; adj HR 2.2 (0.7 to 6.6)	Across 2 studies, hypothyroidism was less common with PBT statistical differences were only seen in one study ⊕⊕◯◯ LOW
	69.6 mos. vs. 84 mos.	Eaton 2016b (N=77) Medulloblastoma					**Hypothyroidism:** 22.5% (9/40) vs 64.9% (24/37); adj OR: 0.13 (0.04 to 0.41)	

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
Other Endocrine toxicities	33.1 mos. vs. 106.1 mos.	Bishop 2014 (N=52) Craniopharyngioma	No	Unknown	No	Yes[3] (-1)	**Panhypopituitarism:** 33% (7/21) vs. 55% (17/31); RR 0.61 (0.31, 1.2) **Other** endocrinopathy: 43% (9/21) vs. 23% (7/31); RR 1.9 (0.84, 4.3)	Other specific endocrinopathies across the two studies tended to be less common in PBT recipients compared with other forms of radiation therapy; however, statistical significance was only achieved for sex hormone deficiency. Endocrine replacement therapy was less common in those receiving PBT vs. photon RT. ⊕⊕◯◯ LOW
	69.6 mos. vs. 84 mos.	Eaton 2016b (N=77) Medulloblastoma	No	Unknown	No	Yes[3] (-1)	**Growth hormone deficiency:** 52.5% (21/40) vs. 56.76% (21/37); adj OR 0.81 (0.26 to 2.59) **Sex hormone deficiency**: 2.5% (1/40) vs. 18.92% (7/37); adj OR 0.06 (0.01 to 0.55) **Endocrine replacement therapy**: 55% (22/40) vs. 78.38% (29/37) adj OR 0.30 (0.09 to 0.99)	
Changes in IQ score changes per year	32.4 mos. vs. 64.8 mos.	Kahalley 2016 (N=150) Various brain tumors	No	Unknown	No	Yes[3] (-1)	**FSIQ (adjusted beta coefficient, 95%CI) PBT vs. Photon RT** All patients** -0.7 (-1.6 to 0.2) vs. -1.1 (-1.8 to -0.4; p= 0.51 **CSI: -** 0.8 vs. -0.9 (CIs NR); p = 0.89 **Focal RT:** 0.6 (-2.0 to 0.8) vs. -1.6 (-3.0 to -0.2); p = 0.34	There were no differences between PBT and photon radiation in with regard to changes in IQ scores. ⊕⊕◯◯ LOW

Final

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
	33.6 mos. to 37.2 mos. post-treatment	Kahalley 2019 (N=93) Various brain tumors	No	Unknown	No	Yes[3] (-1)	**Focal PBT vs. surgery** NS differences FSIQ or for any subscale (all p-values >0.05); scores remained stable for both groups over time. **CSI PBT vs. surgery (adjusted beta coefficient, 95% CI)**** FSIQ: -2.1 (-3.8 to -0.3), p = 0.020 PSI; -2.6 (-4.7 to -0.3), p = 0.019. NS differences for any other subscales (all p-values >0.05)	There were no differences between focal PBT and surgery in changes in FSIQ or subscores after adjustments for baseline differences. CSI PBT was associated with a decline in FSIQ and PSI with time compared with surgery. The clinical significance of the changes is not described. ⊕⊕◯◯ LOW
Other Late toxicities or adverse events (Median f/u by treatment)	PBT 33.1 mos. vs. 106 mos.	Bishop 2014 (N=52) Craniopharyngioma	No	Unknown	No	Yes[3] (-1)	**Vascular Injury (on imaging),** 10% (2/21) vs. 10% (3/31); **Vision changes:** 5% (1/21) vs. 13% (4/31); RR 0.37 (0.04, 3.07) **Hypothalamic obesity:** 19% (4/21) vs. 29% (9/31); RR 0.66 (0.23, 1.9)	Risk of vascular injury, hearing loss and radiation necrosis were similar between PBT and other types of RT; although risk of vision changes and hypothalamic obesity were somewhat lower for PBT in one study, groups were not statistically different. ⊕⊕◯◯ LOW
	55.5 mos. vs.65.5 mos.	Paulino 2018 (N=84) Medulloblastoma	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	**Hearing Loss (worse ear)** • *Grade 3:* 26.3% (10/38) vs. 21.7% (10/46) • *Grade 4:* 2.6% (1/38) vs. 6.5% (3/46)	

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. IMRT* or CRT Effect estimate (95% CI)	Conclusion Quality (SoE)
							• *Grade 3 and 4:* 29.9% (11/38) vs. 28.3% (13/46), p=1.0	
	31.2 mos. vs. 58.8 mos.	Sato 2017 (N=79) Ependymoma	No	Unknown	No	Yes[3] (-1)	**All events:** 7.3% (3/41) vs. 13.2% (5/38); RR 0.56 (0.14, 2.17) **Radiation Necrosis**: 7.3% (3/41) vs. 7.9% (3/38) **Stroke**: 0% (0/41) vs. 2.6% (1/38) **Cavernoma:** 0% (0/41) vs. 2.6% (1/38)	
Acute Toxicities	Acute	Song 2014 (n=30 PBT, n=13 photon) Various tumors	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	**Leukopenia** • Grade 3: 57% (14/30) vs. 46% (6/13) • Grade 4: 7% (2/30) vs. 31% (4/13) • **Grade 3 or 4** RR: 0.68 (0.44, 1.08) **Anemia** • Grade 3: 0% (0/30) vs. 15% (2/13); p=0.493 • Grade 4: 0% (0/30) vs. 0% (0/13) **Thrombocytopenia:** • - Grade 3: 20% (6/30) vs. 31% (4/13) • - Grade 4: 3% (1/30) vs. 23% (3/13); **Grade 3 or 4 RR**: 0.43 (0.19, 0.98)	Frequency of acute Grade 3 or 4 hematological toxicities was lower with PBT vs. photon RT, however the overall sample size is small, particularly in the photon group. There is insufficient evidence to draw conclusions. ⊕◯◯◯ INSUFFICIENT

adj RR= adjusted risk ratio; CI = Confidence Interval; f/u = follow-up; FSIQ = Full Scale Intelligence Quotient; HR = Hazard Ratio; IMRT = Intensity Modulated Radiation Therapy; IQ = Inteligence Quotient; NR = Not Reported; NS = Not significant; OR = Odds ratio; OS = Overall Survival; PBT = Proton Beam Therapy; PFS = Progression Free Survival; PSI = Processing Speed Index; RFS = Recurrence Free Survival; RR = crude Risk Ratio; RT = Radiation Therapy; SOE = Summary of Evidence

* PBT was compared with IMRT in Bishop, Gunther and Sato; IMRT or 3DCRT was used in Eaton; Kopecky had 3 arms; PBT, IMRT and 2D/3D CRT but effect sizes were only reported for PBT vs. 2D/3D CRT not for PBT vs. IMRT;

† Sato and Gunther report on the same underlying patient population. Sato 6 year estimates from author's graph
‡PBT was done as “definitive” treatment in 13% and post-operative/adjuvant treatment in 44%, salvage treatment in 42%
§ 517 pts (of the 1300 identified) diagnosed after 2009 were excluded from survival analysis leaving 783 for survival analysis across three treatment groups but authors do not specify to which treatment group they belong or the number of patient with PBT and CRT which were compared in survival analysis
** Authors do not provide mean changes only beta coefficients and p-values; Beta coefficients represent the increase or if negative, decrease in points per year on each index by treatment group. Inclusion of 0 in the confidence interval signifies results are not statistically significant.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.1.2 Strength of Evidence Summary for Pediatric Head and Neck Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	**Serious Risk of Bias**	**Serious Inconsistency**	**Serious Indirectness**	**Serious Imprecision**	PBT vs. other RT * Effect estimate (95% CI)	Conclusion Quality (SoE)
Head, Neck								
Toxicity	Acute	Grant (N=24) 1 Retro cohort (N=24) Salivary Gland tumors	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	Grade 2/3 acute toxicities: Dysphagia (0 vs. 3/11), otitis externa (1/13 vs. 2/11) mucositis (6/13 vs. 10/11, RR 0.51 (0.27, 0.94).	Mucositis may be less common following adjuvant PBT vs. adjuvant photon RT; risk of other toxicities was similar between groups. ⊕◯◯◯ Insufficient
Ocular (Salvage)								
Effectiveness	Last f/u 3 yrs PBT, 10 yrs RT	Agarwal 2016 (N=39 patients, 47 eyes) Retinoblastoma	Yes[1] (-1)	Unknown	No	Yes[3] (-2)	OS: 97.4% across groups Enucleation-free survival: 38.5% vs. 54.5% Enucleation performed: 37.5% (6/16 eyes) vs. 29.6% (8/27 eyes)	No comparative data reported for OS. Enucleation-free survival was lower with PBT, however small sample size, may preclude detection of statistical difference ⊕◯◯◯ Insufficient
Toxicity	Acute Late						**Acute Toxicity:** PBT 93.8% vs. ERT 74.1%; p =0.22 (mostly skin erythema) **Late/long-term** (number of eyes): **PBT vs. ERT** Any (≥1 event): 62.5% (10/16 eyes) vs. 55.6% (15/27 eyes); p=0.275 **PBT vs. Other Treatment** Cataract: 5 vs. 10 Vitreous hemorrhage: 3 vs. 4	Although acute toxicities were more common with PBT vs. ERT, differences were not statistically significant. Evidence is limited ⊕◯◯◯ Insufficient

Outcome	Time	Studies, Year, N, Tumor	**Serious Risk of Bias**	**Serious Inconsistency**	**Serious Indirectness**	**Serious Imprecision**	PBT vs. other RT * Effect estimate (95% CI)	Conclusion Quality (SoE)
							Radiation retinopathy: 2 vs. 3 Visual acuity Δ: 0 vs. 4 Strabismus: 1 vs. 2	

adj RR= adjusted risk ratio; CI = Confidence Interval; f/u = follow-up; ERT= electron beam radiation therapy; HR = Hazard Ratio; IMRT = Intensity Modulated Radiation Therapy; NR = Not Reported; NS = Not significant; OR = Odds ratio; OS = Overall Survival; PBT = Proton Beam Therapy; PFS = Progression Free Survival; RFS = Recurrence Free Survival; RR = crude Risk Ratio; RT = Radiation Therapy; SOE = Summary of Evidence

* Grant compared PBT (passive scatter n =8, intensity modulated n=5) vs. other RT (electron bean n=8, IMRT n=3); Agarwal compared PBT (passive scatter, n= 16 eyes) vs. photon or electron RT (n=27 eyes) and brachytherapy (n= eyes).

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2 Strength of Evidence Summary: Adult Tumors

5.2.1 Strength of Evidence Summary for Adult Brain, Spinal, Paraspinal Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon* Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)								
Survival outcomes								
Probability, overall survival (OS)†	1-3 years	Adeberg 2017 (N=132) Retro case-matched cohort Glioblastoma (high-grade)	No	Unknown	No	Yes[3] (-1)	PBT boost + photon vs. photon alone: 1 year: 75% vs. 85% 2 years: 40% vs. 43% 3 years: 12% vs. 28% p=NS at all timepoints	Results across studies and tumors types are inconsistent For those with high-grade glioblastoma, PBT boost tended to result in lower OS but higher PFS probability versus photon alone; results were not statistically significant but may be clinically meaningful. In the large database study of primarily high-grade glioma, statistically higher 5-year overall survival was reported following PBT versus photon RT. Of note, the median follow-up period was significantly shorter in the PBT group (50.3 vs. 62.3 months). There is the potential for misclassification in database studies. ⊕⊕◯◯ LOW
	5-years	Jhaveri 2018 (N=49,575) Retro comparative database study; propensity-score matched cohort (n=322) Glioma (91% high-grade)	No	Unknown	No	No	*PBT vs. any photon, entire cohort:* adj. HR 0.66, 95% CI (0.53 to 0.83); favors PBT *PBT vs. any photon, propensity-score matched:* 46.1% vs. 35.5%, p=0.009 vs. IMRT: p=0.01 vs. 3D-CRT: p=0.007	
Probability, Progression free survival (PFS)†		Adeberg 2017 (N=132) Retro case-matched cohort Glioblastoma (high-grade)	No	Unknown	No	Yes[3] (-1)	PBT boost + photon vs. photon alone: 1 year: 31% vs 21% 2 years: 8% vs 2% p=NS at both timepoints	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon* Effect estimate (95% CI)	Conclusion Quality (SoE)
Salvage therapy (KQ2)								
Survival and recurrence outcomes								
Probability, overall survival	6 mos. – 1 years	Gunther 2017 (N=37) Retro cohort CNS involvement in lymphoma or leukemia (pre-SCT)	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	PBT vs. Photon: 6 mos.: 78.6% vs. 69.6%, p=0.15 1 year: 70% vs. 38%,‡ p=NR	No statistical difference between groups in OS at 6 months, statistical testing not reported at 1 year; no statistical difference in CNS relapse risk. Sample size may have played a role in these findings. ⊕◯◯◯ INSUFFICIENT
CNS relapse	5 mos.						PBT vs. Photon: 7% (1/14)§ vs. 0% (0/23); p=1.0	
Safety (KQ3)								
Acute Toxicity (≤3 mos.)	Median 15 mos.	Adeberg 2017 (N=132) Retro case-matched cohort Primary Glioblastoma (high-grade)	No	No	No	Yes[3] (-1)	PBT boost + photon vs. photon alone: • Grade ≥2: 9% (6/66) vs. 14% (9/66), p=NR • Grade 3: 0% (0/66) vs. 7.5% (5/66), p<0.1	No statistical differences between groups; unclear if differences may be clinically important. Sample size may have played a role in these findings. ⊕⊕◯◯ LOW
	During CSI	Gunther 2017 (N=37) Retro cohort CNS-involvement in leukemia/ lymphoma Salvage therapy (pre-SCT RT)	Yes[1] (-1)	No	No	Yes[3] (-1)	PBT vs. Photon • Mucositis, Grade 3: 7% (1/14) vs. 9% (2/23), p=0.1 • Mucositis, any Grade: 7% (1/14) vs. 44% (10/23); RR 0.16 (0.02 to 1.15)** • Gastrointestinal (Grade NR): 29% (4/14) vs. 30% (7/23), p=1.0 • CNS (Grade NR): 21% (3/14) vs. 13% (3/23), p=0.65	PBT resulted in a lower frequency of mucositis (any grade); no other differences were seen over acute or late term. Sample size may have played a role in these findings. ⊕◯◯◯ INSUFFICIENT

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon* Effect estimate (95% CI)	Conclusion Quality (SoE)
	"Late"						• Severe CNS neurotoxicity††: 7% (1/14) vs. 0% (0/23), p=NS	
Radiation necrosis (outside of treatment field)	Median 15 mos.	Adeberg 2017 N=132) Retro case-matched cohort Primary Glioblastoma (high-grade)	No	Unknown	No	Yes[3] (-1)	PBT boost + photon vs. photon alone: 0% (0/66) vs 0% (0/66)	No cases of radiation necrosis outside the treatment field in either group. Sample size may have played a role in the findings. ⊕⊕◯◯ LOW
Change in symptomology, % (n/N)	Median 15 mos.	Adeberg 2017 N=132) Retro case-matched cohort Primary Glioblastoma (high-grade)	No	Unknown	No	Yes[3] (-1)	PBT boost + photon vs. photon alone: <u>Neurocognitive deficits</u>‡‡ • Worse (vs. baseline): 3% (2/66) vs. 6% (4/66) • New: 9% (6/66) vs. 2% (2/66) <u>Sensorimotor deficits</u>‡‡ • Worse (vs. baseline): 3% (2/66) vs. 5% (3/66) • New: 11% (7/66) vs. 14% (9/66) <u>Seizures</u>‡‡ • Worse (vs. baseline): 0% (0/66) vs. 0% (0/66) • New: 2% (1/66) vs. 6% (4/66) p=NS for all	No statistical differences between groups in the proportion of patients experiencing either worsening of preexisting symptoms or new deficits following treatment ⊕⊕◯◯ LOW

CNS = central nervous system; CI = confidence interval; CSI = craniospinal irradiation; KQ = Key Question; NR = not reported; NS = not statistically significant; PBT = proton beam therapy; Retro = retrospective; SCT = stem cell transplantation; SOE = strength of evidence.

* **Adeberg 2017**: Photon + PBT boost vs. Photon alone.
Gunther 2017: PBT (passive scatter) vs. Photon.
Jhaveri 2018: PBT vs. photons (IMRT, 3D-CRT, and other photon not specified).

†All data estimated from graphs provided by authors.

‡Estimated from graph in article.

§Also had concurrent systemic relapse and died from disease.
**Crude RR calculated by AAI using exact methods in Stata.
††Characterized by diffuse demyelination and necrosis, neurocognitive impairment, lower extremity weakness, incontinence, difficulty swallowing
‡‡ Authors describe these as/along with toxicity. As baseline in the PBT vs. photon groups, neurocognitive deficits, sensorimotor deficits, and seizures were presents in 30% (20/66) vs. 42% (28/66), 39% (26/66) vs. 30% (20/66), and 6% (4/66) vs. 3% (2/66), respectively. The majority of patients with pre-therapeutic deficits showed a stable deficit level after radiotherapy.

