Uniform Medical Plan coverage limits

Updates effective 11/1/2018

The benefit coverage limits listed below apply to these UMP plans:
Uniform Medical Plan Classic (UMP Classic)
UMP Consumer-Directed Health Plan (UMP CDHP)
  - UMP Plus–Puget Sound High Value Network
  - UMP Plus–UW Medicine Accountable Care Network

Some services listed under these benefits have coverage limits. These limits are either determined by a Health Technology Clinical Committee (HTCC) decision or a Regence BlueShield medical policy. The table below does not include every limit or exclusion under this benefit. For more details, refer to your plan's Certificate of Coverage.

Laboratory

<table>
<thead>
<tr>
<th>These services or supplies have coverage limits</th>
<th>The rules or policies that define the coverage limits</th>
<th>Limit applies to these codes (chosen by your provider to bill for services)</th>
</tr>
</thead>
</table>
| Laboratory and Genetic Testing for use of Thiopurines                                                          | Regence Medical Policy Lab70                          | • 81335, 81401, 0034U  
  • UMP is subject to HTCC decision for codes 81335 and 0034U.  
  • Codes 81335 and 0034U will deny as not a covered benefit when billed with the following dx: depression, mood disorders, psychosis, anxiety, ADHD and substance use disorders. |

Maternity

Elective early delivery, prior to 39 weeks' gestation is not a covered benefit (not applicable to emergency delivery or spontaneous labor).

Medicine

<table>
<thead>
<tr>
<th>These services or supplies have coverage limits</th>
<th>The rules or policies that define the coverage limits</th>
<th>Limit applies to these codes (chosen by your provider to bill for services)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confocal Laser Endomicroscopy</td>
<td>Regence Medical Policy Med151</td>
<td>• 43206, 43252, 88375</td>
</tr>
<tr>
<td>Coverage of Treatments Provided in a Clinical Trial</td>
<td>Regence Medical Policy Med150</td>
<td>• S9990, S9991, S9988</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Corneal Collagen Cross-Linking</td>
<td>Regence Medical Policy Med159</td>
<td>• 0402T</td>
</tr>
</tbody>
</table>
| Gait Analysis                                    | Regence Medical Policy Med107 | • 96000, 96001, 96002, 96003, 96004  
Gait analysis may be considered medically necessary in children and adolescents with cerebral palsy to select surgical or other therapeutic interventions for gait improvement. All other indications for gait analysis and **Paraspinal Surface Electromyography (EMG)** are considered investigational. |
| Hyperbaric Oxygen Therapy for Disuse Damage Including Wound Care and Treatment of Central Nervous System Conditions | HTCC decision               | • 99183  
• G0277  
**Regence Medical Policy** is used only to determine units of treatment, criteria for diabetic "standard wound therapy" and to address any conditions not addressed in the HTCC decisions under the HTCC "limitations of coverage" or "non-covered indicators". |
| In Vivo Analysis of Colorectal Polyps            | Regence Medical Policy Med104 | • 88375                 |
| Intensity Modulated Radiotherapy (IMRT)           | HTCC decision               | • 77301, 77338, 77385, 77386  
• G6015, G6016 | |
| Orthopedic Applications of Stem-Cell Therapy     | Regence Medical Policy Med05 | • 38206, 38232, 38241    |
| Charged-Particle (Proton or Helium Ion) Radiotherapy | HTCC decision               | • 77520, 77522, 77523, 77525 |
| Radioembolization for Primary and Metastatic Tumors of the Liver | Regence Medical Policy Med140 | • 37243, 79445  
• S2095  
Note: **Regence Medical Policy Ovarian and Internal Iliac Vein Embolization as a Treatment of Pelvic Congestion** is considered investigational. |
<table>
<thead>
<tr>
<th>Surface Electromyography (SEMG)</th>
<th><strong>Regence Medical Policy Med73</strong></th>
<th>• 96002, 96004</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transgender Services</td>
<td><strong>Regence Medical Policy Med153</strong></td>
<td>• <strong>17380, 19325</strong>, 55970, 55980 Note: Codes 55970 and 55980 are non-specific. The specific procedure code(s) must be requested in place of these non-specific codes. Surgical treatments of gender dysphoria require pre-authorization. Check codes for specific procedures listed in other areas of this pre-authorization list (for example, breast reconstruction, blepharoplasty, rhinoplasty and abdominoplasty) that require pre-authorization, which also apply to transgender surgical services. Pre-authorization is not required for mastectomy related to breast cancer or for breast reconstruction and nipple/areola reconstruction following procedure related to breast cancer. • 00103, 15820, 15821, 15822, 15823, 15847, 19303, 19304, 19316, 19318, 19324, 19325, 19350, 30400, 30410, 30420, 30430, 30435, 30450, 31551, 31552, 31553, 31554, 31580, 31584, 31587, 31591, 53400, 53405, 53410, 53415, 53420, 53425, 53430, 54125, 54660, 55175, 55180, 55625, 56800, 56805, 57106, 57110, 57291, 57292, 57295, 57296, 57335, 57426, 58150, 58180, 58260, 58262, 58275, 58290, 58291, 58541, 58542, 58543, 58544, 58550, 58552, 58553, 58554, 58570, 58571, 58572, 58573, 58720, C1813 UMP specific: <a href="#">Transgender policy</a></td>
</tr>
</tbody>
</table>
Laboratory and Genetic Testing for use of Thiopurines

Effective: March 1, 2018

Next Review: January 2019
Last Review: January 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Thiopurines or purine analogues are immunomodulators. They are used to treat malignancies, rheumatic diseases, dermatologic conditions, and irritable bowel disease, and are used in solid organ transplantation. The tests addressed in this policy are used to help identify patients at increased risk of developing severe, life-threatening myelotoxicity from thiopurines and to aid in determining the initial dose and evaluate any ongoing dosing.

MEDICAL POLICY CRITERIA

I. Genotypic or phenotypic analysis of the thiopurine methyltransferase (TPMT) enzyme may be considered medically necessary in patients prior to beginning thiopurine therapy (i.e. azathioprine, mercaptopurine, or thioguanine) OR in patients on thiopurine therapy when there is clinical documentation of abnormal complete blood count results that do not respond to dose reduction.

II. Genotypic and/or phenotypic analysis of the TPMT enzyme is considered investigational in all other situations.

III. Analysis of the metabolite markers azathioprine and mercaptopurine, including 6-methyl-mercaptopurine ribonucleotides and 6-thioguanine nucleotides, is considered investigational.
IV. Genetic testing for NUDT15 may be considered **medically necessary** in patients prior to beginning thiopurine therapy (i.e. azathioprine, mercaptopurine, or thioguanine) OR in patients on thiopurine therapy when there is clinical documentation of abnormal complete blood count results that do not respond to dose reduction.

V. Genetic testing for NUDT15 is considered **investigational** in all other situations.

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.

**POLICY GUIDELINES**

In order to determine the clinical utility of gene test(s), all of the following information must be submitted for review. If any of these items are not submitted, it could impact our review and decision outcome:

1. Name of the genetic test(s) or panel test
2. Name of the performing laboratory and/or genetic testing organization (more than one may be listed)
3. The exact gene(s) and/or mutation(s) being tested
4. Relevant billing codes
5. Brief description of how the genetic test results will guide clinical decisions that would not otherwise be made in the absence testing
6. Medical records related to this genetic test:
   - History and physical exam including any relevant diagnoses related to the genetic testing
   - Conventional testing and outcomes
   - Conservative treatments, if any

Thiopurine methyltransferase (TPMT) testing cannot substitute for complete blood count monitoring in patients receiving thiopurines. Early drug discontinuation may be considered in patients with abnormal complete blood count results. Dosage reduction is recommended in patients with reduced TPMT activity. Alternative therapies may need to be considered for patients who have low or absent TPMT activity (homozygous for nonfunctional alleles). Accurate phenotyping results are not possible in patients who received recent blood transfusions. TPMT genotyping and phenotyping would only need to be performed once.

**CROSS REFERENCES**

None

**BACKGROUND**

The thiopurine drugs—which include azathioprine (a pro-drug for mercaptopurine), mercaptopurine, and thioguanine—are used to treat a variety of diseases; however, it is recommended that the use of thiopurines be limited due to a high rate of drug toxicity. Mercaptopurine and thioguanine are directly metabolized by the thiopurine S-methyltransferase (TPMT) enzyme. Susceptibility to drug toxicity is linked to the level of TPMT activity. The variation in TPMT activity has been related to three distinct TPMT variants. Pharmacogenomic analysis of TPMT status is proposed to identify patients at risk of thiopurine toxicity.
drug toxicity and adjust medication doses accordingly; measurement of metabolite markers has also been proposed.

**THIOPURINES**

Thiopurines or purine analogues are immunomodulators. They include azathioprine (Imuran), mercaptopurine (6-MP; Purinethol), and thioguanine (6-TG; Tabloid). Thiopurines are used to treat malignancies, rheumatic diseases, dermatologic conditions, and irritable bowel disease, and are used in solid organ transplantation. They are considered an effective immunosuppressive treatment of irritable bowel disease, particularly in patients with corticosteroid-resistant disease. However, use of thiopurines is limited by both its long onset of action (3-4 months) and drug toxicities, which include hepatotoxicity, bone marrow suppression, pancreatitis, and allergic reactions.

**Pharmacogenomics**

Thiopurines are converted to 6-MP in vivo, where it is subsequently metabolized to two active metabolites: either 6-thioguanine nucleotides (6-TGN) by the inosine-5'-monophosphate dehydrogenase (IMPDH) enzyme; or to 6-methyl-mercaptopurine ribonucleotides (6-MMPR) by the thiopurine methyltransferase (TPMT) enzyme. TPMT also converts 6-MP into an inactive metabolite, 6-methyl-mercaptopurine. 6-TGNs are considered cytotoxic and thus are associated with bone marrow suppression, while 6-MMPR is associated with hepatotoxicity. In population studies, the activity of the TPMT enzyme has been shown to be trimodal, with 90% of subjects having high activity, 10% intermediate activity, and 0.3% with low or no activity. In patients with intermediate-to-low activity, the metabolism of 6-MP is shunted toward the IMPDH pathway with greater accumulation of 6-TGN; these patients are considered at risk for myelotoxicity (i.e., bone marrow suppression).

This variation in TPMT activity has been related to three distinct TPMT variants and has permitted the development of TPMT genotyping based on a polymerase chain reaction. For example, patients with high TPMT activity are found to have two normal (wild-type) TPMT alleles; those with intermediate activity are heterozygous (i.e., have a variant on one chromosome), while those with low TPMT activity are homozygous for TPMT variants (i.e., a variant is found on both chromosomes). Genetic analysis has been explored as a technique to identify patients at risk for myelotoxicity; those with intermediate TPMT activity may be initially treated with lower doses of thiopurines, while those with low TPMT activity may not be good candidates for thiopurine therapy.