Reasons for downgrade:
1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.2 Strength of Evidence Summary for Adult Breast Cancer for Effectiveness

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon/Electron Boost* Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)								
Survival outcomes								
Probability, overall survival (OS)	5 years	Chowdhary 2019 (N=724,492) Retro comparative database study (NCDB)	No	Unknown	No	No	91.9% vs. 88.9% (unadjusted probabilities) Adjusted HR† 0.85 (95% CI, 0.68 to 1.07), p=0.12 A second additional multivariate analysis conducted after stratifying for factors associated with increase heart doses also showed no difference.	No statistical difference between PBT versus photon/electron boost therapy for the probability of OS at 5 years. ⊕⊕◯◯ LOW

CI = confidence interval; KQ = Key Question; NCDB = National Cancer Data Base; PBT = proton beam therapy; Retro = retrospective; SOE = strength of evidence.
*Aside from the breast, additional lymph node irradiation was indicated in 22% of patients. Other treatments received included chemotherapy in 46% and endocrine therapy in 69%.

†In multivariate analysis, adjusted for: race, Charlson-Deyo comorbidity score, facility (academic vs. nonacademic), household income, regional location, residence (urban vs. rural), laterality, pT-stage, pN-stage, receptor status, receipt of chemotherapy, receipt of endocrine therapy, type of surgery, and year of diagnosis.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.3 Strength of Evidence Summary for Adult Esophageal Cancer for Effectiveness

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
Probability, overall survival (OS)	1 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	88% vs. 85%‡ Log-rank, p=0.01	Probabilities of OS at 1 year were similar, however, over subsequent years OS was better following PBT vs. IMRT or 3DCRT across both studies. However, statistical significance was achieved in only the largest study. ⊕⊕◯◯ LOW
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	80% vs. 78%‡ Log-rank, p=0.10	
	2 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	70% vs. 50%‡ Log-rank, p=0.01	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	66% vs. 49%‡ Log-rank, p=0.10	
	3 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	55% vs. 39%‡ Log-rank, p=0.01	
		Fang 2018 (N=133, stage III/IV subanalysis only)	No	No	No	Yes[3] (-1)	48% vs. 38%‡ Log-rank, p=0.10	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
		Retro propensity-score matched cohort AC (74%) or SCC (26%)						
	4 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	44% vs. 35%‡ Log-rank, p=0.01	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	42% vs. 30%‡ Log-rank, p=0.10	
	5 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	41.6% vs. 31.6%; adj. HR 1.45 (1.09 to 1.94) p=0.010 *Stage III only*: 34.6% vs. 25.0%, p=0.04	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	42% vs. 19%; adj. HR 1.48 (0.93 to 2.35), p=0.10 *All patients*: HR 0.82 (0.56 to 1.20), p=0.30	
Mortality, % (n/N)	3 months	Lin 2017 (N=580) Retro cohort AC (92%) or SCC (8%) Stage III/IV (63%)	No	Unkown	No	Yes[3] (-1)	1 mo. post-op: 0% vs. 1.5% (7/469), p=0.425 2 mos. post-op: 0.9% (1/111) vs. 2.6% (12/469), p=0.59 3 mos. post-op: 0.9% (1/111) vs. 4.3% (20/469), p=0.26	No statistically differences; per authors, the difference at 3 months may be clinically meaningful. ⊕⊕◯◯ LOW

Final

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
	Median 22 months	Makishima 2015 N=44 SCC (100%) Stage III (52%); Stage I/II (48%)	Yes[1] (-1)	Unkown	No	Yes[3] (-1)	20% (5/25) vs. 31.6% (6/19), p=NR	No statistically significant differences; sample sizes are small. ⊕◯◯◯ INSUFFICIENT
Probability, Progression-free survival (PFS) or Disease-free survival (DFS)	1 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	PFS: 62% vs. 50%, p=0.001	At all timepoints, PFS/DFS was better following PBT vs. IMRT or 3DCRT across both studies. However, statistical significance was achieved in only the largest study. ⊕⊕◯◯ LOW
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	DFS: 55% vs. 45%, p=0.11	
	2 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	PFS: 50% vs. 33%, p=0.001	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	DFS: 45% vs. 26%, p=0.11	
	3 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	PFS: 42% vs. 28%, p=0.001	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	DFS: 41% vs. 23%, p=0.11	
	4 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	PFS: 39% vs. 24%, p=0.001	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	DFS: 41% vs. 23%, p=0.11	
	5 year	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	PFS: 34.9% vs. 20.4%; adj. HR 1.56 (95% CI 1.19-2.05), p=0.001 *Stage III*: 33.5% vs. 13.2%, p=0.005	
		Fang 2018 (N=133, stage III/IV subanalysis only) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	DFS: 41% vs. 18%, adj. HR 1.42 (0.92 to 2.19) p=0.11	

adj. = adjusted; 3D-CRT = 3-dimensional conformal radiation therapy; AC = adenocarcinoma; CI = confidence interval; HR = hazard ratio; KQ = Key Question; PBT = proton beam therapy; IMRT = intensity-modulated radiation therapy; NS = not statistically significant; Retro = retrospective study design; SCC = squamous cell carcinoma; XRT = X-ray radiation therapy.

*__Fang 2018__: PBT vs. IMRT

__Lin 2017__: PBT vs. IMRT and vs. 3D-CRT

Makishima 2015: passive scatter PBT vs. XRT
Shiraishi 2018: PBT vs. IMRT
Xi 2017: PBT vs. IMRT
†If no 95% CI is provided in the table, the authors did not report one; log-rank p-values.
‡Estimated from graphs in articles.

Reasons for downgrade:
1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.4 Strength of Evidence Summary for Adult Esophageal Cancer for Safety

Outcome	Time	Studies, Year, N, RoB Tumor Indication	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. various photon Effect estimate (95% CI)	Conclusion Quality (SoE)
RT-related toxicities								
Radiation pneumonitis, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%) Stage III (66%); Stage I/II (34%)	No	No	No	No	PBT vs. IMRT: 1.5% (2/132) vs. 2.8% (6/211), p=NS	For PBT versus IMRT, with the exception of grade 4 radiation-induced lymphopenia (2 studies) and wound events (1 study) which were less common with PBT, the frequency of all other RT-related and treatment-related toxicities and adverse events did not differ statistically between groups. For PBT vs. 3DCRT or XRT, with the exception of GI events, PBT was associated with a statistically lower frequency of any treatment-related toxicity (i.e., pulmonary, cardiac, and wound events; grades ≥2 or not specified) across three studies. There were no differences in the frequency of grade ≥3 radiation pneumonitis and pleural effusion between PBT vs. XRT in one small study. ⊕⊕◯◯ LOW
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	Yes[1] (-1)	No	No	Yes[3] (-1)	PBT vs. XRT: 0% (0/25) vs. 5.3% (1/19), p=NS	
Radiation esophagitis, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%)	No	No	No	No	PBT vs. IMRT: 11.4% (15/132) vs. 14.2% (30/211), p=NS	
Radiation induced lymphopenia, grade 4	Acute (during RT; timing NOS)	Fang 2018 (N=220) Retro propensity-score matched cohort AC (74%) or SCC (26%)	No	No	No	Yes[3] (-1)	PBT vs. IMRT: 31% (34/110) vs. 47% (52/110); adj. OR 0.47 (0.26 to 0.84) p=0.01	
	Acute (during RT; timing NOS)	Shiraishi 2018 (N=272) Retro propensity-score matched cohort AC (97%) or SCC (3%)	No	No	No	Yes[3] (-1)	PBT vs. IMRT: 17.6% (24/136) vs. 40.4% (55/136); adj OR 0.29 (0.16 to 0.52) p<0.0001	
Treatment-related toxicity*								

Outcome	Time	Studies, Year, N, RoB Tumor Indication	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. various photon Effect estimate (95% CI)	Conclusion Quality (SoE)

Esophageal fistula, Esophageal stricture, grade ≥3	NR	Xi 2017
		(N=343)
		Retro cohort
		AC (71%) or SCC (29%)
			No	No	No	No	PBT vs. IMRT:
							• Esophageal fistula: 0% (0/132) vs. 1.4% (3/211)
							  • Grade 5: 0% (0/132) vs. 0.5% (1/211)
							• Esophageal stricture: 9.8% (13/132) vs. 8.1% (17/211)
							  • Grade 5: 0% (0/132) vs. 0.5% (1/211)
							p=NS for all	

Any pulmonary event	Acute (post-op)†	Lin (2017) (N=580) Retro cohort AC (92%) or SCC (8%)	No	No	No	Yes[3] (-1)	*Grade NR* PBT: 16.2% (18/111) IMRT: 24.2% (62/255) 3DCRT: 39.5% (85/214) • PBT *vs. IMRT*: adj. OR 0.58 (95% CI 0.32 to 1.05), p=0.08 • PBT *vs. 3D-CRT*: adj. OR 0.34 (95% CI 0.19 to 0.61), p<0.001	
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	Yes[1] (-1)	No	No	Yes[3] (-1)	PBT vs. XRT: Grade ≥2: 0% (0/25) vs. 42.1% (8/19), p<0.001	
Pleural effusion, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%)	No	No	No	No	PBT vs. IMRT: 0.8% (1/132) vs. 1.9% (4/211), p=0.19	
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	Yes[1] (-1)	No	No	Yes[3] (-1)	PBT vs. XRT: 0% (0/25) vs. 5.3% (1/19), p=NS	

Final

Outcome	Time	Studies, Year, N, RoB Tumor Indication	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. various photon Effect estimate (95% CI)	Conclusion Quality (SoE)
Any cardiac event	Acute (post-op)†	Lin (2017) (N=580) Retro cohort AC (92%) or SCC (8%)	No	No	No	Yes[3] (-1)	*Grade NR* PBT: 11.7% (13/111) IMRT: 11.7% (30/255) 3DCRT: 27.4% (59/214) • PBT *vs. IMRT*: adj. OR 0.87 (95% CI 0.42 to 1.77), p=0.70 • PBT *vs. 3D-CRT*: adj. OR 0.34 (95% CI 0.17 to 0.66), p=0.002	
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	Yes[1] (-1)	No	No	Yes[3] (-1)	PBT vs. XRT: Grade ≥2: 4% (1/25) vs. 52.6% (10/19) , p<0.001 RR 0.08 (0.01 to 0.54)‡	
Pericardial effusion, grade ≥3	NR	Xi 2017 (N=343) Retro cohort AC (71%) or SCC (29%)	No	No	No	No	PBT vs. IMRT: 0.8% (1/132) vs. 2.4% (5/211), p=0.19	
	Late	Makishima 2015 (N=44) Retro cohort SCC (100%)	Yes[1] (-1)	No	No	Yes[3] (-1)	PBT vs. XRT: 0% (0/25) vs. 0% (0/19), p=NS	
Any GI event, any wound event	Acute (post-op)†	Lin (2017) (N=580) Retro cohort AC (92%) or SCC (8%)	No	No	No	Yes[3] (-1)	*Grade NR* <u>GI event</u> PBT: 18.9% (21/111) IMRT: 23.0% (59/255) 3DCRT: 20.9% (45/214) Chi Squared p-value: p=0.656 <u>Wound event</u> PBT: 4.5% (5/111) IMRT: 14.1% (36/255) 3DCRT: 15.3% (33/214)	

Outcome	Time	Studies, Year, N, RoB Tumor Indication	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. various photon Effect estimate (95% CI)	Conclusion Quality (SoE)
							• PBT *vs. IMRT*: adj. OR 0.28 (95% CI 0.11 to 0.73), p=0.009 • PBT vs. *3D-CRT*: OR 0.26 (95% CI 0.10 to 0.68), p=0.006	
Readmission within 60 days or death during same hospitalization	2 mos.†	Lin (2017) (N=580) Retro cohort AC (92%) or SCC (8%)	No	No	No	Yes[3] (-1)	PBT: 17.1% (19/111) IMRT: 15.6% (40/255) 3DCRT: 23.7% (51/214) Chi Squared p-value: p=0.070	

3D-CRT: 3-dimensional conformal radiation therapy; AC: adenocarcinoma; CI: confidence interval; PBT: proton beam therapy; IMRT: intensity-modulated radiation therapy (photons); NOS: not otherwise specified; NS: not statistically significant; OR: odds ratio; post-op: post-operative; Retro: retrospective study design; RR: risk ratio; SCC: squamous cell carcinoma; XRT: X-ray radiation therapy.

*Not directly stated by authors as related to RT – called "treatment-related"; because all patients were receiving concurrent or adjuvant chemotherapy is it unclear the degree to which PBT directly affected these outcomes.

†All patients in the study were treated with neoadjuvant concurrent chemotherapy and radiation therapy followed by surgical resection (most commonly esophagectomy 84%); follow-up period post-op is unclear though appears to be up to 3 months. Postoperative complications were identified from hospital notes, discharge summary, and/or from a prospectively collected surgical database.

‡Crude RR calculated by AAI.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.5 Strength of Evidence Summary for Adult Gastrointestinal (Pancreas) Cancer for Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT (spot scanning) vs. HART Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)								
Probability, overall survival (OS)	1-3 years	Maemura 2017 (N=25) Retro cohort Adenocarcinoma (locally advanced and unresectable)	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	1-year: 80% vs. 86.7% 2-year: 45% vs. 33.3% 3-year: 22.5% vs. 26.6% p=NS at all timepoints	No statistically significant differences were seen between PBT and HART for any primary outcome (OS, disease control, local progression, and metastasis) or for any acute RT-related toxicity (hematological and non-hematological); clinical importance of differences is unclear. The sample size was very small. ⊕◯◯◯ INSUFFICIENT
Disease control, % (n/N)	NR						80% (8/10) vs 93% (14/15), p=NR; RR 0.86 (0.61 to 1.20)*	
Local progression, % (n/N)	NR						40% (4/10) vs 60% (9/15), p=NR; RR 0.60 (0.26 to 1.39)*	
Metastasis, % (n/N)	NR						Any: 30% (3/10) vs. 20% (3/15) • Lung: 10% (1/10) vs 0% (0/15) • Liver: 30% (3/10) vs 6.7% (1/15) • Peritoneum: 10% (1/10) vs 13.3% (2/15) p=NR	
Safety (KQ3) (Curative intent only)								
Acute Toxicity (≤3 mos.)	NR	Maemura 2017 (N=25) Retro cohort Adenocarcinoma (locally advanced and unresectable)	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	RT-related Toxicities, % (n/N) <u>Hematological</u> Leukopenia • Grade 2: 10% (1/10) vs. 13% (2/15) • Grade 3: 0% (0/10) vs. 20% (3/15) Thrombocytopenia: • Grade 2: 10% (1/10) vs. 20% (3/15)	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT (spot scanning) vs. HART Effect estimate (95% CI)	Conclusion Quality (SoE)
							• Grade 3: 0% (0/10) vs. 6.7% (1/15) Neutropenia; Anemia: • Grade 2 or 3: 0% (0/10) vs. 0% (0/15) <u>Non-hematological</u> Ulcer: • Grade 2: 10% (1/10) vs 0% (0/15) • Grade 3: 10% (1/10) vs 0% (0/15) Nausea: • Grade 2: 0% (0/10) vs. 7% (1/15) • Grade 3: 0% (0/10) vs. 0% (0/15) Anorexia: • Grade 2: 0% (0/10) vs. 20% (3/15) • Grade 3: 0% (0/10) vs. 0% (0/15) Malaise • Grade 2 or 3: 0% (0/10) vs. 0% (0/15) *No grade 4 toxicities occurred in either group*	

CI = confidence interval; HART = Hyper-fractionated accelerated RT; KQ = Key Question; NR = not reported; PBT = proton beam therapy; Retro = retrospective study design; RR = risk ratio; SOE = strength of evidence.
*Crude RR calculated by AAI.

Reasons for downgrade:
1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).