TPMT activity can also be measured by phenotypic testing. Phenotypic testing determines the level of thiopurine nucleotides or TPMT activity in erythrocytes and can also be informative. Caution must be taken with phenotyping, because some coadministered drugs can influence the measurement of TPMT activity in blood, and recent blood transfusions will misrepresent a patient's actual TPMT activity.

Prospective TPMT genotyping or phenotyping may help identify patients at increased risk of developing severe, life-threatening myelotoxicity.

The *NUDT15* gene encodes a Nudix hydrolase, a family of enzymes that catalyze the hydrolysis of nucleoside diphosphates. NUDT15 has been proposed to participate in the catabolism of thiopurines and act as a negative regulator of thiopurine activation and toxicity.
Correlations have been shown between NUDT15 variants and thiopurine toxicity. Thus genetic analysis has been examined as a method to identify those at risk of thiopurine-induced toxicity.

**Metabolite Markers**

Monitoring of thiopurine therapy has been based on clinical assessment of response in addition to monitoring blood cell counts, liver function, and pancreatic function tests. However, there has been interest in monitoring intracellular levels of thiopurine metabolites (i.e., 6-TGN, 6-MMPR) to predict response and complications, with the ultimate aim of tailoring drug therapy to each individual patient.

While genotyping and phenotyping of TPMT would only be performed once, metabolite markers might be tested multiple times during the course of the disease to aid in determining the initial dose and also evaluate any ongoing dosing.

**REGULATORY STATUS**

Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; laboratory-developed tests must meet the general regulatory standards of the Clinical Laboratory Improvement Amendments. Several thiopurine genotype, phenotype, and metabolite tests are available under the auspices of the Clinical Laboratory Improvement Amendments. Laboratories that offer laboratory-developed tests must be licensed by the Clinical Laboratory Improvement Amendments for high-complexity testing. To date, the U.S. Food and Drug Administration has chosen not to require any regulatory review of this test.

Prometheus®, a commercial laboratory in San Diego, offers thiopurine genotype, phenotype, and metabolite testing for those on thiopurine therapy. The tests are referred to as Prometheus® TPMT Genetics, Prometheus® TPMT enzyme, and Prometheus® thiopurine metabolites, respectively. Other laboratories that offer TPMT genotyping include Quest Diagnostics (TPMT Genotype; Madison, NJ), ARUP Laboratories (TPMT DNA; Salt Lake City, UT), and Specialty Laboratories (TPMT GenoTypR™; Valencia, CA), PreventionGenetics (TPMT Deficiency via the TPMT Gene; Marshfield, WI), Genelex (TPMT; Seattle, WA), and Fulgent Genetics (TPMT; Temple City, CA).

**EVIDENCE SUMMARY**

Human Genome Variation Society (HGVS) nomenclature[1] is used to describe variants found in DNA and serves as an international standard. It is being implemented for genetic testing medical evidence review updates starting in 2017. According to this nomenclature, the term “variant” is used to describe a change in a DNA or protein sequence, replacing previously-used terms, such as “mutation.” Pathogenic variants are variants associated with disease, while benign variants are not. The majority of genetic changes have unknown effects on human health, and these are referred to as variants of uncertain significance.

**ANALYTIC VALIDITY**

**TPMT Genotype and Phenotype Testing**

The genotypic analysis of the TPMT gene is based on well-established polymerase chain reaction technology to detect three distinct variants. Currently, three alleles (TPMT*2, TPMT*3A, TPMT*3C) account for about 95% of subjects with reduced TPMT enzyme activity. Subjects homozygous for these alleles are TPMT-deficient and those heterozygous for these
alleles have variable TPMT (low or intermediate) activity. A 2012 study from Sweden addressed the concordance between TPMT genotyping and phenotyping.[2] The investigators evaluated data from 7195 unselected and consecutive TPMT genotype and phenotype tests. The genotype tests examined the three most common TPMT variants, previously noted. TPMT genotyping identified 89% as TPMT wild-type, 704 (10%) as TPMT heterozygous, and 37 (0.5%) as TPMT homozygous. The overall agreement between genotyping and phenotyping was 95%. Genotyping alone would have misclassified three (8%) of 37 homozygous patients as heterozygous; these three subjects were found to have uncommon mutations. All three had low TPMT activity. The phenotype test would have misclassified 4 (11%) of 37 of homozygous patients because they had test results above the cutoff level for low TPMT activity (<2.5 U/mL red blood cells).

Metabolite Marker Testing

Reproducibility and sample stability for metabolite testing have been assessed in a small number of studies. Levels of 6-TGN were shown in a study by Cuffari (2000) to be reproducible with <10% variability when tested in patients on stable doses of thiopurines over 2 to 24 months.[3] In 2010, de Graff analyzed metabolite stability in blood samples at three different storage temperatures.[4] Seven-day storage resulted in concentrations of 6-TGN that were 53% and 90% of original in samples stored at 22 ºC and 4 ºC, respectively and concentrations of 6-MMP that were 55% and 86% of original in samples stored at 22 ºC and 4 ºC, respectively. All differences except the 4 ºC 6-TGN decrease were statistically significant.

Section Summary: Analytic Validity

TPMT genotypic analysis via polymerase chain reaction technology is expected to have high performance. Concordance between genotypic and phenotypic analysis for TPMT activity is high in at least one analysis. Storage conditions are a critical component of metabolite marker testing reproducibility.

CLINICAL VALIDITY

TPMT Genotype and Phenotype Testing

Several systematic reviews of studies on the diagnostic performance of TPMT genotyping have been published. A 2011 review by Booth was sponsored by the Agency for Healthcare Research and Quality.[5] Overall, the evidence was judged to be of limited quality. Nineteen studies on test performance were identified; most were cross-sectional or prospective observational studies and approximately 70% included patients with IBD. Among the 1735 total patients, 184 were heterozygous, and 16 were homozygous for variant alleles, a small subsample of subjects with variant alleles. Pooled analysis of data from 19 studies found a sensitivity of 79.9% (95% confidence interval [CI], 74.8% to 84.6%) for correctly identifying subjects with subnormal (intermediate or low) enzymatic activity. The specificity of the wild-type genotype for correctly identifying subjects with normal or high enzymatic activity approached 100%. Seventeen studies addressed the association between TPMT status and thiopurine toxicity. The studies included 2211 patients, 357 of whom had intermediate and 74 had low enzymatic activity. In a pooled analysis of three studies (92 patients, 10 events), there were greater odds of myelotoxicity with low TPMT enzymatic activity than intermediate activity (pooled odds ratio [OR], 14.5; 95% CI, 2.78 to 76.0). Similarly, in a pooled analysis of three studies (403 patients, 29 events), there were greater odds of myelotoxicity with low TPMT...
enzymatic activity than with normal levels (pooled OR=19.1; 95% CI, 4.6 to 80.2). It is worth noting that the CIs were wide due to few events and small sample sizes.

Another systematic review published in 2011, by Donnan, identified 17 studies that reported the performance characteristics of TPMT genotyping tests (12 studies) and phenotyping (six studies) compared with a reference standard.[6] No true criterion standard was available. The enzymatic test was used as the reference standard in nine studies, and the remainder used a genotyping test; three studies compared two methods of genotyping. All studies used a method of genotyping as either the investigational test or the reference standard; the tests varied somewhat in the number and type of variants they were designed to detect. Sixteen of 17 studies either reported sensitivity and specificity, or reported sufficient data for these measures to be calculated. Only three studies considered confounding factors (e.g., concurrent medications, blood transfusions) in their exclusion criteria. Reviewers did not pool study findings. In the included studies, sensitivity of enzymatic tests ranged from 92% to 100% and the specificity ranged from 86% to 98%. The sensitivity of the genotype tests ranged from 55% to 100% and the specificity from 94% to 100%. In general, the enzymatic tests had a high sensitivity and a low positive predictive value when genotype tests were used as the reference standard. Genotype tests showed a lower sensitivity and a high positive predictive value when enzymatic tests were used as the criterion standard. The inconsistent use of a reference standard complicated interpretation of the findings.

A 2015 meta-analysis by Liu evaluated the relation between TPMT variants and adverse drug reactions (ADRs) in patients with IBD taking thiopurine drugs.[7] This 2015 analysis updated a 2010 meta-analysis by Dong, and findings for both were similar.[8] The Liu review included studies that compared TPMT variant frequencies in patients who did and did not experience ADRs. Reviewers initially screened 353 articles, and 14 studies (total N=2276 IBD patients) were ultimately found to meet eligibility criteria. In a meta-analysis of data from 10 studies, 67 (14.1%) of 476 patients with and 57 (4.8%) of 1192 patients without an ADR were TPMT heterozygous or homozygous. The pooled odds ratio was 3.36 (95% CI, 1.82 to 6.19), and the difference between groups was statistically significant. In analyses of specific adverse reactions, there were statistically significant associations between the presence of TPMT alleles and bone marrow toxicity, but not hepatotoxicity, pancreatitis, or other ADRs (e.g., gastric intolerance, skin reactions). The number of events in some analyses was relatively small, and these studies may have been underpowered to detect differences between groups. For example, 2 (3.3%) of 62 IBD patients with pancreatitis were TPMT heterozygous or homozygous compared with 116 (7.7%) of 1500 patients without pancreatitis (OR=0.97; 95% CI, 0.38 to 2.48).

In 2016, Roy reported on the association between TPMT genotype or phenotype tests and a reference standard. Sensitivity, specificity, positive predictive value, negative predictive value, and concordance were calculated in patients receiving thiopurines.[9] Sixty-six studies were included and appraised for quality. Based on data from 25 studies reporting on test performance on genotyping, the calculated sensitivity for TPMT genotyping to detect a heterozygous or homozygous TPMT variant ranged from 13.4% to 100.0%, while the specificity ranged from 90.9% to 100.0%. A smaller 2016 systematic review by Zur reported sensitivity and specificity of 88.8% and 99.2% for TPMT genotyping when the two most common variants were both included.[10]

No systematic reviews of studies on TPMT genotyping or phenotyping tests in patients undergoing solid organ transplantation were identified. One study identified addressed this
population and provided support for genotype analysis. In 2013, Liang published data on 93 heart transplant patients treated with azathioprine (AZA).\textsuperscript{[11]} Eighty-three patients had the wild-type genotype, and 10 were heterozygous for variants. The TPMT activity level was significantly lower in the heterozygous subjects (13.1 U/mL) than in subjects with the wild-type genotype (21 U/mL red blood cells; \( p < 0.001 \)). Moreover, there was a significantly higher rate of severe rejection in heterozygous subjects (7/10 [70\%]) than in subjects with a wild-type genotype (12/83 [15\%; \( p < 0.001 \)). In addition, heterozygous subjects developed severe rejection earlier than wild-type subjects, at a median of 29 days vs 36 days (\( p = 0.046 \)). There were no statistically significant associations between TPMT genotype and the development of hepatotoxicity or leukopenia.