3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.6 Strength of Evidence Summary for Adult Head and Neck Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)								
Survival outcomes								
Probability, overall survival (OS)	1-year	Romesser 2016 (N=41) Retro cohort Salivary gland cancer (primary or metastasis)	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	83.3% vs. 93.3%, p=0.08	Regardless of tumor types, no statistically significant differences were seen between PBT and IMRT in the probability of 1-3 year OS (2 studies) or 3-year PFS (1 study) or in the incidence of all-cause mortality (1 study). Clinical significance of differences is unclear. ⊕⊕◯◯ LOW for primary oropharyngeal and nasopharyngeal cancer ⊕◯◯◯ INSUFFICIENT for salivary cancer (primary or metastatic)
	3-years	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	94.3% vs. 89.3%; adj. HR 0.55 (95% CI 0.1 to 2.5), p=0.44	
Probability, progression free survival (PFS)	3-years	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	86.4% vs. 85.8%; adj. HR 1.0 (95% CI 0.4 to 2.6), p=0.99	
All-cause mortality, % (n/N)	Median 24 mos.	Holliday 2015 (N=30) Retro case-matched cohort	No	Unknown	No	Yes[3] (-1)	10% (1/10) vs. 5% (1/20), p=NS	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
		Nasopharyngeal cancer (primary)						
Safety (KQ3) (Curative intent only)								
Toxicities and other adverse events								
Acute toxicity grade ≥3	≤3 mos.	Romesser 2016 (N=41) Retro cohort Salivary gland cancer (primary or metastasis)	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	• Dermatitis: 27.8% (5/18) vs. 34.8% (8/23) • Mucositis: 0% (0/18) vs. 8.7% (2/23) • Nausea, Dysphagia, Dysgeusia, Fatigue: no events in either group p=NS for all no *grade 4* events in either group	There were no statistically significant differences in the frequency of grade ≥3 acute or late toxicities following PBT versus IMRT across three studies. Clinical significance of differences is unclear. Sample size and residual confounding and/or tumor type and stage may have played a role in some of these findings. ⊕⊕◯◯ LOW for acute (based on highest quality studies) and late toxicity
		Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• Dermatitis: Data NR, p=0.15 • Mucositis: Data NR, p=0.90 • Weight loss (>20% vs. baseline): 8.3% (4/48) vs. 13.5% (13/98); adj OR 0.64 (95 CI 0.19 to 2.11) • Fatigue (grade 2 or 3): 40.8% (20/49) vs. 36.2% (34/94); adj OR 1.1 (95% CI 0.53 to 2.27) • Xerostomia (grade 2 or 3): 42% (21/50) vs. 61.2% (60/98); adj OR 0.38 (95% CI 0.18 to 0.79)	
		Holliday 2015 (N=30) Retro case-matched cohort Nasopharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• Any Grade 3 event: 50% (5/10) [9 events] vs. 90% (18/20) [30 events]; RR 0.56 (95% CI 0.29 to 1.05)† • Dermatitis (Grade 3): 40% (4/10) vs. 25% (5/20); RR 1.6 (0.55 to 4.68)†	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
							• Any Grade 4/5 events: 0% vs. 0% • Swallowing dysfunction: 0% (0/10) vs. 15% (3/20), p=0.175 • Mean percentage (IQR) body weight lost from pre to post RT: 5.7% (4.5% to 11.2%) vs. 7.6% (6.1% to 12.1%), p=0.333	
Late toxicity grade ≥3	1 year	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• Weight loss (>20% vs. baseline): 6.7% (3/45) vs. 19.3% (17/88); adj OR 0.28 (95 CI 0.08 to 1.05) • Fatigue (grade 2 or 3): 14.6% (7/48) vs. 22.1% (17/77); adj OR 0.5 (95% CI 0.18 to 1.36) • Xerostomia (grade 2 or 3): 42% (21/50) vs. 47.2% (42/89); adj OR 0.63 (95% CI 0.30 to 1.33)	
	NR (median 24 mos.)	Holliday (N=30) Retro case-matched cohort Nasopharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• Any Grade 3 event: 30% (3/10) [5 events] vs. 15% (3/20) [3 events]; RR 2.0 (95% CI 0.49 to 8.18)†	
Gastrostomy tube dependence	Acute	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• ≤3 months: 12% (6/50) vs. 23% (23/100); adj OR 0.43 (95% CI 0.16 to 1.17)	GT dependence tended to be lower with PBT, however adjusted estimates from the largest study were not statistically significant, while smaller studies in

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
		Holliday (N=30) Retro case-matched cohort Nasopharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• During or after RT: 20% (2/10) vs. 65% (13/20), p=0.02; adj. OR 9.33 (95% CI 1.74 to 75.96), p=0.008	different cancer types reported statistically significant differences. For the smallest study, the large confidence interval suggest instability of the effect estimate. Clinical significance of differences is unclear. It is unclear what role differences in study populations (including tumor characteristics, etc.) and possible residual confounding may play in these findings. ⊕⊕◯◯ LOW
		McDonald 2016 (N=40) Retro comparative cohort Nasopharynx, nasal cavity or paranasal sinus cancers (primary)	No	Unknown	No	Yes[3] (-1)	• End of RT: adj. OR 0.03 (95 % CI <0.01 to 0.15), p<0.001 • 1 month post-RT: adj. OR 0.11 (95% CI <0.01 to 0.61), p=0.028	
		Romesser 2016 (N=41) Retro cohort Salivary gland cancer (primary or metastasis)	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	• ≤3 months: 0% vs. 0% (reactive gastrostomy tube or tracheostomy)	
	Late	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• 1 year: 2% (1/50) vs. 7.8% (7/90); adj OR 0.16 (95% CI 0.02 to 1.37)	
		Sharma 2018 (N=64) Prospective cohort Oropharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• 6 months: 0% vs. 0%	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. IMRT Effect estimate (95% CI)	Conclusion Quality (SoE)
ED visit or hospital-ization	During RT	Blanchard 2016 (N=150) Retro case-matched cohort Oropharyngeal cancer (primary)	No	Unknown	No	Yes[3] (-1)	• ED visit: 32%(16/50) vs. 32% (32/100); adj. OR 0.95 (95% CI 0.45 to 2.0) • Unscheduled hospitalization: 20% (10/50) vs. 21% (21/100); adj OR 0.92 (95% CI 0.39 to 2.2)	No statistically significant differences in the frequency of ED visits or unplanned hospitalizations following PBT versus IMRT. ⊕⊕◯◯ LOW
Osteoradio-necrosis	Median 34 mos.	Zhang 2017 (N=584) Retro cohort Oropharyngeal cancer (primary)	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	Any grade: 2% (1/50) vs. 7.7% (41/534); RR 0.26 (0.04 to 1.85)† • Grade 3: 0% (0/50) vs. 0.9% (5/534) • Grade 4: 0% (0/50) vs. 2.2% (12/534) • Grade 3 or 4: 0% (50) vs. 3.2% (17/534) p=NS for all	No statistically significant differences in the frequency of osteoradionecrosis following PBT versus IMRT. The small number of patients for PBT may preclude identification of rare events and residual confounding may have played role in some of these findings. ⊕◯◯◯ INSUFFICIENT

adj. = adjusted; CI = confidence interval; ED = emergency department; HR = hazard ratio; KQ = Key Question; OR = odds ratio; PBT = proton beam therapy; IMRT = intensity-modulated radiation therapy; NS = not statistically significant; Retro = retrospective study design; RT = radiation therapy.

* **Blanchard 2016:** intensity modulated spot-scanning PBT vs. IMRT
Holliday 2015: intensity modulated spot-scanning PBT vs. IMRT
McDonald 2016: 3D conformal PBT vs.IMRT
Romesser 2016: Uniform scanning-beam PBT vs. IMRT
Sharma 2018: Adjuvant pencil beam scanning PBT vs. IMRT via volumetric modulated arc therapy (VMAT) following transoral robotic surgery and selective neck dissection
Zhang 2017: intensity modulated spot-scanning PBT vs. IMRT

†Crude RR calculated by AAI. The small number of patients for PBT may preclude identification of rare events.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.7 Strength of Evidence Summary for Skull-base Head and Neck Cancer for Effectiveness

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Surgery + adjuvant PBT vs. Surgery alone RR (95% CI)*	Conclusion Quality (SoE)
Curative intent (KQ1)								
Survival and tumor control outcomes								
Probability, disease-specific survival (DSS)	5-, 10-years	Simon 2018 N=47 (n=34 petroclival only) Retro comparative cohort Chondro-sarcoma (grade II)	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	*5- and 10-year DSS in:* **All patients** 100% vs. 89.8% (76.2% to 100%), p=0.138 **Petroclival patients only** 100% vs. 76.4% (46.1% to 100%), p=0.028	The probability of PFS, but not DSS, at 5 and 10 years was statistically better following surgery with adjuvant PBT versus surgery alone. PBT resulted in improved DSS and PFS at both time points for the subgroup of patients with petroclival tumors. Local control was statistically better following adjuvant PBT. ⊕◯◯◯ INSUFFICIENT
Probability, progression-free survival (PFS)	5-, 10-years						**All patients** • *5-year*: 100% vs. 67.8% (47.7% to 88.0%) • *10-year*: 87.5% (64.6% to 100%) vs. 58.2% (33.5% to 82.8%) p=0.006 **Petroclival patients only** • *5-year*: 100% vs. 50% (15.4% to 84.6%) • *10-year*: 85.7% (59.8% to 100%) vs. 50.0% (15.4% to 84.6%)	

Final

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Surgery + adjuvant PBT vs. Surgery alone RR (95% CI)*	Conclusion Quality (SoE)
							p=0.001	
Proportion of patients experiencing local relapse, or regional or distant metastases% (n/N)	Median 7.5 years						**Local relapse:** 4.3% (1/23) vs. 33% (8/24); RR 0.13, 95% CI 0.02 to 0.96, p=0.01 (5/9 patients went on to receive secondary proton therapy) **Regional or distant metastases:** 0% vs. 0%	
Safety								
Any complication, % (n/N)	Median 7.5 years	Simon 2018 N=28 for PBT and 47 for surgery† Retro comparative cohort Chondro-sarcoma (grade II)					68% (19/28) vs. 26% (12/47), RR 2.7 (1.5 to 4.6)	Unadjusted estimates of treatment-related death and severe complications (grade ≥3 toxicity) did not differ statistically between groups, however, patients who received adjuvant PBT had a higher risk of experiencing any complication, specifically sensorineural and severe hearing loss. However, confidence intervals were wide suggesting instability of the effect estimate. ⊕◯◯◯ INSUFFICIENT
Any grade ≥3 toxicity, % (n/N)							25% (7/28) vs. 11% (5/47), p=0.10	
Treatment-related death, % (n/N)							0% (0/28) vs. 2% (1/47), p=0.44	
Hearing loss and dizziness, % (n/N)							**Sensorineural hearing loss:** 39% (11/28) vs. 6% (3/47), RR 6.2 (1.9 to 20.2) **Severe hearing loss:** 21% (6/28) vs. 4% (2/47), RR 5.0 (1.1 to 23.3) **Conductive hearing loss:** 11% (3/28) vs. 4% (2/47), p=0.28 **Dizziness**: 14% (4/28) vs. 0% (0/47), p=0.008	
Other complications from PBT, % (n/N)							**Vision loss:** 11% (3/28) **Hypopituitarism**: 18% (5/28) **Temporal lobe necrosis**: 18% (5/28)	

CI = confidence interval; KQ = Key Question; NCDB = National Cancer Data Base; PBT = proton beam therapy; Retro = retrospective; SOE = strength of evidence.
*Crude RRs and 95% CIs were calculated by AAI.
†All patients were included in evaluation of complications due to surgery and 28 total patients were included in the evaluation of complications due to PBT (23 primary treatment and 5 secondary PBT treatment follow-up local relapse).

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.8 Strength of Evidence Summary for Adult Liver Tumors for Efficacy and Safety

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT (passive scatter) vs. TACE (RCT) or vs. IMRT (Observational study) Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)								
Randomized controlled trial								
Probability, overall survival (OS)	2-year	Bush 2016 (N=69) RCT Moderately low RoB HCC	No	Unknown	No	Yes[3] (-1)	• All patients: 59% (NR) • Patients receiving liver transplant post-treatment (n=22): 82% (NR) p=NS for both, data not provided	No significant difference between groups in the probability of 2-year OS; patients who received PBT tended to have improved probability of 2-year PFS and local tumor control compared with TACE patients, although the difference did not reach statistical significance. Results are from interim analysis of an ongoing trial. ⊕⊕⊕◯ MODERATE
Probability, progression free survival (PFS)	2-year						48% (NR) vs. 31% (NR); p=0.06	
Probability, local control (LC)	2-year						88% (NR) vs. 45% (NR); p=0.06	
Observational study								
Probability, overall survival (OS)	2-year	Sanford 2019 (N=133) Retrospective cohort study Moderately high RoB HCC	No	Unknown	No	Yes[3] (-1)	59.1% vs. 28.6%; adj. HR 0.47 (95% CI 0.27 to 0.82)	OS was significantly higher following PBT vs. IMRT but there was no difference in local and regional control between groups. ⊕⊕◯◯ LOW
Probability, local and locoregional control	2-year						**Local control** (cumulative incidence): 93% (NR) vs. 90% (NR); HR for cumulative incidence of *local failure* 0.74 (95% CI 0.18 to 3.01)	

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT (passive scatter) vs. TACE (RCT) or vs. IMRT (Observational study) Effect estimate (95% CI)	Conclusion Quality (SoE)
							Locoregional recurrence (cumulative incidence): 53% vs. 42%; adjusted HR 0.98 (95% CI 0.54 to 1.75).	
Safety (KQ3) (Curative intent only)								
Randomized controlled trial								
Acute Toxicity (≤3 mos.)	NR	Bush 2016 (N=69) RCT Moderately low RoB HCC	No	Unknown	No	Yes[3] (-1)	Acute toxicity was generally limited to the following, which were experience by most patients (no data provided): • PBT: fatigue and radiation skin reaction • TACE: abdominal pain and nausea Authors state that serious complications from PBT were uncommon events (no data provided).	Limited information provided on acute toxicity. Significantly fewer patients who received PBT required hospitalization in the month following treatment compared with TACE patients; total days hospitalized was also significantly less in the PBT vs. the TACE group. Results are from interim analysis of an ongoing trial. ⊕⊕⊕◯ MODERATE
Proportion of patients hospitalized for an acute complication, % (n/N)	≤1 mo.						6.1% (2/33) vs. 41.7% (15/36); p<0.001	
Total days hospitalized within 1 month of treatment	≤1 mo.						**Overall**: 24 (0.73 days per patient) vs. 166 (4.6 days per patient); p<0.001 • *for routine observation*: 0 vs. 53 • *for complications:* 24 vs. 113	

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT (passive scatter) vs. TACE (RCT) or vs. IMRT (Observational study) Effect estimate (95% CI)	Conclusion Quality (SoE)
Observational study								
Incidence of nonclassic radiation-induced liver disease (RILD)*	3 mos.	Sandford 2019 (N=100)† Retrospective cohort study Moderately high RoB HCC	No	Unknown	No	Yes[3] (-1)	adj. OR 0.26 (95% CI 0.08 to 0.86) (PBT, n=4 patients; IMRT, n=17 patients) Authors also report that the development of RILD at 3 months was associated with significantly worse OS (HR 3.83; 95% CI 2.12 to 6.92).	Lower risk of RILD in the acute period with PBT versus IMRT ⊕⊕◯◯ LOW
Death due to liver failure	NR (median f/u 14 mos.)	Sandford 2019 (N=36)‡ Retrospective cohort study Moderately high RoB HCC	No	Unknown	No	Yes[3] (-1)	53% (8/15) vs. 91% (19/21); RR 0.59 (95% CI 0.36 to 0.97)§	Lower risk of death due to liver failure with PBT versus IMRT; however data was from a small subset of patients. ⊕◯◯◯ INSUFFICIENT

HCC = hepatocellular carcinoma; IMRT = intensity-modulated radiation therapy (photons); NR = not reported; PBT = proton beam therapy; RCT = randomized controlled trial; RoB = risk of bias; SOE = strength of evidence; TACE = Transarterial chemoembolization

*RILD was defined as worsening of Child-Pugh score by ≥2 points compared with baseline. At baseline, patients treated with photons had worse baseline child-Pugh score (median 6 vs. 5, p=0.008), however, this variable was included in and controlled for via multivariate analyses.

†RILD was calculated in 100 (of 133) patients for whom data was available; denominators for this subset of patients by treatment group were not provided.

‡Death due to liver failure was reported only among the 36 patients (15 PBT, 21 IMRT) without disease progression.

§RR and 95% CI calculated by AAI.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.9 Strength of Evidence Summary for Adult Lung Cancer for Efficacy/Effectiveness and Safety

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
Curative intent (KQ1)								
Randomized controlled trials								
Probability, overall survival (OS)‡	1-5 year	Liao 2018 N=173 (ITT) RCT NSCLC	No	Unclear	No	Yes[3] (-1)	• 1-year: 75% vs. 82% • 2-year: 56% vs. 60% • 3-year: 26% vs. 37% • 4-year: 38% vs. 32% • 5-year: 24% vs. 32% p=0.30	No statistically significant differences between groups in the probability of OS or the cumulative incidence of local failure at any timepoint measured. ⊕⊕⊕◯ MODERATE
Cumulative incidence of local failure (%)‡							• 1-year: 9% vs. 10% • 2-year: 27% vs. 26% • 3-year: 37% vs. 37% • 4-year: 37% vs. 32% • 5-year: 37% vs. 39% p=0.99	
Observational studies								
Probability, overall survival (OS)	1-year	Liao 2018§ N=39 Pro cohort NSCLC	Yes[1] (-1)	No	No	Yes[3] (-1)	69% vs. 57% p=0.97	No statistically significant differences between groups in the probability of OS over 1-5 years (across 4 studies) or LRFS at 1 or 2 years (1 study) or in the incidence of local failure at 2 or 3 years (2 studies) ⊕⊕◯◯ LOW for OS ⊕◯◯◯ INSUFFICIENT for LRFS and local failure
		Remick 2017 N=61 Retro cohort NSCLC	Yes[1] (-1)	No	No	Yes[3] (-1)	85.2% (72.8%–99.7%) vs. 82.4% (70.5%–96.2%) p=0.65	
		Higgins 2017 N=1850 (propensity-matched) Retro database NSCLC	No	No	No	Yes[3] (-1)	62.0% (56.2%–67.2%) vs. 54.2% (51.6%–56.7%) p=NR	
	2-year	Liao 2018§ N=39 Pro cohort NSCLC	Yes[1] (-1)	No	No	Yes[3] (-1)	43% (NR) vs. 43% (NR) p=0.97	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
		Remick 2017 N=61 Retro cohort NSCLC	Yes[1] (-1)	No	No	Yes[3] (-1)	77.8% (63.6%–95.2%) vs. 73.2% (59.6%–89.9%) p=0.65	
		Tucker 2016 N=468 Retro cohort NSCLC	No	No	No	Yes[3] (-1)	• PBT: 56% (40%–69%) • IMRT: 52% (45%–58%) • 3DCRT: 39% (32%–46%) p=NS, PBT vs. IMRT p=0.015, PBT vs. 3DCRT	
	3-year	Liao 2018§ N=39 Pro cohort NSCLC	Yes[1] (-1)	No	No	Yes[3] (-1)	25% (NR) vs. 32.5% (NR) Log-rank p=0.97	
	5-year	Higgins 2017 N=1850 (propensity-matched) Retro database NSCLC	No	No	No	Yes[3] (-1)	5:1 matching: 22.3% (16.3%–28.9%) vs. 15.7% (13.5%–18.1%) adj. HR 1.18 (95% CI 1.02 to 1.37) *a-priori* 1:1 matching: adj. HR 1.16 (95% CI 0.97 to 1.39)	
Probability, Local Recurrence-Free Survival (LRFS)	1-2 year	Remick 2017 N=61 Retro cohort NSCLC	Yes[1] (-1)	Unclear	No	Yes[3] (-1)	• 1-year: 92.3% (82.5%–100%) vs. 93.3% (84.8%–100%) • 2-year: 93.1% vs. 85.7% p=0.82	
Local Failure	1-2 years	Liao 2018§ N=39 Pro cohort NSCLC	Yes[1] (-1)	Unclear	No	Yes[3] (-1)	Cumulative incidence‡: • 1-year: 6% vs. 3% • 2-year: 6% vs. 3% • 3-year: 26% vs. 26% p=0.93	
	2-years	Remick 2017 N=61 Retro cohort NSCLC	Yes[1] (-1)	Unclear	No	Yes[3] (-1)	11.1% (3/27) vs. 5.9% (2/34), p=NS	