**NUDT15 Genotyping**

In 2017, Zhang performed a systematic review and meta-analysis on the association of NUDT15 c.415C>T allele and thiopurine-induced leukocytopenia in Asians.\textsuperscript{[12]} Studies published through July 10, 2016 were searched and seven studies including 1138 patients met inclusion criteria. Six of the studies were cohort studies and one was a case control study. Study quality was evaluated with the Newcastle-Ottawa quality assessment scale (NOS) criteria. All included studies were rated greater than or equal to six, indicating good quality. When assessed with a funnel plot and Egger’s test, no evidence of publication bias was found. A random-effects model meta-analysis indicated that the presence of the T allele was significantly associated with high incidences of leukocytopenia. The risk ratio of developing leukopenia was 3.79 for CT + TT versus CC, 3.41 for CT versus CC, and 6.54 for TT versus CC.

Yin (2017) conducted meta-analyses to analyze the relationship between the NUDT15 c.415C>T allele and two outcomes: thiopurine myelotoxicity susceptibility and thiopurine intolerance dose.\textsuperscript{[13]} Data from two cohorts were assessed for the two outcomes. These cohorts included patients with ALL and those with IBD. ALL and IBD patients were separated in the second analysis because thiopurine dosage used in ALL patients was significantly higher than that used in IBD patients. No publication bias was detected in either meta-analysis. The first meta-analysis aimed to determine the relationship between the NUDT15 variant and thiopurine-induced myelotoxicity. Six of the seven studies from the Zhang systematic review above, plus one additional study, were analyzed. The NUDT15 c.415C>T allele was found to contribute 7.86-fold higher risk for developing leukopenia with 91.74% specificity and 43.19% sensitivity. Specificity for early leukopenia was 84.59%.

The second analysis aimed to determine the association between the NUDT15 c.415C>T variant and thiopurine intolerance dose. 2745 patients from 13 cohorts were assessed. There was high heterogeneity between studies. Patients with CT and TT genotypes required 28\% lower mean thiopurine dose compared to patients with CC genotypes, a difference that was statistically significant.

Recent publications not in a systematic review are included below.

A 2018 cohort study by Sutiman examined IBD patients from an Asian population. NUDT15 and TMPT genes were sequenced and intracellular steady-state metabolite concentrations were quantified. Assessment time points were four, eight, and 12 weeks and six months after thiopurine initiation.\textsuperscript{[14]} Greater copy numbers of the T allele at NUDT15 415C>T were significantly associated with declines in nadir white blood cell, absolute neutrophil count, and platelet counts at all time points. NUDT15 activity was inferred from haplotype pairs, and low
and intermediate activity was significantly associated with higher risks of leukopenia and neutropenia than patients with normal NUDT15 activity.

In a 2017 cohort study, Lee examined 165 patients undergoing thiopurine treatment for Crohn’s disease.[15] Clinical evaluation and laboratory examinations were carried out every two to three months for a median of 12 months. Thiopurine metabolite levels and NUDT15 and TPMT genotypes were also evaluated. NUDT15 variants were reported in 20.6% of patients, but the identity of the variants was not reported. The odds ratio of the patients homozygous for NUDT15 variants developing leukopenia was 3.44 (95% CI, 1.21–9.78). There was an average reduction in white blood cell count of 88.2% at four weeks in patients with NUDT15 homozygous variant genotype.

In 2017, Sato enrolled 163 thiopurine-treated IBD patients in a cohort study.[16] The relationship between five variants (NUDT15 p.Arg139Cys, NUDT15 p.Val18_Val19insGlyVal, NUDT15 p.Val18Ile, and FTO p.Ala134Thr, and RUNX1 rs2834826) and gastrointestinal intolerance to azathioprine and azathioprine discontinuation was examined. No association was found with gastrointestinal intolerance. However, NUDT15 p.Arg139Cys was significantly associated with the interval between azathioprine initiation and discontinuation among patients with gastrointestinal intolerance, early and late leukopenia, and severe hair loss.

Chao (2017) performed a cohort study including 732 Chinese IBD patients prescribed thiopurines for at least two weeks.[17] The NUDT15 variants c.415C>T, c.36_37insGGAGTC, and c.52G>A were significantly associated with thiopurine-induced leukopenia. The predictive sensitivity was 49.2% for NUDT15 c.415C>T alone and 55.4% for NUDT15 c.415C>T combined with c.36_37insGGAGTC and c.52G>A. The median dosage for NUDT15 c.415C>T carriers was significantly lower than that for wild-type.

Kim (2017) examined 84 Korean patients with various neurological diseases.[18] The NUDT15 p.R139C variant was significantly associated with leukopenia, early leukopenia, and severe alopecia. The sensitivity and specificity of this variant predicting AZA-induced early leukopenia were 85.7% and 92.2%, respectively.

Zgheib (2017) examined 137 patients with childhood ALL in a Lebanese population. One patient was heterozygous for the NUDT15 c.415C>T allele.[19] This patient tolerated 33.33% of the planned 6-mercaptopurine dose, which was a statistically significant decrease from the median tolerated 6-mercaptopurine dose. Three patients were found to have the TPMT*3A haplotype and tolerated 40.00-66.66% of the planned 6-mercaptopurine dose, which was also a statistically significant reduction from the median tolerated dose.

In 2016, Lee recruited 81 pediatric Crohn’s disease patients who had used azathioprine for more than three months.[20] The dose and duration of azathioprine treatment and information about adverse events, including leukopenia, were collected. TPMT and NUDT15 gene sequencing was performed only for the patients who experienced AZA-induced early leukopenia, eight of whom experienced early leukopenia. Six of the eight patients were found to have the NUDT15 c.415C>T variant and one patient had the TPMT c.719A>G (TPMT*3C) variant. Three of the patients with NUDT15 c.415C>T variants were homozygous for the variant. These patients all experienced alopecia totalis.

In 2016, Moriyama analyzed coding variants, thiopurine intolerance, and NUDT15 nucleotide diphosphate activity in 270 children enrolled in clinical trials for acute lymphoblastic leukemia.[21] Authors identified four NUDT15 variants. Data related to the c.415C>T
polymorphism are included in the Zhang and Ying meta-analyses discussed above. In addition, they identified three more coding variants, c.416G>A, c.52G>A, and c.36_37insGGAGTC, that resulted in 74.4 – 100% loss of nucleotide diphosphate activity. These variants were significantly associated with thiopurine intolerance.

Metabolite Marker Testing

Studies on the diagnostic accuracy of metabolite testing have focused on assessing the association between metabolite levels and disease remission or ADRs. One systematic review was identified; it focused on studies conducted in the pediatric population. In a literature search through January 2013, Konidari (2014) identified 15 studies (total N=1026 children with IBD).[22] There were nine retrospective, six prospective case series, and no randomized controlled trials (RCTs). Reviewers did not pool findings. Among studies that evaluated the association between metabolite markers and clinical remission, five found significantly higher rates of remission with higher levels of 6-thioguanine nucleotides (6-TGN), and six did not find significant differences in 6-TGN levels between responders and nonresponders. Moreover, five studies found significant associations between 6-methyl-mercaptopurine ribonucleotides (6-MMPR) levels and hepatotoxicity, while three did not.

Wong (2017) reported on the result of a post hoc analysis of the TOPIC trial to address the predictive value of 6-MMPR concentrations one week after treatment initiation for development of hepatotoxicity during the first 20 weeks of treatment.[23] They reported that, in more than 80% of patients, hepatotoxicity could be explained by elevated 6-MMPR concentrations and the independent risk factors of age, sex, and body mass index, allowing personalized thiopurine treatment in IBD to prevent early failure. Placing 174 patients on a stable thiopurine dose showed that those exceeding the 6-MMPR threshold of 3615 pmol/8 X10^8 erythrocytes were more likely to have hepatotoxicity (OR=3.8; 95% CI, 1.8 to 8.0).

A 2014 study by Kopylov found that 6-methyl-mercaptopurine (6-MMP)/6-TGN ratios performed better than 6-TGN levels for predicting relapse in pediatric patients with Crohn disease.[24] The study included 237 patients treated with a thiopurine for at least three months. A total of 7.7% were TPMT heterozygous; none was TPMT homozygous. Patients were followed for 18 months; 6-MP metabolite concentration levels were measured every three to four months, or at the time of a clinical relapse or adverse event. The investigators found that 6-MMP/6-TGN ratios between 4 and 24 were significantly protective against relapse; 6-TGN levels alone were not significantly associated with relapse rates.

Several studies have considered the optimal therapeutic cutoff level of metabolites. A 2000 study by Dubinsky (N=92 patients) and a 2012 study by Gilissen (N=100 patients) both found that 235 pmol/8x10^8 was the optimal therapeutic 6-TGN cutoff.[25,26] A 2012 Dhaliwal studied 70 patients with autoimmune hepatitis who were in remission.[27] Levels of 6-TGN were significantly higher in patients who maintained remission compared with those who did not (mean, 237 pmol/8X10^8 vs 177 pmol/8x10^8; p=0.025). According to receiver operating curve analysis, a cutoff of 220 pmol/8 x108 best discriminated between patients who did and did not stay in remission.

Section Summary: Clinical Validity

Systematic reviews of genotype and phenotype testing have shown a pooled sensitivity of about 80% and specificity near 100% for identifying patients with subnormal enzymatic activity.
In addition, studies have found a greater likelihood of ADRs with low TPMT activity. The evidence is limited by relatively small numbers of events and wide CIs.

The association between metabolite markers and adverse drug events was less consistent, although a post hoc analysis of a large RCT showed that metabolite markers could be used to predict the likelihood of hepatotoxicity with thiopurines.

**CLINICAL UTILITY**

The use of pharmacogenomics and thiopurine metabolite testing creates the possibility of tailoring a drug regimen for each patient, with the ultimate goal of attaining disease remission and eliminating steroid therapy. The preferred study design would compare patient management (e.g., drug choice) and health outcomes in patients managed with and without testing.

**TPMT Genotype and Phenotype Testing**

Three RCTs have compared TPMT testing with no testing and empirical weight-based thiopurine dosing. Genotype testing was used in two studies\(^{[28,29]}\) while the remaining RCTs used the phenotype enzymatic activity. In both RCTs using genotype testing, patients with a normal enzyme and genotype started full-dose thiopurine, while those with intermediate enzymatic activity/heterozygous genotype had a 50% dose reduction. Those with low or absent enzyme activity or homozygous genotype were not given thiopurine or were given a reduced dose at 0% to 10% of the initiation dose. The three RCTs are discussed below.

In 2015, Coenen published results of the TOPIC trial, which randomized 761 patients with IBD across 30 centers to receive empirical weight-based thiopurine dosing (n=378) or genotype-guided dosing (n=405).\(^{15}\) The trial did not meet the primary end point of showing a statistically significant reduction in hematologic ADR among the group that received genotype-guided thiopurines dosing compared with empirical weight-based dosing. After 20 weeks, the percentage of patients with hematologic ADRs was 7.4% vs 7.9% in the genotype-based dosing vs empirical weight-based thiopurine dosing, with a relative risk of 0.93 (95% CI, 0.57 to 1.52). However, among TPMT carriers, only 1 (2.6%) of 39 patients developed a hematologic ADR compared with 8 (22.9%) of 35 patients in the control group (relative risk, 0.11; 95% CI, 0.01 to 0.85). While the results of this secondary analysis were statistically significant, the event rate was low with a wide CI indicating imprecise estimates. Further, there was no statistically significant difference in clinical outcome between the groups in an intention-to-treat analysis at 20 weeks after treatment initiation (p=0.18 for Crohn’s Disease Activity Scale score; p=0.14 for ulcerative colitis). In summary, 200 patients would have to be genotyped to avoid one episode of a hematologic ADR (7.4% vs 7.9%; i.e., 0.5% risk difference). The number needed to treat to avoid one episode of a hematologic ADR would be five for at-risk individuals (risk difference in patients with a genetic variant, 20.3; 2.6% vs 22.9%).

In 2011, Newman reported on the results of the TARGET trial, which randomized 333 IBD patients to genotype-guided dosing for empirical weight-based thiopurine dosing.\(^{[29]}\) Data were available for 322 (97%) of 333 patients at four months. The trial did not meet the primary endpoint of showing a statistically significant reduction in the proportion of patients stopping AZA treatment due to any ADR in genotype-guided dosing arm compared with empirical weight-based dosing. The respective proportion of patients in both arms who stopped taking AZA because of an ADR was 29% (47/163) and 28% (44/159; p=0.74), respectively. The trial included few patients with non-wild-type gene variants (seven heterozygous patients in the
Sayani (2005) reported on the results of a small RCT (N=29) in which IBD patients were randomized to the TPMT assay (n=15) or no assay (n=14) prior to AZA dosing.[30] All 14 patients who received TPMT assay were found to have normal TPMT levels and therefore commenced AZA at 2.5 mg/kg/d while the individuals in the control arm underwent an upward dose-titration protocol to a target dose of 2.5 mg/kg/d. While the trial was small and did not report power calculations, results showed that 53% (8/15) and 57% (8/14) in the no assay and TPMT assay groups, respectively, withdrew as a result of AZA-induced adverse events.

Several prospective studies have examined variations in the efficacy of medication by patient TPMT status. For example, in a study that involved 131 patients with IBD, investigators from Europe did not find that the choice of AZA or 6-MP dose based on red blood cells TPMT activity prevented myelotoxicity; no patients in this study exhibited low activity.[31] In a 2008 study from New Zealand, Gardiner noted that initial target doses to attain therapeutic levels in patients with IBD ranged from 1 to 3 mg/kg/d in intermediate (heterozygous) and normal (wild-type) metabolizers.[32] This conclusion was based on analysis of 52 patients with IBD who were started on AZA or 6-MP and followed for nine months while 6-TGN levels and clinical status were evaluated. This study suggests that knowledge of TPMT activity can assist with initial dosing. In a 2006 study from Europe that included 394 patients with IBD, Gisbert found the probability of myelotoxicity was 14.3% in the TPMT intermediate group compared with 3.5% in groups with high (wild-type) activity.[33] Authors concluded that determining TPMT activity before initiating treatment with AZA could minimize the risk of myelotoxicity.

**NUDT15 Genotyping**

Yi (2017) analyzed the outcomes of NUDT15 genotype-based thiopurine dose adjustments in 258 Korean children with acute lymphoblastic leukemia.[34] Variants identified were c.36_37insGGAGTC; c.415C>T, c.415C>T, c.416G>A, c.52G>A, and c.36_37insGGAGTC. Patients were classified as having normal (wild-type; n = 190), intermediate (heterozygous variant; n = 61), or low (homozygous or compound heterozygous variant; n = 7) NUDT15 activity. Patients with TPMT variants were excluded from the analysis. The low-intermediate- and normal-activity groups were administered 7.5, 24.4, and 31.1 mg/m2/day 6-mercaptopurine, respectively. Therapy interruption was 169, 30, and 16 days for the low-intermediate- and normal-activity groups, respectively, with the low-activity group experiencing significantly longer therapy interruption than the other groups, longer duration of leukopenia (low-activity group 131 days, intermediate-activity group 92 days, normal-activity group 59 days, p < 0.01), and lowest blood cell counts.

**Metabolite Marker Testing**

In 2013, Kennedy retrospectively reviewed medical records of patients who had undergone metabolite testing in South Australia.[35] The analysis reported on 151 patients with IBD who had been taking a thiouprine for at least four weeks, underwent at least one metabolite test, and were managed at a study site. The 151 patients had a total of 157 tests. Eighty (51%) of 157 tests were done because of flare or lack of medication efficacy, 18 (12%) were for adverse events, and 54 (34%) tests were routine. Forty-four (55%) of the 80 patients who had a metabolite test due to flare or lack of efficacy had better outcomes after the test was performed. Outcomes also improved after testing for 5 (28%) of 18 patients with an ADR to a
thiopurine. For patients who had routine metabolite tests, 7 (13%) of 54 had better outcomes following testing. The rate of benefit was significantly higher in patients tested because of flare or lack of efficacy compared with those who underwent routine metabolite testing (p<0.001). Changes in patient management included medication dose adjustments, change in medication, and surgical treatment. The study lacked a control group, and thus, outcomes cannot be compared with patients managed without metabolite testing. It is possible that, even in the absence of metabolite testing, patients who were not seeing a benefit or who were experiencing ADRs would have had their treatments adjusted, which could have improved outcomes.

Other relevant studies have examined the association between drug dose and the level of metabolite markers. In general, studies have reported that there is only weak correlation between metabolite levels and drug dose. One 2013 retrospective study, however, found a positive correlation between levels of 6-TGN and 6-MMP and weight-based AZA dose in children with IBD. In addition, studies have reported that levels obtained with testing are often outside of the therapeutic range. For example, Garry (2005) reported that 41% of values were within the therapeutic range and Armstrong (2011) found that 32% of values were within therapeutic levels.

Section Summary: Clinical Utility

Three RCTs (total N=1145 patients) were identified that compared TPMT genotype and phenotype testing with no testing and empirical weight-based thiopurine dosing. In these studies, only 0.17% (n=2) were homozygous. Genotype testing was used in two studies while one used the phenotype enzymatic activity. Of the three RCTs, only the TOPIC trial with a large sample (N=761) was adequately powered while the remaining two were underpowered. Hematologic adverse events and treatment discontinuation were used as surrogate outcomes for benefits of TPMT testing. There were no significant differences in either outcome based on TPMT testing and treatment discontinuation. Additionally, there was also no significant difference in clinical remission in these groups based on TPMT testing in the largest RCT. However, secondary analysis of individual who were intermediate enzymatic activity/heterozygous genotype or homozygous genotype/low enzymatic activity showed that TPMT testing to guide dosing was associated with an 89% risk reduction of hematologic adverse events. In conclusion, although the risk of harm from not testing a TPMT level before initiating therapy is minimal (indicated by a large number needed to treat) in most cases, there is considerable risk of harm (indicated by a small number needed to harm) in the 0.3% patients who are homozygous genotype or have low/absent TPMT enzymatic activity.

The evidence for metabolite marker testing is limited to retrospective studies that report a benefit of metabolite test in terms of better outcomes. However, due to lack of data from RCTs, outcomes cannot be compared with patients managed without metabolite testing.

SUMMARY OF EVIDENCE

For individuals who are treated with thiopurines who receive TPMT genotype or phenotype analysis, the evidence includes studies of diagnostic performance, systematic reviews, and randomized controlled trials. Relevant outcomes are symptoms, morbid events, and change in disease status. A large number of studies have assessed the diagnostic performance of TPMT genotyping and phenotyping tests. A meta-analysis found a pooled sensitivity of about 80% and specificity near 100% for identifying patients with subnormal enzymatic activity. Three randomized controlled trials (total N=1145 patients) compared TPMT genotype/phenotype
testing with no testing and empirical weight-based thiopurine dosing. There was no significant
difference in the incidence of hematologic adverse events, treatment discontinuation rates, or
clinical remission. However, secondary analysis of a small number of individuals who had
intermediate enzymatic activity/heterozygous genotype or homozygous genotype/low
enzymatic activity showed that TPMT testing to guide dosing was associated with statistically
significant risk reduction in hematologic adverse events with a wide margin of error. In
summary, 200 patients would have to be genotyped to avoid one episode of a hematologic
adverse drug reaction (7.4% vs 7.9%; i.e., 0.5% risk difference). The number needed to treat
to avoid one episode of a hematologic adverse drug reaction would be five for at-risk
individuals (risk difference in patients with a genetic variant, 20.3; 2.6% vs 22.9%). In addition,
a small, inadequately powered randomized controlled trial that assessed phenotype TPMT
testing found no difference in treatment discontinuation rates due to adverse drug reactions
between the two arms. The evidence is sufficient to determine that the technology results in a
meaningful improvement in the net health outcome.

For individuals who are treated with thiopurines who receive azathioprine and/or 6-
mercaptopurine metabolites analysis, the evidence includes a systematic review as well as
prospective and retrospective studies. Relevant outcomes are symptoms, morbid events, and
change in disease status. There is insufficient evidence from prospective studies to determine
whether knowledge of metabolite marker status will lead to improved outcomes (primarily
improved disease control and/or less adverse drug events). Findings for studies evaluating the
association between metabolite markers and clinical remission are mixed, and no prospective
comparative trials have compared health outcomes in patients managed using metabolite
markers with current approaches to care. The evidence is insufficient to determine the effects
of the technology on health outcomes.