Outcome	Time	Studies, Year, N, Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Photon (various)* Effect estimate (95% CI)†	Conclusion Quality (SoE)
Safety (KQ3) (all curative intent)								
Randomized controlled trials								
Rate of radiation pneumonitis, Grade ≥3‡	1-5 years	Liao 2018 N=173 (ITT) RCT NSCLC	No	Unclear	No	Yes[3] (-1)	8% vs. 7% at 1, 2, 3, 4 and 5 years; p=0.58	No statistically significant differences between groups. ⊕⊕⊕◯ MODERATE
Observational studies								
Radiation esophagitis	NR (median 26 months)	Remick 2017 N=61 Retro cohort NSCLC	Yes[1] (-1)	No	No	Yes[3] (-1)	• Grade 2: 18.5% (5/27) vs. 29.4% (10/34), p=NR • Grade 3: 3.7% (1/27) vs. 11.8% (4/34), p=NR	No statistically significant differences between groups for any grade 3 outcome; however differences may be clinically important. ⊕◯◯◯ INSUFFICIENT
	NR	Niedzielski 2017 N=134 Retro cohort NSCLC	Yes[1] (-1)	No	No	Yes[3] (-1)	• Grade 2: 59.2% (29/49) vs. 54.1% (46/85), p=NS • Grade 3: 22.4% (11/49) vs. 17.6% (15/85); OR 1.4 (0.7 to 2.9), p=0.37	
Radiation pneumonitis	NR (median 26 months)	Remick 2017 N=61 Retro cohort NSCLC	Yes[1] (-1)	No	No	Yes[3] (-1)	• Grade 2: 3.7% (1/27) vs. 8.8% (3/34), p=NR • Grade 3: 3.7% (1/27) vs. 2.9% (1/34), p=NR	
Radiation dermatitis							• Grade 2: 37% (10/27) vs. 12% (4/34), p=NR • Grade 3: 0% (0/27) vs. 0% (0/34), p=NR	

3D-CRT = Three-dimension conformal radiation therapy; adj. = adjusted; CI = confidence interval; HR = hazard ratio; IMRT = intensity-modulated radiation therapy; ITT = intention-to-treat analysis; KQ = Key Question; NR = not reported; NS = not statistically significant; NSCLC = non-small cell lung cancer; RCT = randomized controlled trial; Retro = retrospective study design; Pro = prospective study design.

*__Liao 2018__ (RCT and observational): passive scatter PBT vs. IMRT
__Higgins 2017__: PBT vs. various photon (external beam, 3D-conformal, IMRT, "photons")
__Niedzielski 2017__: passively scattered PBT vs. IMRT
__Remick 2017__: double scatter or pencil beam PBT vs. IMRT
__Tucker 2016__: pencil beam PBT vs. IMRT vs. 3DCRT

†If no 95% CI is provided in the table, the authors did not report one; log-rank p-values.

‡Estimated from figures/graphs in article.
§This cohort is comprised of patients from the RCT who could not be randomized because their PBT or IMRT plans did not allow for random assignment (i.e., did not meet prespecified dose-volume constraints developed for photon radiation); they were followed as an observational cohort.

Reasons for downgrade:
1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.10 Strength of Evidence Summary for Adult Ocular Tumors for Effectiveness and Safety

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Brachytherapy or Stereotactic Radiosurgery* Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)								
Survival and tumor recurrence outcomes								
Probability of overall survival (OS)	2, 5, years	Lin 2017 (N=452) Retro propensity-score matched cohort (NCD) Choroid melanoma	No	Unknown	No	Yes[3] (-1)	• 2-year OS: 93% vs. 97%, p=NS • 5-year OS: 51% vs. 77% • adj. HR for *risk of mortality*: 1.89, 95% CI 1.24 to 2.95	Similar OS/mortality at 2 and 3 years for PBT vs. brachytherapy or SRS in 2 studies of choroid and uveal melanoma. In the larger database study of choroid melanoma, PBT was associated with a statistically higher risk of mortality at 5 years vs. brachytherapy. ⊕⊕◯◯ LOW
Mortality, % (n/N)	3 years	Sikuade 2015 (N=191) Retro cohort Uveal Melanoma	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	13% (14/106) vs. 16% (14/85), p=NS	
Local recurrence	3, 5, 10 years	Böker (2018), N=140 Retro case-matched cohort Large Uveal Melanoma	No	Unknown	No	Yes[3] (-1)	Rate (95% CI) • 3-years: 4% (1.2% to 17.8%) vs. 24.6% (15.8% to 37.1%), p<0.001 • 5-years: 9.1% (2.9% to 27.3%) vs. 27.5% (17.8% to 41.1%), p<0.001 • 10-years: 9.1% (2.8% to 27.3%) vs. 36.5% (20.7% to 59.1%); adj. HR 7.69 (95% CI 2.22 to 26.06) *for brachytherapy*	PBT was associated with a statistically lower frequency of local recurrence over 10 years compared with brachytherapy (+TSR for both). ⊕⊕◯◯ LOW

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT vs. Brachytherapy or Stereotactic Radiosurgery* Effect estimate (95% CI)	Conclusion Quality (SoE)
	Mean3 years	Sikuade 2015 (N=191) Retro cohort Uveal Melanoma	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	2.8% (3/106) vs. 0% (0/85), p=NS	No statistical difference in local recurrence between PBT versus SRS ⊕◯◯◯ INSUFFICIENT
Safety (KQ3) (Curative intent only)								
Frequency of adverse events, % (n/N)	Mean 3.3 years	Böker (2018), N=140 Retro case-matched cohort Large Uveal Melanoma	No	Unknown	No	Yes[3] (-1)	• Enucleation: 8.5% (6 eyes) vs. 15.7% (11 eyes), p=0.196 • Rubeosis of the iris: 1.4% (1/70) vs. 0% (0/70), p=0.316 • Neovascular glaucoma: 1.4% (1/70) vs. 1.4% (1/70), p=NS	With the exception of optic neuropathy which was statistically lower following PBT versus SRS in one study of uveal melanoma, no other statistical differences were seen in the frequency of adverse events over 3 years between PBT versus brachytherapy or SRS. ⊕⊕◯◯ LOW
	Mean 3 years	Sikuade 2015 (N=191) Retro cohort Uveal Melanoma	Yes[1] (-1)	Unknown	No	Yes[3] (-1)	• Enucleation: 1.9% (2/106) vs. 2.4% (2/85), p=NS • Rubeotic glaucoma: 4.7% (5/106)† vs. 11% (9/85)†, p=NS • Radiation retinopathy: 30% (31/106) vs. 24% (20/85), p=NS • Optic Neuropathy: 13% (14/106) vs. 28% (23/85); RR=0.49 (0.27 to 0.89)‡	

adj. = adjusted; CI = confidence interval; HR = hazard ratio; KQ = Key Question; NCD = National Cancer Database; NS = not statistically significant; PBT = proton beam therapy; Retro = retrospective study design; RR = risk ratio; SRS = stereotactic radiosurgery; TSR = transscleral resection.

*__Boker 2018__: Neoadjuvant PBT + TSR vs. Adjuvant Brachytherapy + TSR
__Lin 2017__: PBT vs. Brachytherapy
__Sikuade 2015__: PBT vs. SRS

†*Requiring enucleation:* 1.9% (2/106) [40% (2/5) with rubeotic glaucoma] vs. 2.4% (2/85) [22% (2/9) with rubeotic glaucoma].

‡Calculated by AAI.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.2.11 Strength of Evidence Summary for Adult Prostate Cancer for Effectiveness and Safety

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
Curative intent (KQ1)								
Survival outcomes – *quasi-RCT*								
Probability, overall survival (OS)	5-year	Khmelevsky 2018 quasi-RCT (N=289) Moderately high RoB Prostate Cancer Risk: High (53%), Intermediate (42%), Low (5%)	No	Unclear	No	Yes[3] (-1)	74% ± 5.0% vs. 78.8% ± 4.1%, p=NS	No statistically significant differences between Photon plus PBT boost vs. Photon alone in the probability of 5- and 10-year OS or BRFS ⊕⊕◯◯ LOW
	10-year						55.9% ± 9.0% vs. 60.6% ± 5.7%, p=NS	
Probability, Biochemical Relapse Free Survival (BRFS)	5-year						60% ± 5.4% vs. 61.9% ± 4.4%, p=NS	
	10-year						45.5% ± 8.5% vs. 42.8% ± 7.1%, p=NS	
Safety (KQ3) (curative intent only)								
Quasi-RCT								
GI toxicity, probability	Acute	Khmelevsky 2018 quasi-RCT (N=289) Moderately high RoB Prostate Cancer Risk: High (53%), Intermediate (42%), Low (5%)	No	Unknown	No	Yes[3] (-1)	• **Grade 2**: 54.4% ± 5.4% vs. 69.2% ± 5.7%, p<0.01 • **Grade 3 or 4**: 0% vs. 0%	There were no statistically significant differences in the probabilities of grade 3 or 4 toxicities; however, acute and late Grade 2 GI, but not GU, toxicity, were significantly lower in patients who received the PBT boost versus photons only. The actuarial frequency of grade ≥3 GI and GU toxicities was lower in the PBT boost group but statistical testing was not done. ⊕⊕◯◯
	Late						• **Grade 2**: 10.2% ± 5.5% vs. 34.8% ± 7.4%, p<0.01 • **Grade 3 or 4**: 0.9% ± 1.7% vs. 1.3% ± 1.8%, p=NS	
GU toxicity, probability	Acute						• **Grade 2**: 33.3% ± 4.6% vs. 36.1% ± 3.5%, p=NS • **Grade 3 or 4**: PBT: 0% vs. 1.9% ± 1.8%, p=NS	
	Late						• **Grade 2**: 8.3% ± 5.0% vs. 9.1% ± 4.5%, p=NS • **Grade 3 or 4**: 2.8% ± 2.6% vs. 3.8% ± 3.0%, p=NS	

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
Actuarial frequency of GI and GU toxicities, Grade ≥3	10 years						1.7% vs. 8.7%, p=NR	LOW
Observational studies								
GI toxicity	Acute	Dutz 2019 (N=58) Retro propensity score-matched cohort Prostate Cancer Risk: Low (3%), Intermediate (78%), High (19%)	No	No	No	Yes[3] (-1)	*Proportion of patients* • **Grade 1:** 48% (14/29) vs. 38% (11/29); RR 1.27 (95% CI 0.70 to 2.32)† • **Grade 2:** 14% (4/29) vs. 17% (5/29); RR 0.80 (95% CI 0.24 to 2.68)† • **Grade 3:** 3% (1/29) vs. 0% (0/29), p=0.60	In the two clinical studies, there were no statistical difference between PBT and IMRT in acute or late toxicity (GI or GU). In the large database study, PBT resulted in lower cumulative incidences of any grade GI and GU toxicity and erectile dysfunction compared with IMRT; differences between groups were small and clinical significance is unknown. However, only the incidence of urethral stricture remained significant in a sensitivity analysis using validated diagnosis and procedure codes for severe toxicities post-pelvic radiation. ⊕⊕◯◯ LOW
		Fang 2015 (N=188) Retro case-matched cohort Prostate Cancer Risk: Low (55%), Intermediate (31%), High (7%)	No	No	No	Yes[3] (-1)	*Proportion of patients* • **Grade 0 to 1:** 95.7% (90/94) vs. 86.2% (81/94) • **Grade 2 to 3:** 4.3% (4/94) vs. 13.8% (13/94); adj. OR 0.27 (0.06 to 1.24); p=0.09	
	Late	Dutz 2019 (N=58) Retro propensity score-matched cohort Prostate Cancer Risk: Low (3%), Intermediate (78%), High (19%)	No	No	No	Yes[3] (-1)	*Proportion of patients* • **Grade 1:** 9% (2/22) vs. 27% (6/22); RR 0.33 (95% CI 0.08 to 1.47)† • **Grade 2:** 9% (2/22) vs. 9% (2/22) • **Grade 3:** 5% (1/22) vs. 0% (0/22), p=0.32	
		Fang 2015 (N=188)	No	No	No	Yes[3] (-1)	*Proportion of patients* • **Grade 0 to 1:** 87.2% (82/94) vs. 88.3% (83/94)	

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
		Retro case-matched cohort Prostate Cancer Risk: Low (55%), Intermediate (31%), High (7%)					• **Grade 2 to 3:** 12.8% (12/94) vs. 10.8% (10/94); adj. HR 1.24 (0.53 to 2.94) p=0.62	
		Pan 2018 (N=4158) Retro propensity-score matched database study‡ Prostate Cancer Risk: NR	No	No	No	Yes[3] (-1)	*Cumulative incidence, any bowel toxicity (any grade)* • **6-months**: 1.6% (n=693) vs. 3.2% (n=3465) • **12-months**: 7.4% (n=572) vs. 7.7% (n=2862) • **24-months**: 19.5% (n=341) vs. 15.4% (n=1718) • **36-months**: 24.9% (n=205) vs. 19.2% (n=1003) HR 1.27 (1.05 to 1.55); p=0.02 *Sensitivity analysis* based on validated diagnosis and procedure codes for severe toxicities post-pelvic radiation showed no difference in rectal complications between groups at 24 months (1.5% vs. 2.0%; HR 1.19, 95% CI 0.62 to 2.30)	
GU toxicity	Acute	Dutz 2019 (N=58) Retro propensity score-matched cohort Prostate Cancer Risk: Low (3%), Intermediate (78%), High (19%)	No	No	No	Yes[3] (-1)	*Proportion of patients* • **Grade 1:** 66% (19/29) vs. 45% (13/29); RR 1.46 (95% CI 0.90 to 2.37)† • **Grade 2:** 24% (7/29) vs. 41% (12/29); RR 0.58 (95% CI 0.27 to 1.27)† • **Grade 3:** 3% (1/29) vs. 3% (1/29)	
		Fang 2015	No	No	No	Yes[3] (-1)	*Proportion of patients*	

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
		(N=188) Retro case-matched cohort Prostate Cancer Risk: Low (55%), Intermediate (31%), High (7%)					• **Grade 0 to 1:** 78.7% (74/94) vs. 71.3% (67/94) • **Grade 2 to 3:** 21.3% (20/94) vs. 28.7% (27/94); adj OR 0.69 (0.32 to 1.51); p= 0.36	
	Late	Dutz 2019 (N=58) Retro propensity score-matched cohort Prostate Cancer Risk: Low (3%), Intermediate (78%), High (19%)	No	No	No	Yes[3] (-1)	*Proportion of patients* • **Grade 1:** 23% (5/22) vs. 32% (7/22); RR 0.71 (95% CI 0.27 to 1.91)† • **Grade 2:** 23% (5/22) vs. 27% (6/22); RR 0.83 (95% CI 0.30 to 2.33)† • **Grade 3:** 0% (0/22) vs. 5% (1/22), p=0.32	
		Fang 2015 (N=188) Retro case-matched cohort Prostate Cancer Risk: Low (55%), Intermediate (31%), High (7%)	No	No	No	Yes[3] (-1)	*Proportion of patients* • **Grade 0 to 1:** 87.2% (82/94) vs. 80.9% (76/94) • **Grade 2 to 3:** 12.8% (12/94) vs. 18.3% (17/94); adj. HR 0.56 (0.22 to 1.41); p=0.22	
		Pan 2018 (N=4158) Retro propensity-score matched database study‡ Prostate Cancer Risk: NR	No	No	No	Yes[3] (-1)	*Cumulative incidence, any urinary toxicity (any grade)* • **6-months**: 12.1% (n=693) vs. 21.5% (n=3465) • **12-months**: 23.1% (n=572) vs. 31.6% (n=2862) • **24-months**: 33.3% (n=341) vs. 42.2% (n=1718) • **36-months**: 39.1% (n=205) vs. 48.3% (n=1003) HR 0.72 (0.63 to 0.83); p<0.001	

Outcome	Time	Studies, Year, N, RoB Tumor	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	PBT* vs. Photon (various)* Effect estimate (95% CI)	Conclusion Quality (SoE)
							Sensitivity analysis based on validated diagnosis and procedure codes for severe toxicities post-pelvic radiation showed less urethral stricture with PBT at 24 months (1.3% vs. 0%; HR 0.12, 95% CI 0.02 to 0.86); no differences in cystitis, ureteral stricture, or urinary/rectal fistula.	
Erectile dysfunction (cumulative incidence)	36 mos.	Pan 2018 (N=4158) Retro propensity-score matched database study‡ Prostate Cancer Risk: NR	No	No	No	Yes[3] (-1)	• **6-months**: 5.0% (n=693) vs. 9.7% (n=3465) • **12-months**: 10.6% (n=572) vs. 18.1% (n=2862) • **24-months**: 20.7% (n=341) vs. 27.8% (n=1718) • **36-months**: 28.6% (n=205) vs. 34.3% (n=1003) HR 0.71 (0.59 to 0.84); p=0.001 *Sensitivity analysis* using procedure codes only (as surrogate for toxicity severity), 24 month incidence: 2.0% vs. 3.1%, HR 0.63, 95% CI 0.36 to 1.10	

adj. = adjusted; CI = confidence interval; GI = gastrointestinal; GU = genitourinary; HR = hazard ratio; IMRT = intensity-modulated radiation therapy; KQ = Key Question; NR = not reported; NS = not statistically significant; OR = odds ratio; PBT = proton beam therapy; Retro = retrospective study design; RR = risk ratio

* **Khmelevsky 2018**: Photon (standard conformal) + PBT boost vs. Photon (standard conformal) alone.
Dutz 2019: PBT (passive scatter) vs. IMRT
Fang 2015: PBT (passive scatter) vs. IMRT
Pan 2018: PBT vs. IMRT

†RR and 95% CI were calculated by AAI.