PRACTICE GUIDELINE SUMMARY

NATIONAL COMPREHENSIVE CANCER NETWORK

National Comprehensive Cancer Network (v.5.2017) guidelines on acute lymphoblastic
leukemia state that testing for thiopurine methyltransferase (TPMT) gene variants should be
considered for patients receiving mercaptopurine (6-MP)—in particular, patients who develop
severe neutropenia after starting 6-MP.[40]

NORTH AMERICAN SOCIETY FOR PEDIATRIC GASTROENTEROLOGY, HEPATOLOGY
AND NUTRITION

In 2013, the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition
on inflammatory bowel disease (IBD) published consensus recommendations on the role of the
TPMT enzyme and thiopurine metabolite testing in pediatric IBD.[41] Recommendations (high
and moderate) included:

1. “TPMT testing is recommended before initiation of TPs [thiopurines] to identify
   individuals who are homozygous recessive or have extremely low TPMT activity…
2. Individuals who are homozygous recessive or have extremely low TPMT activity should
   avoid use of TPs because of concerns for significant leucopenia.
3. … All individuals on TPs should have routine monitoring of CBC [complete blood cell]
   and WBC [white blood cell] counts to evaluate for leucopenia regardless of TPMT
   testing results.
4. Metabolite testing can be used to determine adherence to TP therapy.
5. Metabolite testing can be used to guide dosing increases or modifications in patients with active disease.…

6. Routine and repeat metabolite testing has little or no role in patients who are doing well and taking an acceptable dose of a TP.”

**BRITISH ASSOCIATION OF DERMATOLOGISTS**

The 2011 guidelines from the British Association of Dermatologists addressed the safe and effective prescribing of azathioprine for the management of autoimmune and inflammatory skin diseases.[42] The guidelines included the following recommendations on analysis of TPMT activity and azathioprine toxicity:

- “There is strong evidence that baseline testing predicts severe neutropenia in patients with absent TPMT activity.
- There is good evidence that intermediate TPMT activity is associated with myelotoxicity in patients using conventional azathioprine doses.
- TPMT testing only identifies … haematological toxicity, hence the continued need for regular monitoring of blood counts irrespective of TPMT status.”

**AMERICAN GASTROENTEROLOGICAL ASSOCIATION INSTITUTE**

Recommendations from a 2017 American Gastroenterological Association Institute guidelines on therapeutic drug monitoring in IBD are summarized in Table 1.[43]

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>SOR</th>
<th>QOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>In adults with IBD being started on thiopurines, AGA suggests routine <strong>TPMT</strong> testing (enzymatic activity or genotype) to guide thiopurine dosing</td>
<td>Conditional</td>
<td>Low</td>
</tr>
<tr>
<td>In adults treated with thiopurines with active IBD or adverse effects thought to be due to thiopurine toxicity, AGA suggests reactive thiopurine metabolite monitoring to guide treatment changes</td>
<td>Conditional</td>
<td>Very low</td>
</tr>
<tr>
<td>In adults with quiescent IBD treated with thiopurines, AGA suggests against routine thiopurine metabolite monitoring</td>
<td>Conditional</td>
<td>Very low</td>
</tr>
</tbody>
</table>


**SUMMARY**

There is enough research to show that TPMT genotype or phenotype analysis prior to thiopurine therapy improves health outcomes. In addition, there is enough research to show that TPMT genotype or phenotype analysis in patients on thiopurine therapy with abnormal complete blood count results that do not respond to dose reduction improves health outcomes. These improved outcomes include reduced adverse events due to drug toxicity. Clinical guidelines based on research recommend genotype or phenotype analysis prior to thiopurine therapy and in patients on thiopurine therapy with abnormal complete blood count results that do not respond to dose reduction. Therefore, TPMT genotype or phenotype analysis may be considered medically necessary prior to thiopurine therapy and in patients on thiopurine therapy when there is clinical documentation of abnormal complete blood count results that do not respond to dose reduction when policy criteria are met.
There is not enough research to show that TPMT genotype or phenotype analysis improves health outcomes in people who do not meet the criteria. Therefore, genotypic and/or phenotypic analysis of the TPMT enzyme is considered investigational in all other situations.

There is not enough research to show that analysis of the metabolite markers azathioprine and mercaptopurine improves health outcomes. Therefore, analysis of the metabolite markers azathioprine and mercaptopurine, including 6-methyl-mercaptopurine ribonucleotides and 6-thioguanine nucleotides, is considered investigational.

There is enough research to show that genetic testing of NUDT15 prior to thiopurine therapy improves health outcomes. In addition, there is enough research to show that NUDT15 in patients on thiopurine therapy with abnormal complete blood count results that do not respond to dose reduction improves health outcomes. These improved outcomes include reduced adverse events due to drug toxicity. Therefore, genetic testing of NUDT15 may be considered medically necessary prior to thiopurine therapy and in patients on thiopurine therapy when there is clinical documentation of abnormal complete blood count results that do not respond to dose reduction when policy criteria are met.

There is not enough research to show that genetic testing of NUDT15 improves health outcomes in people who do not meet the criteria. Therefore, genetic testing of NUDT15 is considered investigational in all other situations.

REFERENCES


44. BlueCross BlueShield Association Medical Policy Reference Manual "Pharmacogenomic and Metabolite Markers for Patients Treated With Thiopurines." Policy No. 2.04.19

**CODES**

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>81335</td>
<td>TPMT (thiopurine S-methyltransferase) (eg, drug metabolism), gene analysis, common variants (eg, *2, *3)</td>
</tr>
<tr>
<td></td>
<td>81401</td>
<td>Molecular Pathology Procedure Level 2</td>
</tr>
<tr>
<td>HCPCS</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

**Date of Origin:** January 2018

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Regence

Medical Policy Manual

Confocal Laser Endomicroscopy

Effective: September 1, 2018

Next Review: July 2019
Last Review: July 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Confocal laser endomicroscopy (CLE), also known as confocal fluorescent endomicroscopy and optical endomicroscopy, allows in vivo microscopic imaging of cells during endoscopy. CLE is proposed for a variety of purposes, especially as a real-time alternative to histology during colonoscopy and for targeting areas to undergo biopsy in patients with inflammatory bowel disease and Barrett esophagus.

MEDICAL POLICY CRITERIA

Use of confocal laser endomicroscopy is considered investigational for all indications.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES

1. Analysis of Human DNA in Stool Samples as a Technique for Colorectal Cancer Screening, Genetic Testing, Policy No. 12
2. In Vivo Analysis of Colorectal Polyps, Medicine, Policy No. 104
3. Electromagnetic Navigation Bronchoscopy, Surgery, Policy No. 179
BACKGROUND

CLE involves using light from a low-power laser to illuminate tissue and, subsequently, the same lens detects light reflected from the tissue through a pinhole. The term confocal refers to having both illumination and collection systems in the same focal plane. Light reflected and scattered at other geometric angles that is not reflected through the pinhole is excluded from detection, which dramatically increases the special resolution of CLE images.

Endoscope-based and probe-based systems have been cleared by the U.S. Food and Drug Administration (FDA). Endoscope-based systems incorporate a confocal probe onto the tip of a conventional endoscope. Image collection scan rates vary by device. Probe-based systems place a probe through the biopsy channel of a conventional endoscope. Depth of imaging and field of view varies by device. As pointed out in review articles, the limited viewing area emphasizes the need for careful conventional endoscopy to target the areas for evaluation. Both CLE systems are optimized using a contrast agent. The most widely used agent is intravenous fluorescein, which is FDA-approved for ophthalmologic imaging of blood vessels when used with a laser scanning ophthalmoscope.

Unlike techniques such as chromoendoscopy, which are primarily intended to improve the sensitivity of colonoscopy, CLE is unique in that it is designed to immediately characterize the cellular structure of lesions. CLE can thus potentially be used to make a diagnosis of polyp histology, particularly in association with screening or surveillance colonoscopy, which could allow for small hyperplastic lesions to be left in place rather than removed and sent for histologic evaluation. This would reduce risks associated with biopsy and reduce the number of biopsies and histologic evaluations. Another key potential application of CLE technology is targeting areas for biopsy in patients with Barrett esophagus undergoing surveillance endoscopy. This is an alternative to conducting random biopsies during surveillance and has the potential to reduce the number of biopsies and/or improve the detection of dysplasia. Other potential uses of CLE under investigation include better diagnosis and differentiation of conditions such as gastric metaplasia, lung cancer and bladder cancer.

As noted previously, limitations of CLE systems include a limited viewing area and depth of view. An additional limitation is the lack of standardized systems for classifying lesions viewed with CLE devices. Although there is not currently an internationally accepted classification system for colorectal lesions, two systems have been developed that have been used in a number of studies conducted in different countries. These are the Mainz criteria for endoscopy-based CLE devices and the Miami classification system for probe-based CLE devices.[1] Lesion classification systems are less developed for non-gastrointestinal lesions viewed by CLE devices, e.g., those in the lung or bladder. Another potential limitation of CLE is the learning curve for obtaining high-quality images and classifying lesions. Although several recent studies have found that the ability to acquire high-quality images and interpret them accurately can be learned relatively quickly, these studies were limited to colorectal applications of CLE.[2,3]

Regulatory Status

Several CLE devices have been cleared for marketing by the FDA. These include:

Cellvizio® (Mauna Kea Technologies): This device consists of a confocal laser system, proprietary software, a flat-panel display and miniaturized fiber optic probes. Since 2006, Mauna Kea has received ten FDA approvals for Cellvizio® systems, most recently in May 2016.

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
EC-3870CLIK Confocal Video Colonoscope (Pentax Medical Company): This is an endoscopy-based CLE system which consists of the EC-3870CLIK, Confocal Video Colonoscope (K042741) and the ISC-1000 Pentax Confocal Laser System (K042740). The device must be used with a Pentax Video Processor. According to FDA materials, the intended use of the device is to provide optical and microscopic visualization of and therapeutic access to the lower gastrointestinal tract.

On June 28, 2016, the FDA issued a Class 2 device recall for the EC-3870CLIK device.[4]

EVIDENCE SUMMARY

COLORECTAL LESIONS

Ideally, the evaluation of the safety and efficacy of confocal laser endomicroscopy (CLE) as a diagnostic tool would be based on randomized controlled trials (RCTs) comparing CLE to conventional diagnostic methods, such as biopsy with histology for analysis of colorectal lesions. The evidence for the use of CLE is best evaluated in the framework of a diagnostic test, as the test provides diagnostic information that assists in treatment decisions. Validation of the clinical use of any diagnostic test focuses on three main principles:

1. The analytic validity of the test, which refers to the technical accuracy of the test in detecting abnormal histology that is present or in excluding an abnormality that is absent;
2. The clinical validity of the test, which refers to the diagnostic performance of the test (sensitivity, specificity, positive and negative predictive values) in detecting clinical disease; and
3. The clinical utility of the test, i.e., how the results of the diagnostic test will be used to change management of the patient and whether these changes in management lead to clinically important improvements in health outcomes.