‡MarketScan Commercial Claims and Encounters database.

Reasons for downgrade:

1. Serious risk of bias: Majority of studies did not control for confounding and/or did not account for time at risk for survival outcomes. Studies which did control for confounding via study design and/or statistical analyses (e.g. Adequate randomization and concealment, matching, multivariate regression, propensity matching) were not downgrade for risk of bias.
2. Inconsistency: differing estimates of effects across studies; If effect size estimates across studies are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was not downgraded. Consistency is also unknown if there is of overlap of study populations, use different treatment protocols and/or different treatment types (including use of co-intervention such as chemotherapy).
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded

6 References

1. Cancer Prevalence and Cost of Care Projections. National Cancer Institute. at https://costprojections.cancer.gov/.)
2. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.
3. Transarterial Chemoembolization (TACE). RadiologyInfo.org. at https://www.radiologyinfo.org/en/info.cfm?pg=chemoembol#benefits-risks.)
4. Sec. 892.5050 Medical charged-particle radiation therapy system. U.S. Food and Drug Administration (FDA), 1988. at https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPCD/classification.cfm?ID=IYE.)
5. Evolution of Cancer Treatments: Radiation. American Cancer Society, 2014. at https://www.cancer.org/cancer/cancer-basics/history-of-cancer/cancer-treatment-radiation.html.)
6. A collection of R functions supporting the text book Modern Epidemiology, Second Edition, by Kenneth J. Rothman and Sander Greenland. GitHub, Inc., 2017. (Accessed October 13, 2017, at https://github.com/epijim/episheet.)
7. [RETIRED] Centers for Medicare and Medicaid Services (CMS). Medicare Local Coverage Determination (LCD L34634) for Proton Beam Therapy. 2017. (Accessed 02/01/19,
8. Cancer Statistics. National Cancer Institute, 2018. at https://www.cancer.gov/about-cancer/understanding/statistics.)
9. Centers for Medicare and Medicaid Services (CMS). Medicare Local Coverage Determination (LCD L35075) for Proton Beam Therapy. 2018. (Accessed 02/01/19,
10. Centers for Medicare and Medicaid Services (CMS). Medicare Local Coverage Determination (LCD L36658) for Proton Beam Therapy. 2018. (Accessed 02/01/19,
11. Stereotactic radiosurgery. Mayo Clinic, 2018. at https://www.mayoclinic.org/tests-procedures/stereotactic-radiosurgery/about/pac-20384526.)
12. Proton Therapy Centers in the U.S. The National Association For Proton Therapy, 2019. at https://www.proton-therapy.org/map/.)
13. Adeberg S, Bernhardt D, Harrabi SB, et al. Sequential proton boost after standard chemoradiation for high-grade glioma. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;125:266-72.
14. Proton Beam and Neutron Beam Radiotherapy. 2018. (Accessed August 2018, at http://www.aetna.com/cpb/medical/data/200_299/0270.html.)
15. Agarwal A, Thaker NG, Tawk B, et al. The evolution of radiation therapy for retinoblastoma: the MD Anderson Cancer Center experience. International Journal of Particle Therapy 2016;2:490-8.
16. Aibe N, Demizu Y, Sulaiman NS, et al. Outcomes of Patients With Primary Sacral Chordoma Treated With Definitive Proton Beam Therapy. International journal of radiation oncology, biology, physics 2018;100:972-9.
17. Albers CA, Grieve AJ. Test review: Bayley, N.(2006). Bayley scales of infant and toddler development–third edition. San Antonio, TX: Harcourt assessment. Journal of Psychoeducational Assessment 2007;25:180-90.

18. Proton Beam Radiation Therapy. Anthem Blue Cross Blue Shield, 2018. (Accessed August 2018, at https://www.anthem.com/medicalpolicies/policies/mp_pw_a053258.htm.)
19. Ares C, Albertini F, Frei-Welte M, et al. Pencil beam scanning proton therapy for pediatric intracranial ependymoma. Journal of neuro-oncology 2016;128:137-45.
20. Arimura T, Yoshiura T, Matsukawa K, Kondo N, Kitano I, Ogino T. Proton Beam Therapy Alone for Intermediate- or High-Risk Prostate Cancer: An Institutional Prospective Cohort Study. Cancers 2018;10.
21. Armoogum K, Thorp N. Dosimetric comparison and potential for improved clinical outcomes of paediatric CNS patients treated with protons or IMRT. Cancers 2015;7:706-22.
22. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. Bmj 2004;328:1490.
23. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401-6.
24. Barney CL, Brown AP, Grosshans DR, et al. Technique, outcomes, and acute toxicities in adults treated with proton beam craniospinal irradiation. Neuro-oncology 2014;16:303-9.
25. Barry MJ, Fowler Jr FJ, O'Leary MP, et al. The American Urological Association symptom index for benign prostatic hyperplasia. The Journal of urology 1992;148:1549-57.
26. Baskar R, Lee KA, Yeo R, Yeoh KW. Cancer and radiation therapy: current advances and future directions. International journal of medical sciences 2012;9:193-9.
27. Bayley N. Bayley scales of infant and toddler development: administration manual: Harcourt Assessment; 2006.
28. Bellocq D, Roy P, Kodjikian L, et al. 20-year assessment of metastatic latency and subsequent time to death after proton therapy for uveal melanomas. Melanoma Res 2018.
29. Berkman ND, Lohr KN, Ansari M, et al. Grading the strength of a body of evidence when assessing health care interventions for the effective health care program of the Agency for Healthcare Research and Quality: an update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]: Agency for Healthcare Research and Quality (US); 2013.
30. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD)2013.
31. Bielamowicz K, Okcu MF, Sonabend R, et al. Hypothyroidism after craniospinal irradiation with proton or photon therapy in patients with medulloblastoma. Pediatric hematology and oncology 2018;35:257-67.
32. Bishop AJ, Greenfield B, Mahajan A, et al. Proton beam therapy versus conformal photon radiation therapy for childhood craniopharyngioma: multi-institutional analysis of outcomes, cyst dynamics, and toxicity. International journal of radiation oncology, biology, physics 2014;90:354-61.
33. Blanchard P, Garden AS, Gunn GB, et al. Intensity-modulated proton beam therapy (IMPT) versus intensity-modulated photon therapy (IMRT) for patients with oropharynx cancer - A case matched analysis. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;120:48-55.
34. Bojaxhiu B, Ahlhelm F, Walser M, et al. Radiation Necrosis and White Matter Lesions in Pediatric Patients With Brain Tumors Treated With Pencil Beam Scanning Proton Therapy. International journal of radiation oncology, biology, physics 2018;100:987-96.

35. Boker A, Pilger D, Cordini D, et al. Neoadjuvant proton beam irradiation vs. adjuvant ruthenium brachytherapy in transscleral resection of uveal melanoma. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie 2018.
36. Brewington BY, Shao YF, Davidorf FH, Cebulla CM. Brachytherapy for patients with uveal melanoma: historical perspectives and future treatment directions. Clinical ophthalmology (Auckland, NZ) 2018;12:925-34.
37. Bronk JK, Guha-Thakurta N, Allen PK, Mahajan A, Grosshans DR, McGovern SL. Analysis of pseudoprogression after proton or photon therapy of 99 patients with low grade and anaplastic glioma. Clinical and translational radiation oncology 2018;9:30-4.
38. Bruininks RH. Scales of Independent Behavior: Psycho-educational Battery: DLM; 1984.
39. Bryant C, Smith TL, Henderson RH, et al. Five-Year Biochemical Results, Toxicity, and Patient-Reported Quality of Life After Delivery of Dose-Escalated Image Guided Proton Therapy for Prostate Cancer. International journal of radiation oncology, biology, physics 2016;95:422-34.
40. Bujold A, Massey CA, Kim JJ, et al. Sequential phase I and II trials of stereotactic body radiotherapy for locally advanced hepatocellular carcinoma. Journal of Clinical Oncology 2013;31:1631-9.
41. Bush DA, Do S, Lum S, et al. Partial breast radiation therapy with proton beam: 5-year results with cosmetic outcomes. International journal of radiation oncology, biology, physics 2014;90:501-5.
42. Bush DA, Smith JC, Slater JD, et al. Randomized Clinical Trial Comparing Proton Beam Radiation Therapy with Transarterial Chemoembolization for Hepatocellular Carcinoma: Results of an Interim Analysis. International journal of radiation oncology, biology, physics 2016;95:477-82.
43. Canadian Agency For Drugs Technologies in Health (CADTH). Proton beam therapy versus photon radiotherapy for adult and pediatric oncology patients: a review of the clinical and cost-effectiveness. Ottawa, ON Canada: CADTH; 2016.
44. Chang JY, Verma V, Li M, et al. Proton Beam Radiotherapy and Concurrent Chemotherapy for Unresectable Stage III Non-Small Cell Lung Cancer: Final Results of a Phase 2 Study. JAMA oncology 2017;3:e172032.
45. Chang JY, Zhang W, Komaki R, et al. Long-term outcome of phase I/II prospective study of dose-escalated proton therapy for early-stage non-small cell lung cancer. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;122:274-80.
46. Chao HH, Berman AT, Simone CB, 2nd, et al. Multi-Institutional Prospective Study of Reirradiation with Proton Beam Radiotherapy for Locoregionally Recurrent Non-Small Cell Lung Cancer. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer 2017;12:281-92.
47. Chi A, Chen H, Wen S, Yan H, Liao Z. Comparison of particle beam therapy and stereotactic body radiotherapy for early stage non-small cell lung cancer: A systematic review and hypothesis-generating meta-analysis. Radiotherapy and Oncology 2017;123:346-54.
48. Chowdhary M, Lee A, Gao S, et al. Is Proton Therapy a "Pro" for Breast Cancer? A Comparison of Proton vs. Non-proton Radiotherapy Using the National Cancer Database. Frontiers in oncology 2018;8:678.
49. Chowdhry VK, Liu L, Goldberg S, et al. Thoracolumbar spinal cord tolerance to high dose conformal proton-photon radiation therapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;119:35-9.

50. Chuong MD, Hartsell W, Larson G, et al. Minimal toxicity after proton beam therapy for prostate and pelvic nodal irradiation: results from the proton collaborative group REG001-09 trial. Acta oncologica (Stockholm, Sweden) 2018;57:368-74.
51. Cleeland CS. The M. D. Anderson Symptom Inventory User Guide. 2009.
52. Cleeland CS, Mendoza TR, Wang XS, et al. Assessing symptom distress in cancer patients: the M.D. Anderson Symptom Inventory. Cancer 2000;89:1634-46.
53. Colaco RJ, Hoppe BS, Flampouri S, et al. Rectal toxicity after proton therapy for prostate cancer: an analysis of outcomes of prospective studies conducted at the university of Florida Proton Therapy Institute. International journal of radiation oncology, biology, physics 2015;91:172-81.
54. Connell PP, Hellman S. Advances in radiotherapy and implications for the next century: a historical perspective. Cancer research 2009;69:383-92.
55. Conners CK, Staff M, Connelly V, Campbell S, MacLean M, Barnes J. Conners' continuous performance Test II (CPT II v. 5). Multi-Health Syst Inc 2000;29:175-96.
56. Crott R, Ali F, Burdette-Radoux S. Cost–Utility of Adjuvant High-Dose Interferon Alpha Therapy in Stage III Cutaneous Melanoma in Quebec. Value in Health 2004;7:423-32.
57. Cuaron JJ, Chon B, Tsai H, et al. Early toxicity in patients treated with postoperative proton therapy for locally advanced breast cancer. International journal of radiation oncology, biology, physics 2015;92:284-91.
58. Cummings DE, Merriam GR. Growth hormone therapy in adults. Annual Review of Medicine 2003;54:513-33.
59. Dagan R, Bryant C, Li Z, et al. Outcomes of Sinonasal Cancer Treated With Proton Therapy. International journal of radiation oncology, biology, physics 2016;95:377-85.
60. Damato B, Kacperek A, Chopra M, Campbell IR, Errington RD. Proton beam radiotherapy of choroidal melanoma: the Liverpool-Clatterbridge experience. International Journal of Radiation Oncology* Biology* Physics 2005;62:1405-11.
61. Darby SC, Ewertz M, McGale P, et al. Risk of ischemic heart disease in women after radiotherapy for breast cancer. New England Journal of Medicine 2013;368:987-98.
62. De B, Cahlon O, Sine K, Mah D, Hug EB, Wolden SL. Early Axial Growth Outcomes of Pediatric Patients Receiving Proton Craniospinal Irradiation. J Pediatr Hematol Oncol 2018;40:574-9.
63. DeLaney TF, Liebsch NJ, Pedlow FX, et al. Long-term results of Phase II study of high dose photon/proton radiotherapy in the management of spine chordomas, chondrosarcomas, and other sarcomas. Journal of surgical oncology 2014;110:115-22.
64. Demizu Y, Mizumoto M, Onoe T, et al. Proton beam therapy for bone sarcomas of the skull base and spine: A retrospective nationwide multicenter study in Japan. Cancer science 2017;108:972-7.
65. Dendale R, Lumbroso-Le Rouic L, Noel G, et al. Proton beam radiotherapy for uveal melanoma: results of Curie Institut–Orsay proton therapy center (ICPO). International Journal of Radiation Oncology* Biology* Physics 2006;65:780-7.
66. Department of Veterans Affairs. Comparative Effectiveness of Proton Irradiation Treatment. In: Peterson K, McCleery E, Waldrip K, Helfand M, eds. Washington, DC: VA Evidence-Based Synthesis Program (ESP); 2015.
67. Deraniyagala RL, Yeung D, Mendenhall WM, et al. Proton therapy for skull base chordomas: an outcome study from the university of Florida proton therapy institute. Journal of neurological surgery Part B, Skull base 2014;75:53-7.

68. Doyen J, Falk AT, Floquet V, Hérault J, Hannoun-Lévi J-M. Proton beams in cancer treatments: clinical outcomes and dosimetric comparisons with photon therapy. Cancer treatment reviews 2016;43:104-12.
69. Dutz A, Agolli L, Baumann M, et al. Early and late side effects, dosimetric parameters and quality of life after proton beam therapy and IMRT for prostate cancer: a matched-pair analysis. Acta oncologica (Stockholm, Sweden) 2019:1-10.
70. Dutz A, Luhr A, Agolli L, et al. Development and validation of NTCP models for acute side-effects resulting from proton beam therapy of brain tumours. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2018.
71. Eaton BR, Esiashvili N, Kim S, et al. Endocrine outcomes with proton and photon radiotherapy for standard risk medulloblastoma. Neuro-oncology 2016b;18:881-7.
72. Eaton BR, Esiashvili N, Kim S, et al. Clinical Outcomes Among Children With Standard-Risk Medulloblastoma Treated With Proton and Photon Radiation Therapy: A Comparison of Disease Control and Overall Survival. International journal of radiation oncology, biology, physics 2016a;94:133-8.
73. El Shafie RA, Czech M, Kessel KA, et al. Clinical outcome after particle therapy for meningiomas of the skull base: toxicity and local control in patients treated with active rasterscanning. Radiation oncology (London, England) 2018;13:54.
74. Proton Beam Therapy. NHS England, 2018. (Accessed August 2018, at https://www.england.nhs.uk/commissioning/spec-services/highly-spec-services/pbt/.)
75. Fakhry C, Zhang Q, Nguyen-Tan PF, et al. Human papillomavirus and overall survival after progression of oropharyngeal squamous cell carcinoma. Journal of clinical oncology 2014;32:3365.
76. Fang P, Mick R, Deville C, et al. A case-matched study of toxicity outcomes after proton therapy and intensity-modulated radiation therapy for prostate cancer. Cancer 2015;121:1118-27.
77. Fang P, Shiraishi Y, Verma V, et al. Lymphocyte-Sparing Effect of Proton Therapy in Patients with Esophageal Cancer Treated with Definitive Chemoradiation. Int J Part Ther 2018;4:23-32.
78. Feuvret L, Bracci S, Calugaru V, et al. Efficacy and Safety of Adjuvant Proton Therapy Combined With Surgery for Chondrosarcoma of the Skull Base: A Retrospective, Population-Based Study. International journal of radiation oncology, biology, physics 2016;95:312-21.
79. Fukuda K, Okumura T, Abei M, et al. Long-term outcomes of proton beam therapy in patients with previously untreated hepatocellular carcinoma. Cancer science 2017;108:497-503.
80. Fukumitsu N, Okumura T, Takizawa D, et al. Proton beam therapy for metastatic liver tumors. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2015;117:322-7.
81. Fung V, Calugaru V, Bolle S, et al. Proton beam therapy for skull base chordomas in 106 patients: A dose adaptive radiation protocol. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2018;128:198-202.
82. Garibaldi C, Jereczek-Fossa BA, Marvaso G, et al. Recent advances in radiation oncology. Ecancermedicalscience 2017;11:785.
83. Gentile MS, Yeap BY, Paganetti H, et al. Brainstem Injury in Pediatric Patients With Posterior Fossa Tumors Treated With Proton Beam Therapy and Associated Dosimetric Factors. International journal of radiation oncology, biology, physics 2018;100:719-29.