Multiple studies have evaluated the diagnostic accuracy of CLE for patients undergoing screening or surveillance colonoscopy. Several systematic reviews of studies evaluating the diagnostic accuracy of CLE compared to a reference standard have been published. Descriptions of several systematic reviews and representative diagnostic accuracy studies are included below.

Systematic Reviews

A 2018 systematic review by Lord analyzed the diagnostic accuracy of several optical imaging techniques for in vivo lesion characterization in colonic inflammatory bowel disease (IBD).[5] A total of 22 studies were identified assessing performance of virtual chromoendoscopy, dye-based chromoendoscopy, magnification endoscopy, and confocal laser endomicroscopy. A bivariate meta-analysis was performed. Pooled sensitivities of real-time CLE, magnification endoscopy, virtual chromoendoscopy, and dye-based chromoendoscopy were 91% (95%CI: 66%-98%), 90% (95%CI: 77%-96%), 86% (95%CI: 62%-95%), and 67% (95%CI: 44%-84%), respectively. Pooled specificities were 97% (95%CI: 94%-98%), 87% (95%CI: 81%-91%), 87% (95% CI: 72%-95%), 86% (95%CI: 72%-94%), for the same methods, respectively. The authors concluded that real-time CLE is highly accurate for differentiating neoplastic from non-
neoplastic lesions in patients with colonic IBD, but also note that most CLE studies were performed by single expert users within tertiary centers, which may confound results.

In 2013, Su reviewed studies on the efficacy of CLE for discriminating colorectal neoplasms from non-neoplasms.[6] Studies needed to use histologic biopsy as the reference standard and in which the pathologist and endoscopist were blinded to each other’s findings. Included studies also used a standardized CLE classification system. Patient populations included individuals at increased risk of colorectal cancer due to personal or family history, patients with previously identified polyps, and/or patients with IBD. Two reviewers independently assessed the quality of individual studies using the modified Quality Assessment Of Diagnostic Accuracy Studies (QUADAS) tool, and studies considered to be at high risk of bias were excluded from further consideration. A total of 15 studies with 719 adult patients were found to be eligible for the systematic review. All were single-center trials and two were available only as abstracts. In all the studies, suspicious lesions were first identified by conventional white-light endoscopy with or without chromoendoscopy and then further examined by CLE. A pooled analysis of the 15 studies found an overall sensitivity of CLE of 94% (95% confidence interval [CI], 0.88 to 0.97) and specificity of 95% (95% CI, 0.89 to 0.97), compared to histology. Six of the studies included patients at increased risk of colorectal cancer (CRC) who were undergoing surveillance endoscopy, five studies included patients with colorectal polyps and four studies included patients with IBD. In a predefined subgroup analysis by indication for screening, the pooled sensitivity and specificity for surveillance studies was 94% (95% CI, 90% to 97%) and 98% (95% CI, 97% to 99%), respectively. For patients presenting with colorectal polyps, the pooled sensitivity of CLE was 91% (95% CI, 87% to 94%) and specificity was 85% (95% CI, 78% to 90%). For patients with IBD, the pooled sensitivity was 83% (95% CI, 70% to 92%) and specificity was 90% (95% CI, 87% to 93%). In other predefined subgroup analyses, the summary sensitivity and specificity was significantly higher (p<0.001) in studies of endoscopy-based CLE (97% and 99%, respectively) than studies of probe-based CLE (87% and 82%, respectively). In addition, the summary sensitivity and specificity was significantly higher (p<0.01) with real-time CLE in which the macroscopic endoscopy findings were known (96% and 97%, respectively) than with blinded CLE in which recorded confocal images were subsequently analyzed without knowledge of macroscopic endoscopy findings (85% and 82%, respectively).

Another systematic review was published in 2013 by Dong.[7] The investigators included studies that assessed the diagnostic accuracy of CLE compared with conventional endoscopy. They did not explicitly state that the reference standard was histologic biopsy, but this was the implied reference standard. A total of six studies were included in a meta-analysis. All of the studies were prospective, and at least five included blinded interpretation of CLE findings (in one study, it was unknown whether interpretation was blinded). In a pooled analysis of data from all six studies, the sensitivity was 81% (95% CI, 77% to 85%) and the specificity was 88% (95% CI, 85% to 90%). The authors also conducted a subgroup analysis by type of CLE used. When findings from the two studies on endoscopy-based CLE were pooled, the sensitivity was 82% (95% CI, 69% to 91%) and the specificity was 94% (95% CI, 91% to 96%). Two studies may not have been a sufficient number to obtain a reliable estimate of diagnostic accuracy. When findings from the 4 studies on probe-based endoscopy were pooled, the sensitivity was 81% (95% CI, 76% to 85%) and the specificity was 75% (95% CI, 69% to 81%).

A 2013 systematic review by Wanders searched for studies that reported diagnostic accuracy of studies on any of several new technologies used to differentiate between colorectal neoplasms and non-neoplasms.[8] To be included in the review, studies needed to use the
technology to differentiate between non-neoplastic and neoplastic lesions and to use histopathology as the reference standard. Blinding was not an inclusion criterion. Eleven eligible studies were identified that included an analysis of CLE. A pooled analysis of study findings yielded an estimated sensitivity of 93.3% (95% CI, 88.4 to 96.2) and a specificity of 89.9% (95% CI, 81.8% to 94.6%). A meta-analysis limited to the five studies that used endoscopy-based CLE found a sensitivity of 94.8% (95% CI, 90.6% to 98.92%) and a specificity of 94.4% (95% CI, 90.7% to 99.2%). When findings of the six studies on probe-based CLE were pooled, the sensitivity was 91.5% (86.0% to 97.0%) and the specificity was 80.9% (95% CI, 69.4% to 92.4%).

Nonrandomized Studies

Ohmiya (2017) evaluated the ability of CLE to differentiate among ulcerative colitis (UC)-associated neoplasia (differentiated type or undifferentiated type), sporadic adenoma, and circumscribed regenerative lesions.[9] The authors examined 12 patients with suspected UC-associated neoplasia with probe-based CLE and compared findings with pathological diagnoses determined by magnifying chromoendoscopy with crystal violet and narrow band imaging. Sensitivity, specificity, and accuracy of CLE were 100%, 83%, and 92%, respectively. The authors stated that CLE was helpful in evaluating suspected UC-associated neoplasia, but it is limited by the small sample size.

In 2017, Kim evaluated probe-based CLE for feasibility and safety in evaluating colorectal submucosa following removal of colorectal neoplasms.[10] Colorectal submucosa were classified as negative or indicative of carcinoma infiltration. The results were compared to pathological findings. The sensitivity, specificity, and accuracy of the classifications were 91.7, 86.8, and 88.0%, respectively. The authors concluded that CLE is useful but that large-scale prospective studies are needed.

In a 2012 study by Shadid two methods of analyzing CLE images, real-time diagnosis and blinded review of video images after endoscopy (known as “offline” diagnosis), were compared.[11] The study included 74 patients with a total of 154 colorectal lesions. Eligibility criteria were similar to the Buchner study (see above); the included patients undergoing surveillance or screening colonoscopy. Patients underwent white-light colonoscopy and identified polyps were also evaluated with virtual chromoendoscopy and probe-based CLE. Intravenous fluorescein sodium was administered after the first polyp was identified. At the time of examination, an endoscopist made a real-time diagnosis based on CLE images. Based on that diagnosis, the patient underwent polypectomy, biopsy or endoscopic mucosal resection, and histopathologic analysis was done on the specimens. The CLE images were then de-identified and then reviewed offline by the same endoscopist at least one month later. At the second review, the endoscopist was blinded to the endoscopic and histopathologic diagnosis. Of the 154 polyps, 74 were found by histopathologic analysis to be non-neoplastic and 80 were neoplastic (63 tubular adenomas, 12 tubulovillous adenomas, three mixed hyperplastic-adenoma polyps and two adenocarcinomas). Overall, there was not a statistically significant difference in the diagnostic accuracy of real-time CLE diagnosis and blinded offline CLE diagnosis (i.e., confidence intervals overlapped). The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for real-time CLE diagnosis was 81%, 76%, 87%, and 79%, respectively. For offline diagnosis, these numbers were 88%, 77%, 81% and 85%, respectively. However, in the subgroup of 107 smaller polyps, less than 10 mm in size, the accuracy of real-time CLE was significantly lower than offline CLE. For the smaller polyps, sensitivity, specificity, PPV and NPV of real-time CLE was 71%, 83%, 78%, and 78%.
and for offline CLE was 86%, 78%, 76%, and 87%, all respectively. For larger polyps, in contrast, there was a nonsignificant trend in favor of better diagnostic accuracy with real-time compared to offline CLE.

A 2011 study by Hlavaty included patients with ulcerative colitis or Crohn disease. Thirty patients were examined with standard white-light colonoscopy, chromoendoscopy and an endoscopy-based CLE system. An additional 15 patients were examined only with standard colonoscopy. All lesions identified by white-light colonoscopy or chromoendoscopy were examined using CLE to identify neoplasia using the Mainz classification system. Suspicious lesions underwent biopsy and, additionally, random biopsies were taken from four quadrants every 10 cm per the standard surveillance colonoscopy protocol. All specimens underwent histologic analysis by a gastrointestinal pathologist who was blinded to the CLE diagnosis. Diagnostic accuracy of CLE was calculated for examinable lesions only. Compared to histologic diagnosis, the sensitivity of CLE for diagnosing low-grade and high-grade intraepithelial neoplasia was 100%, the specificity was 98.4%, the PPV was 66.7%, and the NPV was 100%. However, whereas CLE was able to examine 28 of 30 (93%) flat lesions, it could examine only 40 of 70 (57%) protruding polyps. Moreover, 6 of 10 (60%) dysplastic lesions, including three of five low-grade and high-grade intraepithelial neoplasms were not evaluable by CLE. It is also worth noting that the diagnostic accuracy of chromoendoscopy was similar to that of CLE. The sensitivity, specificity, PPV and NPV of chromoendoscopy was 100%, 97.9%, 75%, and 100%, respectively.