84. Giantsoudi D, Sethi RV, Yeap BY, et al. Incidence of CNS Injury for a Cohort of 111 Patients Treated With Proton Therapy for Medulloblastoma: LET and RBE Associations for Areas of Injury. International journal of radiation oncology, biology, physics 2016;95:287-96.
85. Giap H, Giap B. Historical perspective and evolution of charged particle beam therapy. Translational Cancer Research 2012;1:127-36.
86. Gioia GA, Isquith PK, Guy SC, Kenworthy L. Test review behavior rating inventory of executive function. Child Neuropsychology 2000;6:235-8.
87. Giovannini G, Bohlen T, Cabal G, et al. Variable RBE in proton therapy: comparison of different model predictions and their influence on clinical-like scenarios. Radiation oncology (London, England) 2016;11:68.
88. Goitein M. Magical protons? International Journal of Radiation Oncology• Biology• Physics 2008;70:654-6.
89. Grant SR, Grosshans DR, Bilton SD, et al. Proton versus conventional radiotherapy for pediatric salivary gland tumors: Acute toxicity and dosimetric characteristics. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2015;116:309-15.
90. Grassberger C, Hong TS, Hato T, et al. Differential Association Between Circulating Lymphocyte Populations With Outcome After Radiation Therapy in Subtypes of Liver Cancer. International journal of radiation oncology, biology, physics 2018;101:1222-5.
91. Gray ST, Lin A, Curry WT, et al. Delayed complications after anterior craniofacial resection of malignant skull base tumors. Journal of neurological surgery Part B, Skull base 2014;75:110-6.
92. Greenberger BA, Pulsifer MB, Ebb DH, et al. Clinical outcomes and late endocrine, neurocognitive, and visual profiles of proton radiation for pediatric low-grade gliomas. International journal of radiation oncology, biology, physics 2014;89:1060-8.
93. Gunn GB, Blanchard P, Garden AS, et al. Clinical Outcomes and Patterns of Disease Recurrence After Intensity Modulated Proton Therapy for Oropharyngeal Squamous Carcinoma. International journal of radiation oncology, biology, physics 2016;95:360-7.
94. Gunther JR, Rahman AR, Dong W, et al. Craniospinal irradiation prior to stem cell transplant for hematologic malignancies with CNS involvement: Effectiveness and toxicity after photon or proton treatment. Pract Radiat Oncol 2017;7:e401-e8.
95. Gunther JR, Sato M, Chintagumpala M, et al. Imaging Changes in Pediatric Intracranial Ependymoma Patients Treated With Proton Beam Radiation Therapy Compared to Intensity Modulated Radiation Therapy. International journal of radiation oncology, biology, physics 2015;93:54-63.
96. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383-94.
97. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 2011;64:380-2.
98. Ha B, Cho KH, Lee KH, et al. Long-term results of a phase II study of hypofractionated proton therapy for prostate cancer: moderate versus extreme hypofractionation. Radiation oncology (London, England) 2019;14:4.
99. Hall MD, Bradley JA, Rotondo RL, et al. Risk of Radiation Vasculopathy and Stroke in Pediatric Patients Treated With Proton Therapy for Brain and Skull Base Tumors. International journal of radiation oncology, biology, physics 2018;101:854-9.

100. Hashimoto T, Isobe T, Hashii H, et al. Influence of secondary neutrons induced by proton radiotherapy for cancer patients with implantable cardioverter defibrillators. Radiation oncology 2012;7:10.
101. Hatayama Y, Nakamura T, Suzuki M, et al. Clinical Outcomes and Prognostic Factors of High-Dose Proton Beam Therapy for Peripheral Stage I Non-Small-Cell Lung Cancer. Clinical lung cancer 2016;17:427-32.
102. Hawkins BS. The Collaborative Ocular Melanoma Study (COMS) randomized trial of pre-enucleation radiation of large choroidal melanoma: IV. Ten-year mortality findings and prognostic factors. COMS report number 24. American journal of ophthalmology 2004;138:936-51.
103. Hayashi Y, Nakamura T, Mitsudo K, et al. Re-irradiation using proton beam therapy combined with weekly intra-arterial chemotherapy for recurrent oral cancer. Asia-Pacific journal of clinical oncology 2017;13:e394-e401.
104. Hayashi Y, Nakamura T, Mitsudo K, et al. Retrograde intra-arterial chemotherapy and daily concurrent proton beam therapy for recurrent oral cavity squamous cell carcinoma: Analysis of therapeutic results in 46 cases. Head & neck 2016;38:1145-51.
105. Health AS. Clinical Appropriateness Guidelines: Radiation Oncology Proton Beam Therapy Guidelines. Chicago, IL 60631: AIM Specialty Health; 2018:1-13.
106. Health Evidence Review Commission (HERC) Oregon Health Authority. Coverage Guidance: Proton Beam Therpay. 2016.
107. Higgins J. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.02011.
108. Higgins KA, O'Connell K, Liu Y, et al. National Cancer Database Analysis of Proton Versus Photon Radiation Therapy in Non-Small Cell Lung Cancer. International journal of radiation oncology, biology, physics 2017;97:128-37.
109. Hirano E, Fuji H, Onoe T, Kumar V, Shirato H, Kawabuchi K. Cost-effectiveness analysis of cochlear dose reduction by proton beam therapy for medulloblastoma in childhood. J Radiat Res 2014;55:320-7.
110. Ho CK, Bryant CM, Mendenhall NP, et al. Long-term outcomes following proton therapy for prostate cancer in young men with a focus on sexual health. Acta oncologica (Stockholm, Sweden) 2018;57:582-8.
111. Holliday EB, Garden AS, Rosenthal DI, et al. Proton therapy reduces treatment-related toxicities for patients with nasopharyngeal cancer: a case-match control study of intensity-modulated proton therapy and intensity-modulated photon therapy. International journal of particle therapy 2015;2:19-28.
112. Hong TS, Ryan DP, Borger DR, et al. A phase 1/2 and biomarker study of preoperative short course chemoradiation with proton beam therapy and capecitabine followed by early surgery for resectable pancreatic ductal adenocarcinoma. International journal of radiation oncology, biology, physics 2014;89:830-8.
113. Hong TS, Wo JY, Borger DR, et al. Phase II Study of Proton-Based Stereotactic Body Radiation Therapy for Liver Metastases: Importance of Tumor Genotype. J Natl Cancer Inst 2017;109.
114. Hong TS, Wo JY, Yeap BY, et al. Multi-Institutional Phase II Study of High-Dose Hypofractionated Proton Beam Therapy in Patients With Localized, Unresectable Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma. J Clin Oncol 2016;34:460-8.
115. Hoppe BS, Hill-Kayser CE, Tseng YD, et al. Consolidative proton therapy after chemotherapy for patients with Hodgkin lymphoma. Ann Oncol 2017;28:2179-84.

116. Hoppe BS, Michalski JM, Mendenhall NP, et al. Comparative effectiveness study of patient-reported outcomes after proton therapy or intensity-modulated radiotherapy for prostate cancer. Cancer 2014;120:1076-82.
117. Hoppe BS, Tsai H, Larson G, et al. Proton therapy patterns-of-care and early outcomes for Hodgkin lymphoma: results from the Proton Collaborative Group Registry. Acta oncologica (Stockholm, Sweden) 2016;55:1378-80.
118. Indelicato DJ, Bradley JA, Rotondo RL, et al. Outcomes following proton therapy for pediatric ependymoma. Acta oncologica (Stockholm, Sweden) 2018;57:644-8.
119. Indelicato DJ, Bradley JA, Sandler ES, et al. Clinical outcomes following proton therapy for children with central nervous system tumors referred overseas. Pediatric blood & cancer 2017;64.
120. Indelicato DJ, Flampouri S, Rotondo RL, et al. Incidence and dosimetric parameters of pediatric brainstem toxicity following proton therapy. Acta oncologica (Stockholm, Sweden) 2014;53:1298-304.
121. Indelicato DJ, Rotondo RL, Begosh-Mayne D, et al. A Prospective Outcomes Study of Proton Therapy for Chordomas and Chondrosarcomas of the Spine. International journal of radiation oncology, biology, physics 2016;95:297-303.
122. Ishikawa H, Hashimoto T, Moriwaki T, et al. Proton beam therapy combined with concurrent chemotherapy for esophageal cancer. Anticancer Res 2015;35:1757-62.
123. Ishikawa Y, Nakamura T, Kato T, et al. Dosemetric Parameters Predictive of Rib Fractures after Proton Beam Therapy for Early-Stage Lung Cancer. Tohoku J Exp Med 2016;238:339-45.
124. Iwata H, Ishikawa H, Takagi M, et al. Long-term outcomes of proton therapy for prostate cancer in Japan: a multi-institutional survey of the Japanese Radiation Oncology Study Group. Cancer Med 2018;7:677-89.
125. Jacola LM, Conklin HM, Scoggins MA, et al. Investigating the Role of Hypothalamic Tumor Involvement in Sleep and Cognitive Outcomes Among Children Treated for Craniopharyngioma. Journal of pediatric psychology 2016;41:610-22.
126. Janssen KC, Phillipson S, O'Connor J, Johns MW. Validation of the Epworth Sleepiness Scale for Children and Adolescents using Rasch analysis. Sleep medicine 2017;33:30-5.
127. Jarlskog CZ, Paganetti H. Sensitivity of different dose scoring methods on organ-specific neutron dose calculations in proton therapy. Physics in Medicine & Biology 2008;53:4523.
128. Jeldres C, Cullen J, Hurwitz LM, et al. Prospective quality-of-life outcomes for low-risk prostate cancer: Active surveillance versus radical prostatectomy. Cancer 2015;121:2465-73.
129. Jhaveri J, Cheng E, Tian S, et al. Proton vs. Photon Radiation Therapy for Primary Gliomas: An Analysis of the National Cancer Data Base. Frontiers in oncology 2018;8:440.
130. Kabolizadeh P, Chen YL, Liebsch N, et al. Updated Outcome and Analysis of Tumor Response in Mobile Spine and Sacral Chordoma Treated With Definitive High-Dose Photon/Proton Radiation Therapy. International journal of radiation oncology, biology, physics 2017;97:254-62.
131. Kahalley LS, Ris MD, Grosshans DR, et al. Comparing Intelligence Quotient Change After Treatment With Proton Versus Photon Radiation Therapy for Pediatric Brain Tumors. J Clin Oncol 2016;34:1043-9.
132. Kahalley LS, Ris MD, Mahajan A, et al. Prospective, Longitudinal Comparison of Neurocognitive Change in Pediatric Brain Tumor Patients Treated with Proton Radiotherapy versus Surgery Only. Neuro-oncology 2019.

133. Kammerer E, Le Guevelou J, Chaikh A, et al. Proton therapy for locally advanced breast cancer: a systematic review of the literature. Cancer treatment reviews 2018;63:19-27.
134. Kamran SC, Collier JM, Lane AM, et al. Outcomes of proton therapy for the treatment of uveal metastases. International journal of radiation oncology, biology, physics 2014;90:1044-50.
135. Kamran SC, Goldberg SI, Kuhlthau KA, et al. Quality of life in patients with proton-treated pediatric medulloblastoma: Results of a prospective assessment with 5-year follow-up. Cancer 2018;124:3390-400.
136. Kanemoto A, Okumura T, Ishikawa H, et al. Outcomes and prognostic factors for recurrence after high-dose proton beam therapy for centrally and peripherally located stage I non--small-cell lung cancer. Clinical lung cancer 2014;15:e7-12.
137. Kang KH, Schapira E, Niemierko A, et al. The role of proton beam therapy in central neurocytoma: A single-institution experience. Pract Radiat Oncol 2018;8:e305-e11.
138. Kawashima M, Furuse J, Nishio T, et al. Phase II study of radiotherapy employing proton beam for hepatocellular carcinoma. Journal of clinical oncology 2005;23:1839-46.
139. Khmelevsky EV, Kancheli IN, Khoroshkov VS, Kaprin AD. Morbidity dynamics in proton-photon or photon radiation therapy for locally advanced prostate cancer. Reports of practical oncology and radiotherapy : journal of Greatpoland Cancer Center in Poznan and Polish Society of Radiation Oncology 2018;23:21-7.
140. Kim DY, Park JW, Kim TH, et al. Risk-adapted simultaneous integrated boost-proton beam therapy (SIB-PBT) for advanced hepatocellular carcinoma with tumour vascular thrombosis. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;122:122-9.
141. Kim J, Wells C, Khangura S, et al. Proton Beam Therapy for the Treatment of Cancer in Children and Adults: A Health Technology Assessment. Ottawa, CA: CADTH; 2017.
142. Kim TH, Lee WJ, Woo SM, et al. Effectiveness and Safety of Simultaneous Integrated Boost-Proton Beam Therapy for Localized Pancreatic Cancer. Technol Cancer Res Treat 2018;17:1533033818783879.
143. Kim TH, Park JW, Kim BH, et al. Optimal time of tumour response evaluation and effectiveness of hypofractionated proton beam therapy for inoperable or recurrent hepatocellular carcinoma. Oncotarget 2018;9:4034-43.
144. Kindler HL, Ismaila N, Armato III SG, et al. Treatment of malignant pleural mesothelioma: American Society of Clinical Oncology clinical practice guideline. Journal of Clinical Oncology 2018;36:1343-73.
145. King G, Nielsen R. Why propensity scores should not be used for matching. Copy at http://j mp/1sexgVw Download Citation BibTex Tagged XML Download Paper 2016;378.
146. Kjellberg R, Sweet W, Preston W, Koehler A. The Bragg peak of a proton beam in intracranial therapy of tumors. Transactions of the American Neurological Association (US) 1962;87.
147. Konstantinidis L, Roberts D, Errington RD, Kacperek A, Heimann H, Damato B. Transpalpebral proton beam radiotherapy of choroidal melanoma. The British journal of ophthalmology 2015;99:232-5.
148. Kopecky AS, Khan AJ, Pan W, Drachtman R, Parikh RR. Outcomes and patterns of care in a nationwide cohort of pediatric medulloblastoma: Factors affecting proton therapy utilization. Adv Radiat Oncol 2017;2:588-96.

149. Kralik SF, Ho CY, Finke W, Buchsbaum JC, Haskins CP, Shih CS. Radiation Necrosis in Pediatric Patients with Brain Tumors Treated with Proton Radiotherapy. AJNR American journal of neuroradiology 2015;36:1572-8.
150. Kralik SF, Mereniuk TR, Grignon L, et al. Radiation Induced Cerebral Microbleeds in Pediatric Patients with Brain Tumors Treated with Proton Radiotherapy. International journal of radiation oncology, biology, physics 2018.
151. Kralik SF, Watson GA, Shih CS, Ho CY, Finke W, Buchsbaum J. Radiation-Induced Large Vessel Cerebral Vasculopathy in Pediatric Patients With Brain Tumors Treated With Proton Radiation Therapy. International journal of radiation oncology, biology, physics 2017;99:817-24.
152. Ladra MM, Mandeville HC, Niemierko A, et al. Local failure in parameningeal rhabdomyosarcoma correlates with poor response to induction chemotherapy. International journal of radiation oncology, biology, physics 2015;92:358-67.
153. Ladra MM, Szymonifka JD, Mahajan A, et al. Preliminary results of a phase II trial of proton radiotherapy for pediatric rhabdomyosarcoma. J Clin Oncol 2014;32:3762-70.
154. Lane AM, Kim IK, Gragoudas ES. Long-term Risk of Melanoma-Related Mortality for Patients With Uveal Melanoma Treated With Proton Beam Therapy. JAMA ophthalmology 2015;133:792-6.
155. Larsson B, Leksell L, Rexed B, Sourander P, Mair W, Andersson B. The high-energy proton beam as a neurosurgical tool. Nature 1958;182:1222.
156. Lee SU, Moon SH, Cho KH, et al. Ablative dose proton beam therapy for stage I and recurrent non-small cell lung carcinomas : Ablative dose PBT for NSCLC. Strahlenther Onkol 2016;192:649-57.
157. Leiser D, Calaminus G, Malyapa R, et al. Tumour control and Quality of Life in children with rhabdomyosarcoma treated with pencil beam scanning proton therapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;120:163-8.
158. Leroy R, Benahmed N, Hulstaert F, Van Damme N, De Ruysscher D. Proton therapy in children: a systematic review of clinical effectiveness in 15 pediatric cancers. International Journal of Radiation Oncology* Biology* Physics 2016;95:267-78.
159. Leung HW, Chan AL. Cost-utility of stereotactic radiation therapy versus proton beam therapy for inoperable advanced hepatocellular carcinoma. Oncotarget 2017;8:75568.
160. Levin W, Kooy H, Loeffler J, DeLaney T. Proton beam therapy. British journal of Cancer 2005;93:849.
161. Liao Z, Lee JJ, Komaki R, et al. Bayesian Adaptive Randomization Trial of Passive Scattering Proton Therapy and Intensity-Modulated Photon Radiotherapy for Locally Advanced Non-Small-Cell Lung Cancer. J Clin Oncol 2018;36:1813-22.
162. Lichtenberger EO, Kaufman AS. Essentials of WAIS-IV assessment: John Wiley & Sons; 2009.
163. Lin AJ, Rao YJ, Acharya S, Schwarz J, Rao PK, Grigsby P. Patterns of care and outcomes of proton and eye plaque brachytherapy for uveal melanoma: Review of the National Cancer Database. Brachytherapy 2017;16:1225-31.
164. Lin SH, Merrell KW, Shen J, et al. Multi-institutional analysis of radiation modality use and postoperative outcomes of neoadjuvant chemoradiation for esophageal cancer. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;123:376-81.
165. Lucas JT, Jr., Ladra MM, MacDonald SM, et al. Proton therapy for pediatric and adolescent esthesioneuroblastoma. Pediatric blood & cancer 2015;62:1523-8.
166. Ludwig Boltzmann Institut. [Stereotactic radiotherapy, proton therapy and irreversible electroporation for the treatment of localized prostate cancer: Systematic Review] Stereotaktische

Strahlentherapie, Protonentherapie und Irreversible Elektroporation zur Behandlung des lokalisierten Prostatakarzinoms. Systematischer Review. In: Schmidt L, Lohr P, Wild C, eds.: Ludwig Boltzmann Institut; 2018.

167. Macdonald SM, Sethi R, Lavally B, et al. Proton radiotherapy for pediatric central nervous system ependymoma: clinical outcomes for 70 patients. Neuro-oncology 2014;15:1552-9.
168. Maemura K, Mataki Y, Kurahara H, et al. Comparison of proton beam radiotherapy and hyper-fractionated accelerated chemoradiotherapy for locally advanced pancreatic cancer. Pancreatology 2017;17:833-8.
169. Mahdjoubi A, Dendale R, Desjardins L, et al. Treatment of Exudative Circumscribed Choroidal Hemangioma: Efficacy of Fractionated Proton Therapy (20 Gray Relative Biological Effectiveness in 8 Fractions). Retina 2017.
170. Mailhot Vega R, Kim J, Hollander A, et al. Cost effectiveness of proton versus photon radiation therapy with respect to the risk of growth hormone deficiency in children. Cancer 2015;121:1694-702.
171. Mailhot Vega RB, Ishaq O, Raldow A, et al. Establishing cost-effective allocation of proton therapy for breast irradiation. International Journal of Radiation Oncology* Biology* Physics 2016;95:11-8.
172. Makishima H, Ishikawa H, Tanaka K, et al. A retrospective study of late adverse events in proton beam therapy for prostate cancer. Mol Clin Oncol 2017;7:547-52.
173. Makishima H, Ishikawa H, Terunuma T, et al. Comparison of adverse effects of proton and X-ray chemoradiotherapy for esophageal cancer using an adaptive dose-volume histogram analysis. J Radiat Res 2015;56:568-76.
174. Makita C, Nakamura T, Takada A, et al. High-dose proton beam therapy for stage I non-small cell lung cancer: Clinical outcomes and prognostic factors. Acta oncologica (Stockholm, Sweden) 2015;54:307-14.
175. Maquilan G, Grover S, Alonso-Basanta M, Lustig RA. Acute toxicity profile of patients with low-grade gliomas and meningiomas receiving proton therapy. Am J Clin Oncol 2014;37:438-43.
176. Mathis T, Hofverberg P, Caujolle JP, et al. Occurrence of Phosphenes in Patients Undergoing Proton Beam Therapy for Ocular Tumor. American journal of ophthalmology 2018;192:31-8.
177. McDonald MW, Linton OR, Calley CS. Dose-volume relationships associated with temporal lobe radiation necrosis after skull base proton beam therapy. International journal of radiation oncology, biology, physics 2015;91:261-7.
178. McDonald MW, Liu Y, Moore MG, Johnstone PA. Acute toxicity in comprehensive head and neck radiation for nasopharynx and paranasal sinus cancers: cohort comparison of 3D conformal proton therapy and intensity modulated radiation therapy. Radiation oncology (London, England) 2016;11:32.
179. McDonald MW, Zolali-Meybodi O, Lehnert SJ, et al. Reirradiation of Recurrent and Second Primary Head and Neck Cancer With Proton Therapy. International journal of radiation oncology, biology, physics 2016;96:808-19.
180. McGovern SL, Okcu MF, Munsell MF, et al. Outcomes and acute toxicities of proton therapy for pediatric atypical teratoid/rhabdoid tumor of the central nervous system. International journal of radiation oncology, biology, physics 2014;90:1143-52.
181. Mendenhall NP, Hoppe BS, Nichols RC, et al. Five-year outcomes from 3 prospective trials of image-guided proton therapy for prostate cancer. International journal of radiation oncology, biology, physics 2014;88:596-602.

182. Merchant TE, Rose SR, Bosley C, Wu S, Xiong X, Lustig RH. Growth hormone secretion after conformal radiation therapy in pediatric patients with localized brain tumors. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2011;29:4776-80.
183. Mizuhata M, Takamatsu S, Shibata S, et al. Respiratory-gated Proton Beam Therapy for Hepatocellular Carcinoma Adjacent to the Gastrointestinal Tract without Fiducial Markers. Cancers 2018;10:58.
184. Mizumoto M, Murayama S, Akimoto T, et al. Proton beam therapy for pediatric malignancies: a retrospective observational multicenter study in Japan. Cancer Med 2016;5:1519-25.
185. Mizumoto M, Murayama S, Akimoto T, et al. Long-term follow-up after proton beam therapy for pediatric tumors: a Japanese national survey. Cancer science 2017;108:444-7.
186. Mizumoto M, Murayama S, Akimoto T, et al. Preliminary results of proton radiotherapy for pediatric rhabdomyosarcoma: a multi-institutional study in Japan. Cancer Med 2018;7:1870-4.
187. Mizumoto M, Oshiro Y, Okumura T, et al. Association between pretreatment retention rate of indocyanine green 15 min after administration and life prognosis in patients with HCC treated by proton beam therapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2014;113:54-9.
188. Mizumoto M, Yamamoto T, Ishikawa E, et al. Proton beam therapy with concurrent chemotherapy for glioblastoma multiforme: comparison of nimustine hydrochloride and temozolomide. Journal of neuro-oncology 2016;130:165-70.
189. Mokhtech M, Rotondo RL, Bradley JA, et al. Early outcomes and patterns of failure following proton therapy for nonmetastatic intracranial nongerminomatous germ cell tumors. Pediatric blood & cancer 2018;65:e26997.
190. Moreno AC, Zhang N, Giordano SH, et al. Trends and Outcomes of Proton Radiation Therapy Use for Non–Small Cell Lung Cancer. International Journal of Particle Therapy 2018.
191. Morgan RL, Thayer KA, Santesso N, et al. A risk of bias instrument for non-randomized studies of exposures: A users' guide to its application in the context of GRADE. Environment international 2019;122:168-84.
192. Moriarty JP, Borah BJ, Foote RL, Pulido JS, Shah ND. Cost-effectiveness of proton beam therapy for intraocular melanoma. PloS one 2015;10:e0127814.
193. Morimoto K, Demizu Y, Hashimoto N, et al. Particle radiotherapy using protons or carbon ions for unresectable locally advanced head and neck cancers with skull base invasion. Japanese journal of clinical oncology 2014;44:428-34.
194. Moskvin V, Lasley FD, Ray GL, et al. Acute skin toxicity associated with proton beam therapy in spine and brain patients. Journal of Radiation Oncology 2014;3:195-203.
195. Mouw KW, Sethi RV, Yeap BY, et al. Proton radiation therapy for the treatment of retinoblastoma. International journal of radiation oncology, biology, physics 2014;90:863-9.
196. Mozes P, Dittmar JO, Habermehl D, et al. Volumetric response of intracranial meningioma after photon or particle irradiation. Acta oncologica (Stockholm, Sweden) 2017;56:431-7.
197. Murray FR, Snider JW, Bolsi A, et al. Long-Term Clinical Outcomes of Pencil Beam Scanning Proton Therapy for Benign and Non-benign Intracranial Meningiomas. International journal of radiation oncology, biology, physics 2017;99:1190-8.
198. Nakajima K, Iwata H, Ogino H, et al. Acute toxicity of image-guided hypofractionated proton therapy for localized prostate cancer. International journal of clinical oncology 2018;23:353-60.

199. Nakamura N, Zenda S, Tahara M, et al. Proton beam therapy for olfactory neuroblastoma. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;122:368-72.
200. Nanda R, Flampouri S, Mendenhall NP, et al. Pulmonary Toxicity Following Proton Therapy for Thoracic Lymphoma. International journal of radiation oncology, biology, physics 2017;99:494-7.
201. Nayan M, Hamilton RJ, Juurlink DN, Finelli A, Kulkarni GS, Austin PC. Critical appraisal of the application of propensity score methods in the urology literature. BJU international 2017;120:873-80.
202. Network NCC. NCCN Guidelines for Treatment of Cancer by Site. National Comprehensive Cancer Network; 2018.
203. Newberry SJ, Ahmadzai N, Motala A, et al. AHRQ Methods for Effective Health Care. Surveillance and Identification of Signals for Updating Systematic Reviews: Implementation and Early Experience. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013.
204. Nguyen-Tan PF, Zhang Q, Ang KK, et al. Randomized phase III trial to test accelerated versus standard fractionation in combination with concurrent cisplatin for head and neck carcinomas in the Radiation Therapy Oncology Group 0129 trial: long-term report of efficacy and toxicity. Journal of Clinical Oncology 2014;32:3858.
205. Nguyen QN, Ly NB, Komaki R, et al. Long-term outcomes after proton therapy, with concurrent chemotherapy, for stage II-III inoperable non-small cell lung cancer. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2015;115:367-72.
206. Niedzielski JS, Yang J, Mohan R, et al. Differences in Normal Tissue Response in the Esophagus Between Proton and Photon Radiation Therapy for Non-Small Cell Lung Cancer Using In Vivo Imaging Biomarkers. International journal of radiation oncology, biology, physics 2017;99:1013-20.
207. Nishioka K, Prayongrat A, Ono K, et al. Prospective study to evaluate the safety of the world-first spot-scanning dedicated, small 360-degree gantry, synchrotron-based proton beam therapy system. J Radiat Res 2018;59:i63-i71.
208. Noel G, Gondi V. Proton therapy for tumors of the base of the skull. Chinese clinical oncology 2016;5.
209. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. J Manag Care Pharm 2003;9:53-61.
210. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. American journal of clinical oncology 1982;5:649-56.
211. ACR/ASTRO Practice Parameter for the Performance of Proton Beam Radiation Therapy. 2018. (Accessed 08/01/18, at https://www.astro.org/Patient-Care-and-Research/Clinical-Practice-Statements/Practice-Parameter-for-Proton-Therapy.)
212. Ono T, Nakamura T, Yamaguchi H, et al. Clinical results of proton beam therapy for elderly patients with non-small cell lung cancer. Radiation oncology (London, England) 2018;13:19.
213. Oshiro Y, Mizumoto M, Okumura T, et al. Analysis of repeated proton beam therapy for patients with hepatocellular carcinoma. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;123:240-5.
214. Ovalle V, Strom EA, Shaitelman S, et al. Proton Partial Breast Irradiation: Detailed Description of Acute Clinico-Radiologic Effects. Cancers 2018;10:111.

215. Paganetti H, Niemierko A, Ancukiewicz M, et al. Relative biological effectiveness (RBE) values for proton beam therapy. International Journal of Radiation Oncology* Biology* Physics 2002;53:407-21.
216. Pan HY, Jiang J, Hoffman KE, et al. Comparative Toxicities and Cost of Intensity-Modulated Radiotherapy, Proton Radiation, and Stereotactic Body Radiotherapy Among Younger Men With Prostate Cancer. J Clin Oncol 2018;36:1823-30.
217. Papakostas TD, Lane AM, Morrison M, Gragoudas ES, Kim IK. Long-term Outcomes After Proton Beam Irradiation in Patients With Large Choroidal Melanomas. JAMA ophthalmology 2017;135:1191-6.
218. Park Y, Yu ES, Ha B, Park HJ, Kim JH, Kim JY. Neurocognitive and Psychological Functioning of Children with an Intracranial Germ Cell Tumor. Cancer research and treatment : official journal of Korean Cancer Association 2017;49:960-9.
219. Patel AV, Lane AM, Morrison MA, et al. Visual Outcomes after Proton Beam Irradiation for Choroidal Melanomas Involving the Fovea. Ophthalmology 2016;123:369-77.
220. Patel SH, Wang Z, Wong WW, et al. Charged particle therapy versus photon therapy for paranasal sinus and nasal cavity malignant diseases: a systematic review and meta-analysis. The Lancet Oncology 2014;15:1027-38.
221. Paulino AC, Mahajan A, Ye R, et al. Ototoxicity and cochlear sparing in children with medulloblastoma: Proton vs. photon radiotherapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2018;128:128-32.
222. Pennicooke B, Laufer I, Sahgal A, et al. Safety and local control of radiation therapy for chordoma of the spine and sacrum: a systematic review. Spine 2016;41:S186.
223. Petrovic A, Bergin C, Schalenbourg A, Goitein G, Zografos L. Proton therapy for uveal melanoma in 43 juvenile patients: long-term results. Ophthalmology 2014;121:898-904.
224. Phan J, Sio TT, Nguyen TP, et al. Reirradiation of Head and Neck Cancers With Proton Therapy: Outcomes and Analyses. International journal of radiation oncology, biology, physics 2016;96:30-41.
225. Polishchuk AL, Mishra KK, Weinberg V, et al. Temporal Evolution and Dose-Volume Histogram Predictors of Visual Acuity After Proton Beam Radiation Therapy of Uveal Melanoma. International journal of radiation oncology, biology, physics 2017;97:91-7.
226. Puga Gonzalez B, Ferrandez Longas A, Oyarzabal M, Nosas R. The effects of growth hormone deficiency and growth hormone replacement therapy on intellectual ability, personality and adjustment in children. Pediatr Endocrinol Rev 2010;7:328-38.
227. Pugh TJ, Choi S, Nogueras-Gonzalaez GM, et al. Proton beam therapy for localized prostate cancer: Results from a prospective quality-of-life trial. International Journal of Particle Therapy 2016;3:27-36.
228. Pulsifer MB, Duncanson H, Grieco J, et al. Cognitive and Adaptive Outcomes After Proton Radiation for Pediatric Patients With Brain Tumors. International journal of radiation oncology, biology, physics 2018.
229. Radiology ACo. ACR Appropriateness Criteria. 2014-2018.
230. Rahmi A, Mammar H, Thariat J, et al. Proton beam therapy for presumed and confirmed iris melanomas: a review of 36 cases. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie 2014;252:1515-21.

231. Rassi MS, Hulou MM, Almefty K, et al. Pediatric Clival Chordoma: A Curable Disease that Conforms to Collins' Law. Neurosurgery 2018;82:652-60.
232. Ray GL, Buchsbaum JC, McMullen KP, et al. Definitive treatment of leptomeningeal spinal metastases in children. Pediatric blood & cancer 2013;60:1839-41.
233. Remick JS, Schonewolf C, Gabriel P, et al. First Clinical Report of Proton Beam Therapy for Postoperative Radiotherapy for Non-Small-Cell Lung Cancer. Clinical lung cancer 2017;18:364-71.
234. Reynolds CR, Kamphaus RW, Vannest KJ. Behavior assessment system for children (BASC). Encyclopedia of clinical neuropsychology: Springer; 2011:366-71.
235. Riechardt AI, Cordini D, Dobner B, et al. Salvage proton beam therapy in local recurrent uveal melanoma. American journal of ophthalmology 2014;158:948-56.
236. Riechardt AI, Pilger D, Cordini D, et al. Neovascular glaucoma after proton beam therapy of choroidal melanoma: incidence and risk factors. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie 2017;255:2263-9.
237. Roid GH, Miller LJ, Pomplun M, Koch C. Leiter international performance scale,(Leiter-3). Los Angeles: Western Psychological Services 2013.
238. Romesser PB, Cahlon O, Scher E, et al. Proton beam radiation therapy results in significantly reduced toxicity compared with intensity-modulated radiation therapy for head and neck tumors that require ipsilateral radiation. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;118:286-92.
239. Romesser PB, Cahlon O, Scher ED, et al. Proton Beam Reirradiation for Recurrent Head and Neck Cancer: Multi-institutional Report on Feasibility and Early Outcomes. International journal of radiation oncology, biology, physics 2016;95:386-95.
240. Rosenthal DI, Mendoza TR, Chambers MS, et al. Measuring head and neck cancer symptom burden: the development and validation of the M. D. Anderson symptom inventory, head and neck module. Head & neck 2007;29:923-31.
241. Rotondo RL, Folkert W, Liebsch NJ, et al. High-dose proton-based radiation therapy in the management of spine chordomas: outcomes and clinicopathological prognostic factors. Journal of neurosurgery Spine 2015;23:788-97.
242. Rwigema JM, Verma V, Lin L, et al. Prospective study of proton-beam radiation therapy for limited-stage small cell lung cancer. Cancer 2017;123:4244-51.
243. Sandinha MT, Kacperek A, Errington RD, Coupland SE, Damato B. Recurrence of iris melanoma after proton beam therapy. The British journal of ophthalmology 2014;98:484-7.
244. Sanford NN, Pursley J, Noe B, et al. Protons versus Photons for Unresectable Hepatocellular Carcinoma: Liver Decompensation and Overall Survival. International journal of radiation oncology, biology, physics 2019.
245. Sato M, Gunther JR, Mahajan A, et al. Progression-free survival of children with localized ependymoma treated with intensity-modulated radiation therapy or proton-beam radiation therapy. Cancer 2017;123:2570-8.
246. Schonfeld S, Cordini D, Riechardt AI, et al. Proton beam therapy leads to excellent local control rates in choroidal melanoma in the intermediate fundus zone. American journal of ophthalmology 2014;158:1184-91.