A 2011 study by Xie included 116 consecutive patients who had polyps found during CLE; one patient was excluded from the analysis. All patients had an indication for colonoscopy (19 were undergoing surveillance postpolypectomy, two had a family history of colorectal cancer, three had IBD and 91 were seeking a diagnosis). All patients first underwent white-light colonoscopy. Endoscopy-based CLE was used on the first polyp identified during withdrawal of the endoscope (i.e., one polyp per patient was analyzed). Intravenous fluorescein sodium was used. Real-time diagnosis of the polyp was performed based on criteria used at the study center (which is adapted from the Mainz classification system). The polyps were biopsied or were removed and histopathologic diagnosis was determined. Real-time CLE diagnosis correctly identified 109 of 115 (95%) adenomas or hyperplastic polyps. Four adenomas were misdiagnosed by CLE as hyperplastic polyps (two were tubulous adenomas and two were tubulovillous adenomas) and two hyperplastic polyps were misdiagnosed as adenomas. The overall sensitivity, specificity, PPV, and NPV of CLE diagnosis was 93.9% (95% CI, 85.4% to 97.6%), 95.9% (95% CI, 86.2% to 98.9%), 96.9% (95% CI, 89% to 99%), and 94.8% (95% CI, 89.1% to 97.6%), respectively. For polyps less than 10 mm, the CLE diagnosis had a sensitivity of 90.3% and specificity of 95.7%, and for polyps 10 mm and larger, sensitivity was 97.1% and specificity was 100%.[13]

In 2010, Buchner published findings on 75 patients who had a total of 119 polyps. Patients were eligible for study participation if they were undergoing surveillance or screening colonoscopy or undergoing evaluation of known or suspected polyps identified by other imaging modalities or endoscopic resection of larger flat colorectal neoplasia. White-light colonoscopy was used as the primary screening method. When a suspicious lesion was identified, it was evaluated by virtual chromoendoscopy system and a probe-based CLE system. Intravenous fluorescein sodium was administered after the first polyp was identified. Following the imaging techniques, the appropriate intervention, i.e., polypectomy, biopsy, or endoscopic mucosal resection, of lesions were performed and all resected specimens underwent histopathologic analysis by a pathologist blinded to CLE information. Confocal
images of the 199 polyps were evaluated after all procedures were completed; the evaluator was blinded to histology diagnosis and endoscopic appearance of the lesion. Diagnosis of confocal images used modified Mainz criteria; polyps were classified as benign or neoplastic. According to histopathologic analysis, there were 38 hyperplastic polyps and 81 neoplastic lesions (58 tubular adenomas, 15 tubulovillous adenomas and 4 adenocarcinomas). CLE correctly identified 74 of 81 neoplastic polyps (sensitivity, 91%; 95% CI, 83% to 96%). In addition, CLE correctly identified 29 of 38 hyperplastic polyps (specificity, 76%; 95% CI, 60% to 89%). In contrast, virtual chromoendoscopy correctly identified 62 neoplastic polyps (sensitivity, 77%; 95% CI, 66% to 85%) and 27 hyperplastic polyps (specificity, 71%; 95% CI, 54% to 85%).

Section Summary

Multiple studies have evaluated the accuracy of confocal laser endoscopy compared with histopathology for diagnosing colorectal lesions. In three published systematic reviews, pooled estimates of overall sensitivity of CLE ranged from 81% to 94% and pooled estimates of specificity ranged from 88% to 95%. Although the reported diagnostic accuracy tended to be relatively high, it is not clear whether the accuracy is high enough to replace biopsy/polypectomy and histologic analysis.

BARRETT ESOPHAGUS

The ideal study would determine whether CLE with targeted biopsy can distinguish Barrett’s Esophagus (BE) without dysplasia from BE with low- and high-grade dysplasia. In addition, study results would need to determine if CLE with target biopsy led to fewer biopsies of benign tissue compared to surveillance with random biopsies. The ideal study to address the above questions would include an unselected clinical population of patients with BE presenting for surveillance and would randomly assign patients to CLE with targeted biopsy or a standard biopsy protocol without CLE. Relevant outcomes include diagnostic accuracy for detecting dysplasia, the detection rate for dysplasia, and the number of biopsies. Several studies with most or all of these elements of study design were identified, including randomized controlled trials (RCTs).

Systematic Reviews

In 2017, Xiong published a systematic review and meta-analysis to assess the accuracy of within-patient comparisons of narrow band imaging and CLE for the diagnosis of high-grade dysplasia and esophageal adenocarcinoma in BE patients. The quality of studies was assessed using the QUADAS-2 tool. A total of five studies with 251 patients were included in the meta-analysis. The pooled sensitivities were not significantly different, with values of 62.8% (95% CI: 0.56-0.69, I²=94.6%) for narrow band imaging and 72.3% (95% CI: 0.66-0.78, I²=89.3%) for CLE. Pooled specificities were also not significantly different (narrow band imaging 85.3% [95% CI: 0.84-0.87, I²=92.1%] vs CLE 83.8% [95% CI: 0.82-0.85, I²=96.8%]). The pooled additional detection rate of CLE compared to narrow band imaging for per-lesion detection of neoplasia was 19.3% (95% CI: 0.05-0.33, I²=74.6%).

In 2016, Xiong published a meta-analysis of prospective studies evaluating the diagnostic accuracy of CLE in patients with BE and using histopathologic analysis as the criterion standard. Studies were not required to compare CLE to standard four-quadrant biopsy. Fourteen studies were included. Three were reported to have a high risk of bias and the rest a low risk of bias. There was no statistically significant publication bias. In a pooled analysis of

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
seven studies (n=473 patients) reporting a per-patient analysis, the sensitivity of CLE for
detecting neoplasia was 89% (95% CI, 82% to 94%) and the specificity was 83% (95% CI, 
78% to 86%). The pooled positive and negative likelihood ratios were 6.53 (95% CI, 3.12 to 
13.4) and 0.17 (95% CI, 0.11 to 0.29, respectively). Reviewers did not report PPV or NPV. 
Sensitivity and specificity were similar to those reported below in the 2014 meta-analysis by Gupta. Limitations to this analysis include heterogeneity of the results and a lack of 
relationship between the diagnostic odds ratio and the characteristics of the studies.

Gupta (2014) conducted a systematic review and meta-analysis to evaluate the diagnostic 
accuracy of the CLE-based targeted biopsies in detecting high grade dysplasia 
(HGD)/adenocarcinoma compared with four-quadrant random biopsies.[17] All the studies that 
compared the diagnostic yield from CLE-based targeted biopsies to detect 
HGD/adenocarcinoma with a gold standard of histopathology were included and a meta-
analysis was carried out to estimate the pooled sensitivity, specificity, and positive and 
negative likelihood. Seven studies with 345 patients and 3080 lesions were included in the 
meta-analysis. All the studies had reported per-lesion analyses; however, only four of the 
seven studies had data reported on per-patient analyses. ‘Per-lesion’ analysis for the diagnosis of HGD/adenocarcinoma yielded a pooled sensitivity and specificity of 68% (95% CI of 64-
73%) and 88% (95% CI of 87-89%), respectively. The pooled positive and negative likelihood ratios were 6.56 (95% CI of 3.61-11.90) and 0.24 (95% CI of 0.09-0.63), respectively. Similar 
numbers were calculated on the basis of ‘per-patient’ basis, which showed a pooled sensitivity 
and specificity of 86% (95% CI of 74-96%) and 83% (95% CI of 77-88%), respectively. The 
pooled positive and negative likelihood ratios were 5.61 (95% CI of 2.00-15.69) and 0.21 (95% 
CI of 0.08-0.59), respectively. Authors noted that CLE, by providing targeted biopsies, has a 
good diagnostic accuracy in identifying HGD/EAC; however, the overall prevalence of 
HGD/EAC in the studies included was much higher than what would be seen in clinical 
practice and these results should be interpreted with caution. Due to its relatively low 
sensitivity and negative predictive value, CLE may currently not replace standard biopsy 
techniques for the diagnosis of HGD/EAC in Barrett’s esophagus.

In 2013, a meta-analysis by Wu of observational studies and RCTs focused on the diagnostic 
accuracy of CLE for detecting neoplasia in BE patients.[18] In a pooled analysis of data from 
four studies that reported per-patient accuracy of CLE, the pooled sensitivity for detection of 
neoplasia was 89% (95% CI, 0.80% to 0.95%), and the pooled specificity was 75% (95% CI, 
69% to 81%). Seven studies reported per-location accuracy of CLE. The pooled sensitivity for 
CLE was 70% (95% CI, 65% to 74%) and the pooled specificity was 91% (95% CI, 90% to 
92%). This study did not address other outcomes such as number of biopsies and did not 
compare CLE for detection of neoplasia in patients with BE with white-light endoscopy.

Randomized Controlled Trials

In 2013, Canto published findings from a single-blind multicenter RCT conducted at academic 
centers with experienced endoscopists.[19] The trial included consecutive patients undergoing 
endoscopy for routine surveillance of BE or for suspected or known neoplasia. Patients were 
randomized to high-definition white-light endoscopy with random biopsy (n=98) or white-light 
endoscopy with endoscopy-based CLE and targeted biopsy (n=94). In the white-light 
endoscopy-only group, four-quadrant random biopsies were taken every one to two cm of the 
entire length of the BE for patients undergoing surveillance and every one cm in patients with 
suspected neoplasia. In the CLE group, biopsy specimens were obtained only when there was 
CLE evidence of neoplasia. The final pathology diagnosis was the reference standard. A per-
patient analysis of diagnostic accuracy for diagnosing BE-related neoplasia found a sensitivity of 40% with white-light endoscopy alone and 95% with white-light endoscopy plus CLE. Specificity was 98% with white-light endoscopy alone and 92% with white-light endoscopy plus CLE. When the analysis was done on a per-biopsy specimen basis, when CLE was added, the sensitivity was substantially higher and the specificity was slightly lower. The median number of biopsies per patient was significantly higher in the white-light endoscopy group compared with the group that also received CLE (4 vs 2, p<0.001). The investigators conducted an analysis of the number of cases in which CLE resulted in a different diagnosis. Thirty-two of 94 (34%) patients in the white-light plus CLE group had a correct change in dysplasia grade after CLE compared to the initial endoscopic findings. Six of the 32 (19%) patients had lesions and the remaining 26 did not. In 21 of the 26 patients without lesions, CLE changed the plan from biopsy to no biopsy. The remaining 62 of 94 (65%) patients in the white-light endoscopy plus CLE group had concordant diagnoses with the two techniques. The study was conducted at academic centers and used endoscopy-based CLE. Findings may not be generalizable to other clinical settings or to probe-based CLE.