247. Schünemann HJ, Tugwell P, Reeves BC, et al. Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions. Research synthesis methods 2013;4:49-62.
248. Seibel I, Cordini D, Hager A, et al. Predictive risk factors for radiation retinopathy and optic neuropathy after proton beam therapy for uveal melanoma. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie 2016;254:1787-92.
249. Seibel I, Riechardt AI, Heufelder J, Cordini D, Joussen AM. Adjuvant Ab Interno Tumor Treatment After Proton Beam Irradiation. American journal of ophthalmology 2017;178:94-100.
250. Sethi RV, Giantsoudi D, Raiford M, et al. Patterns of failure after proton therapy in medulloblastoma; linear energy transfer distributions and relative biological effectiveness associations for relapses. International journal of radiation oncology, biology, physics 2014;88:655-63.
251. Sharma S, Zhou O, Thompson R, et al. Quality of Life of Postoperative Photon versus Proton Radiation Therapy for Oropharynx Cancer. International Journal of Particle Therapy 2018;5:11-7.
252. Shekelle PG, Newberry SJ, Wu H, et al. AHRQ Methods for Effective Health Care. Identifying Signals for Updating Systematic Reviews: A Comparison of Two Methods. Rockville (MD): Agency for Healthcare Research and Quality (US); 2011.
253. Sher DJ, Tishler RB, Pham N-L, Punglia RS. Cost-Effectiveness Analysis of Intensity Modulated Radiation Therapy Versus Proton Therapy for Oropharyngeal Squamous Cell Carcinoma. International journal of radiation oncology, biology, physics 2018;101:875-82.
254. Radiation Therapy: Which type is right for me? OncoLink Penn Medicine, 2018. at https://www.oncolink.org/cancer-treatment/radiation/introduction-to-radiation-therapy/radiation-therapy-which-type-is-right-for-me.)
255. Shiraishi Y, Fang P, Xu C, et al. Severe lymphopenia during neoadjuvant chemoradiation for esophageal cancer: A propensity matched analysis of the relative risk of proton versus photon-based radiation therapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2018;128:154-60.
256. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Annals of internal medicine 2007;147:224-33.
257. Sieghart W, Hucke F, Peck-Radosavljevic M. Transarterial chemoembolization: modalities, indication, and patient selection. Journal of hepatology 2015;62:1187-95.
258. Sikuade MJ, Salvi S, Rundle PA, Errington DG, Kacperek A, Rennie IG. Outcomes of treatment with stereotactic radiosurgery or proton beam therapy for choroidal melanoma. Eye (London, England) 2015;29:1194-8.
259. Simon F, Feuvret L, Bresson D, et al. Surgery and protontherapy in Grade I and II skull base chondrosarcoma: A comparative retrospective study. PLoS One 2018;13:e0208786.
260. Sio TT, Lin HK, Shi Q, et al. Intensity Modulated Proton Therapy Versus Intensity Modulated Photon Radiation Therapy for Oropharyngeal Cancer: First Comparative Results of Patient-Reported Outcomes. International journal of radiation oncology, biology, physics 2016;95:1107-14.
261. Skolarus TA, Dunn RL, Sanda MG, et al. Minimally important difference for the expanded prostate cancer index composite short form. Urology 2015;85:101-6.
262. Sloan JA, Dueck A. Issues for statisticians in conducting analyses and translating results for quality of life end points in clinical trials. Journal of biopharmaceutical statistics 2004;14:73-96.

263. Smith AR. Vision 20/20: proton therapy. Medical physics 2009;36:556-68.
264. Snider JW, Schneider RA, Poelma-Tap D, et al. Long-Term Outcomes and Prognostic Factors After Pencil-Beam Scanning Proton Radiation Therapy for Spinal Chordomas: A Large, Single-Institution Cohort. International journal of radiation oncology, biology, physics 2018;101:226-33.
265. Song S, Park HJ, Yoon JH, et al. Proton beam therapy reduces the incidence of acute haematological and gastrointestinal toxicities associated with craniospinal irradiation in pediatric brain tumors. Acta oncologica (Stockholm, Sweden) 2014;53:1158-64.
266. StataCorp. Stata Statistical Software: Release 11. College Station, TX. 2009.
267. Stieb S, Snider JW, 3rd, Placidi L, et al. Long-Term Clinical Safety of High-Dose Proton Radiation Therapy Delivered With Pencil Beam Scanning Technique for Extracranial Chordomas and Chondrosarcomas in Adult Patients: Clinical Evidence of Spinal Cord Tolerance. International journal of radiation oncology, biology, physics 2018;100:218-25.
268. Swisher-McClure S, Hahn SM, Bekelman J. Proton beam therapy: the next disruptive innovation in healthcare? : The Fellowship of Postgraduate Medicine; 2015.
269. Takada A, Nakamura T, Takayama K, et al. Preliminary treatment results of proton beam therapy with chemoradiotherapy for stage I-III esophageal cancer. Cancer Med 2016;5:506-15.
270. Takagi M, Demizu Y, Terashima K, et al. Long-term outcomes in patients treated with proton therapy for localized prostate cancer. Cancer Med 2017;6:2234-43.
271. Takaoka EI, Miyazaki J, Ishikawa H, et al. Long-term single-institute experience with trimodal bladder-preserving therapy with proton beam therapy for muscle-invasive bladder cancer. Japanese journal of clinical oncology 2017;47:67-73.
272. Takayama K, Nakamura T, Takada A, et al. Treatment results of alternating chemoradiotherapy followed by proton beam therapy boost combined with intra-arterial infusion chemotherapy for stage III-IVB tongue cancer. Journal of cancer research and clinical oncology 2016;142:659-67.
273. Taylor CW, Wang Z, Macaulay E, Jagsi R, Duane F, Darby SC. Exposure of the Heart in Breast Cancer Radiation Therapy: A Systematic Review of Heart Doses Published During 2003 to 2013. International journal of radiation oncology, biology, physics 2015;93:845-53.
274. Teichman SL, Do S, Lum S, et al. Improved long-term patient-reported health and well-being outcomes of early-stage breast cancer treated with partial breast proton therapy. Cancer Med 2018;7:6064-76.
275. Thariat J, Grange JD, Mosci C, et al. Visual Outcomes of Parapapillary Uveal Melanomas Following Proton Beam Therapy. International journal of radiation oncology, biology, physics 2016;95:328-35.
276. Thariat J, Jacob S, Caujolle JP, et al. Cataract Avoidance With Proton Therapy in Ocular Melanomas. Investigative ophthalmology & visual science 2017b;58:5378-86.
277. Thariat J, Maschi C, Lanteri S, et al. Dry Eye Syndrome After Proton Therapy of Ocular Melanomas. International journal of radiation oncology, biology, physics 2017a;98:142-51.
278. Thariat J, Rahmi A, Salleron J, et al. Proton Beam Therapy for Iris Melanomas in 107 Patients. Ophthalmology 2017c;125:606-14.
279. Thorp N. Proton therapy for children. RAD Magazine 2010;36:17-8.
280. Tian X, Liu K, Hou Y, Cheng J, Zhang J. The evolution of proton beam therapy: Current and future status. Mol Clin Oncol 2018;8:15-21.
281. Toyomasu Y, Demizu Y, Matsuo Y, et al. Outcomes of Patients With Sinonasal Squamous Cell Carcinoma Treated With Particle Therapy Using Protons or Carbon Ions. International journal of radiation oncology, biology, physics 2018;101:1096-103.

282. Tucker SL, Liu A, Gomez D, et al. Impact of heart and lung dose on early survival in patients with non-small cell lung cancer treated with chemoradiation. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;119:495-500.
283. Unkelbach J, Chan TC, Bortfeld T. Accounting for range uncertainties in the optimization of intensity modulated proton therapy. Physics in Medicine & Biology 2007;52:2755.
284. Proton Therapy: Behind the Scenes. OncoLink Penn Medicine, 2018. at https://www.oncolink.org/cancer-treatment/proton-therapy/overviews-of-proton-therapy/proton-therapy-behind-the-scenes.)
285. Vargas CE, Hartsell WF, Dunn M, et al. Image-guided hypofractionated proton beam therapy for low-risk prostate cancer: Analysis of quality of life and toxicity, PCG GU 002. Reports of practical oncology and radiotherapy : journal of Greatpoland Cancer Center in Poznan and Polish Society of Radiation Oncology 2016;21:207-12.
286. Vargas CE, Hartsell WF, Dunn M, et al. Hypofractionated Versus Standard Fractionated Proton-beam Therapy for Low-risk Prostate Cancer: Interim Results of a Randomized Trial PCG GU 002. Am J Clin Oncol 2018;41:115-20.
287. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in pediatric cancer: reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. Cancer 2002;94:2090-106.
288. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL™* 4.0 as a pediatric population health measure: feasibility, reliability, and validity. Ambulatory pediatrics 2003;3:329-41.
289. Varni JW, Limbers CA, Burwinkle TM. Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL 4.0 Generic Core Scales. Health and quality of life outcomes 2007;5:2.
290. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Medical care 1999;37:126-39.
291. Ventura LM, Grieco JA, Evans CL, et al. Executive functioning, academic skills, and quality of life in pediatric patients with brain tumors post-proton radiation therapy. Journal of neuro-oncology 2018;137:119-26.
292. Verma V, Iftekaruddin Z, Badar N, et al. Proton beam radiotherapy as part of comprehensive regional nodal irradiation for locally advanced breast cancer. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;123:294-8.
293. Verma V, Lin SH, Simone CB. Clinical outcomes and toxicities of proton radiotherapy for gastrointestinal neoplasms: a systematic review. Journal of gastrointestinal oncology 2016;7:644.
294. Verma V, Mehta M. Clinical outcomes of proton radiotherapy for uveal melanoma. Clinical Oncology 2016;28:e17-e27.
295. Verma V, Rwigema J-CM, Malyapa RS, Regine WF, Simone II CB. Systematic assessment of clinical outcomes and toxicities of proton radiotherapy for reirradiation. Radiotherapy and Oncology 2017;125:21-30.
296. Verma V, Shah C, Mehta MP. Clinical outcomes and toxicity of proton radiotherapy for breast cancer. Clinical breast cancer 2016;16:145-54.
297. Verma V, Simone CB, Mishra MV. Quality of life and patient-reported outcomes following proton radiation therapy: a systematic review. JNCI: Journal of the National Cancer Institute 2017;110:341-53.

298. Vern-Gross TZ, Indelicato DJ, Bradley JA, Rotondo RL. Patterns of Failure in Pediatric Rhabdomyosarcoma After Proton Therapy. International journal of radiation oncology, biology, physics 2016;96:1070-7.
299. Vlachogiannis P, Gudjonsson O, Montelius A, et al. Hypofractionated high-energy proton-beam irradiation is an alternative treatment for WHO grade I meningiomas. Acta neurochirurgica 2017;159:2391-400.
300. Vogel J, Both S, Kirk M, et al. Proton therapy for pediatric head and neck malignancies. Pediatric blood & cancer 2018;65:e26858.
301. Vogin G, Calugaru V, Bolle S, et al. Investigation of ectopic recurrent skull base and cervical chordomas: The Institut Curie's proton therapy center experience. Head & neck 2016;38 Suppl 1:E1238-46.
302. Walsh J. Proton Beam Therapy for Prostate Cancer: A Technology Assessment. San Francisco, CA: California Technology Assessment Forum (CTAF)
2012.
303. Wang XS, Shi Q, Williams LA, et al. Prospective Study of Patient-Reported Symptom Burden in Patients With Non-Small-Cell Lung Cancer Undergoing Proton or Photon Chemoradiation Therapy. J Pain Symptom Manage 2016;51:832-8.
304. Wattson DA, Tanguturi SK, Spiegel DY, et al. Outcomes of proton therapy for patients with functional pituitary adenomas. International journal of radiation oncology, biology, physics 2014;90:532-9.
305. Weber B, Paton K, Ma R, Pickles T. Outcomes of Proton Beam Radiotherapy for Large Non-Peripapillary Choroidal and Ciliary Body Melanoma at TRIUMF and the BC Cancer Agency. Ocular oncology and pathology 2015;2:29-35.
306. Weber DC, Ares C, Albertini F, et al. Pencil Beam Scanning Proton Therapy for Pediatric Parameningeal Rhabdomyosarcomas: Clinical Outcome of Patients Treated at the Paul Scherrer Institute. Pediatric blood & cancer 2016;63:1731-6.
307. Weber DC, Ares C, Malyapa R, et al. Tumor control and QoL outcomes of very young children with atypical teratoid/rhabdoid tumor treated with focal only chemo-radiation therapy using pencil beam scanning proton therapy. Journal of neuro-oncology 2015;121:389-97.
308. Weber DC, Malyapa R, Albertini F, et al. Long term outcomes of patients with skull-base low-grade chondrosarcoma and chordoma patients treated with pencil beam scanning proton therapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;120:169-74.
309. Weber DC, Murray F, Combescure C, et al. Long term outcome of skull-base chondrosarcoma patients treated with high-dose proton therapy with or without conventional radiation therapy. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2018.
310. Weber DC, Murray FR, Correia D, et al. Pencil beam scanned protons for the treatment of patients with Ewing sarcoma. Pediatric blood & cancer 2017;64:e26688.
311. Wei JT, Dunn RL, Litwin MS, Sandler HM, Sanda MG. Development and validation of the expanded prostate cancer index composite (EPIC) for comprehensive assessment of health-related quality of life in men with prostate cancer. Urology 2000;56:899-905.
312. Willerding GD, Cordini D, Hackl C, et al. Proton beam radiotherapy of diffuse iris melanoma in 54 patients. The British journal of ophthalmology 2015;99:812-6.

313. Willerding GD, Cordini D, Moser L, Krause L, Foerster MH, Bechrakis NE. Neoadjuvant proton beam irradiation followed by transscleral resection of uveal melanoma in 106 cases. The British journal of ophthalmology 2016;100:463-7.
314. Williams PE, Weiss PDLG, Rolfhus PDE. WISC–IV Technical Report# 1 Theoretical Model and Test Blueprint. San Antonio, TX: The Psychological Corporation 2003.
315. Woodcock RW, McGrew KS, Mather N, Schrank F. Woodcock-Johnson R III NU Tests of Achievement. Itasca, IL: Riverside 2001.
316. Wray J, Flampouri S, Slayton W, et al. Proton Therapy for Pediatric Hodgkin Lymphoma. Pediatric blood & cancer 2016;63:1522-6.
317. Xi M, Xu C, Liao Z, et al. Comparative Outcomes After Definitive Chemoradiotherapy Using Proton Beam Therapy Versus Intensity Modulated Radiation Therapy for Esophageal Cancer: A Retrospective, Single-Institutional Analysis. International journal of radiation oncology, biology, physics 2017;99:667-76.
318. Yeung R, Bowen SR, Chapman TR, MacLennan GT, Apisarnthanarax S. Chest wall toxicity after hypofractionated proton beam therapy for liver malignancies. Pract Radiat Oncol 2018;8:287-93.
319. Yock TI, Tarbell NJ. Technology insight: Proton beam radiotherapy for treatment in pediatric brain tumors. Nature Reviews Clinical Oncology 2004;1:97.
320. Yock TI, Yeap BY, Ebb DH, et al. Long-term toxic effects of proton radiotherapy for paediatric medulloblastoma: a phase 2 single-arm study. The Lancet Oncology 2016;17:287-98.
321. Yoon H, Shah K, Small W, Mehta MP, Hayes JP. Basic Concepts of Clinical Radiation Oncology. In: Small W, ed. Clinical Radiation Oncology: Indications, Techniques, and Results, 3rd Edition: John Wiley & Sons, Inc.; 2017:1-14.
322. Yu JI, Yoo GS, Cho S, et al. Initial clinical outcomes of proton beam radiotherapy for hepatocellular carcinoma. Radiation oncology journal 2018;36:25-34.
323. Zeisberg A, Seibel I, Cordini D, et al. Long-term (4 years) results of choroidal hemangioma treated with proton beam irradiation. Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie 2014;252:1165-70.
324. Zenda S, Akimoto T, Mizumoto M, et al. Phase II study of proton beam therapy as a nonsurgical approach for mucosal melanoma of the nasal cavity or para-nasal sinuses. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2016;118:267-71.
325. Zenda S, Kawashima M, Arahira S, et al. Late toxicity of proton beam therapy for patients with the nasal cavity, para-nasal sinuses, or involving the skull base malignancy: importance of long-term follow-up. International journal of clinical oncology 2015;20:447-54.
326. Zhang W, Zhang X, Yang P, et al. Intensity-modulated proton therapy and osteoradionecrosis in oropharyngeal cancer. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 2017;123:401-5.
327. Zhang Z, Zhu Y, Zhang L, Wang Z, Wan H. Prediction model of critical weight loss in cancer patients during particle therapy. Japanese journal of clinical oncology 2018;48:75-81.
328. Zhou J, Yang B, Wang X, Jing Z. Comparison of the effectiveness of radiotherapy with photons and particles for chordoma after surgery: a meta-analysis. World neurosurgery 2018.