In 2011, Sharma published an international, multicenter RCT that included 122 consecutive patients presenting for surveillance of BE or endoscopic treatment of high-grade dysplasia or early carcinoma. This study was described in the systematic review and meta-analysis described by Gupta in the previous section. Patients were randomly assigned to receive, in random order, both standard white-light endoscopy and narrow-band imaging. Following these two examinations, which were done in a blinded fashion, the location of lesions was unblinded and, subsequently, all patients underwent probe-based CLE. All examinations involved presumptive diagnosis of suspicious lesions. Also, in both groups, after all evaluations were performed, there were biopsies of all suspicious lesions, as well as biopsies of random locations (four quadrants every two cm). Histopathologic analysis was the reference standard. Twenty-one patients were excluded from the analysis. Of the remaining 101 patients, 66 (65%) were found on histopathologic analysis to have no dysplasia, four (4%) had low-grade dysplasia, six (6%) had high-grade dysplasia and 25 (25%) had early carcinoma. The sensitivity of CLE with white-light endoscopy for detecting high-grade dysplasia or early carcinoma was 68.3% (95% CI, 60.0% to 76.7%), which was significantly higher than white-light endoscopy alone; 34.2% (95% CI, 25.7% to 42.7%, p=0.002). However, the specificity of CLE and white-light endoscopy was significantly lower than white-light endoscopy alone: 92.7% (95% CI, 90.8% to 94.6%) versus 87.8% (95% CI, 85.5% to 90.1%; p<0.001). For white-light endoscopy alone, the PPV was 42.7% (32.8% to 52.6%) and the NPV was 89.8% (95% CI, 87.7% to 92.0%). For white-light endoscopy with probe-based CLE, the PPV was 47.1% (95% CI, 39.7% to 54.5%) and the NPV was 94.6% (95% CI, 92.9% to 96.2%). White-light endoscopy alone missed 79 of 120 (66%) areas with high-grade dysplasia or early carcinoma and white-light endoscopy with CLE missed 38 (32%) areas. On a per-patient basis, 31 patients were diagnosed with high-grade dysplasia or early carcinoma. White-light endoscopy alone failed to identify four of these patients (sensitivity, 87%), whereas white-light endoscopy and CLE failed to identify two patients (sensitivity, 93.5%).

Another RCT was published in 2012 by Bertani in Italy; this was a single-center study. The study compared the dysplasia detection rate of biopsies obtained by standard white-light endoscopy only to the detection rate with standard endoscopy followed by probe-based CLE in patients with BE who were enrolled in a surveillance program. One hundred consecutive patients were included, and 50 were randomly assigned to each group. In both groups, targeted biopsies of suspicious lesions and random four-quadrant biopsies (one biopsy every one cm) were taken. The authors described the criteria they used for classifying CLE images
as dysplastic or neoplastic. According to histopathologic analysis, the reference standard, high-grade dysplasia, was diagnosed in three patients and low-grade dysplasia was diagnosed in 16 patients, for an overall detection rate of 19 in 100 (19%) cases. Five cases were in the standard endoscopy group (one case of high-grade dysplasia and four cases of low-grade dysplasia) and 14 were in the CLE group (two cases of high-grade dysplasia and 12 cases of low-grade dysplasia). No suspicious lesions were identified in the standard endoscopy group and thus, only random biopsies were performed. In the CLE group, no suspicious lesions were identified when patients were initially evaluated with standard endoscopy but CLE detected areas suspicious for neoplasia in 21 of 50 (42%) of patients. All the cases of dysplasia were in patients with areas suspicious for neoplasia at CLE but not standard endoscopy. The sensitivity, specificity, PPV and NPV of probe-based CLE for detecting dysplasia were 100%, 83%, 67%, and 100%, respectively. Overall, the mean number of biopsies did not differ between groups (mean of 6.6 per patient in the standard endoscopy group and 6.1 in the CLE group, p=0.77), so the increased detection rate in the CLE group cannot be explained by a larger number of biopsies.

A single-center crossover RCT was published in 2009 by Dunbar.[22] This study was able to evaluate whether CLE can reduce the biopsy rate. This study was described in the systematic review and meta-analysis described by Gupta (2014) in the previous section. Forty-six patients with BE were enrolled, and 39 (95%) completed the study protocol. Of these, 23 were undergoing BE surveillance and 16 had BE with suspected neoplasia. All patients received endoscopy-based CLE and standard endoscopy, in random order. One endoscopist performed all CLE procedures and another endoscopist performed all standard endoscopy procedures; endoscopists were blinded to the finding of the other procedure. During the standard endoscopy procedure, biopsies were taken of any discrete lesions followed by four-quadrant random biopsy (every one cm for suspected neoplasia and every two cm for BE surveillance). During the CLE procedure, only lesions suspicious of neoplasia were biopsied. Endoscopists interpreted CLE images using the Confocal Barrett’s Classification system, developed in a previous research study. Histopathologic analysis was the reference standard. Among the 16 study completers with suspected high-risk dysplasia, there were significantly fewer biopsies per patient with CLE compared to standard endoscopy (mean of 9.8 biopsies vs 23.9 biopsies per patient, p=0.002). Although there were fewer biopsies, the mean number of biopsy specimens showing high-grade dysplasia or cancer was similar in the two groups: 3.1 during CLE and 3.7 during standard endoscopy, respectively. The diagnostic yield for neoplasia was 33.7% with CLE and 17.2% with standard endoscopy. None of the 23 patients undergoing BE for surveillance were found to have high-grade dysplasia or cancer. The mean number of mucosal specimens obtained for patients in this group was 12.6 with white-light endoscopy and 1.7 with CLE (p<0.001).

**Section Summary**

Several RCTs and a meta-analysis of RCTs and non-randomized, observational studies suggest that CLE has high accuracy for identifying dysplasia in patients with BE. A 2014 meta-analysis found that the pooled sensitivity, specificity, and negative predictive value in available studies is not sufficiently high to replace the standard Seattle protocol, according to criteria adopted by the American Society for Gastrointestinal Endoscopy (ASGE).

The sensitivity of CLE in the individual studies was higher than for white-light endoscopy alone, but the specificity was not consistently higher. There are limited data comparing standard protocols using random biopsies to protocols using CLE and targeted biopsies, so data are
inconclusive regarding the potential for CLE to reduce the number of biopsies in patients with
BE undergoing surveillance without compromising diagnostic accuracy. Moreover, studies do
not appear to use a consistent approach to classifying lesions viewed using CLE as dysplastic.

ASSESSING THE ADEQUACY OF ENDOSCOPIE TREATMENT OF GASTROINTESTINAL
LESIONS

Evidence is not clear regarding whether use of CLE improves the determination of residual
disease compared with conventional techniques (i.e. white-light endoscopy). In 2014,
Ypsilantis published a systematic review of the literature.[23] They included retrospective and
prospective studies that reported diagnostic accuracy of CLE for the detection of residual
disease after endoscopic mucosal resection (EMR) of gastrointestinal lesions. After examining
full-text articles, a total of three studies (one RCT and two prospective, non-randomized
comparative studies) met the eligibility criteria. Studies included patients with BE, gastric
neoplasia, and colorectal neoplasia. There was significant heterogeneity among studies. In a
per-lesion meta-analysis, pooled sensitivity of CLE for detecting neoplasmia was 91% (95% CI:
83% to 96%), and pooled specificity was 69% (95% CI: 61 to 76%). Based on the small
number of studies and heterogeneity among studies, the authors concluded that evidence on
the usefulness of CLE in assessing the adequacy of EMR is weak. The single RCT was
published in 2012 by Wallace[24]. This multicenter trial included patients with BE who were
undergoing ablation. After an initial attempt at ablation, patients were randomized to follow-up
with either high-definition white light (HDWL) endoscopy or HDWL endoscopy plus CLE.
The primary outcome was the proportion of optimally treated patients, defined as those with no
evidence of disease at follow-up, and those with residual disease who were identified and
treated. Enrollment in the study was halted after an interim analysis showed no difference
between groups. Among the 119 patients who had enrolled by the time of the interim analysis,
15 (26%) of 57 in the HDWL group and 17 (27%) of 62 in the HDWL plus CLE group were
optimally treated; the difference was not statistically significant. Moreover, other outcomes
were similar in the two groups.

Section summary

There is insufficient evidence that CLE improves upon standard practice for assessing the
adequacy of endoscopic treatment of gastrointestinal lesions. The single RCT on this topic was
stopped early because an interim analysis reported that CLE did not improve upon high-
definition white light endoscopy.

OTHER POTENTIAL APPLICATIONS OF CLE

Preliminary studies have been published evaluating CLE for diagnosing a variety of conditions
including lung cancer,[25-27] bladder cancer,[28,29] head and neck cancer,[30-33] gastric cancer,[34-
strictures and stenosis,[45,47-49] gastric intestinal metaplasia,[50-52] basal and squamous cell
carcinoma,[53] liver[54] and peritoneal nodules[55], inflammatory bowel disease,[56] and bile duct
malignancies.[57,58] There are insufficient studies to determine the accuracy of CLE for these
applications and their potential role in clinical care.

PRACTICE GUIDELINE SUMMARY

AMERICAN GASTROENTEROLOGICAL ASSOCIATION (AGA)
In 2011 the AGA published a position statement on the management of Barrett esophagus.\(^{[59]}\) The statement includes the following recommendations regarding endoscopic surveillance of Barrett esophagus:

The AGA suggest that endoscopic surveillance be performed in patients with Barrett esophagus (weak recommendation, moderate-quality evidence).

The AGA suggest the following surveillance intervals (weak recommendation, low-quality evidence):

- No dysplasia: three to five years
- Low-grade dysplasia: 6 to 12 months
- High-grade dysplasia in the absence of eradication therapy: three months

For patients with Barrett esophagus who are undergoing surveillance, the AGA recommended:

- Endoscopic evaluation be performed using white light endoscopy (strong recommendation, moderate-quality evidence).
- Four-quadrant biopsy specimens be taken every 2 cm (strong recommendation, moderate-quality evidence).
- Specific biopsy specimens of any mucosal irregularities be submitted separately to the pathologist (strong recommendation, moderate-quality evidence).
- Four-quadrant biopsy specimens be obtained every 1 cm in patients with known or suspected dysplasia (strong recommendation, moderate-quality evidence).

The AGA recommend against requiring chromoendoscopy or advanced imaging techniques for the routine surveillance of patients with Barrett esophagus at this time (weak recommendation, low-quality evidence).

**AMERICAN SOCIETY FOR GASTROINTESTINAL ENDOSCOPY (ASGE)**

In 2006 (reaffirmed in 2011), the ASGE published a guideline on the role of endoscopy in the surveillance of premalignant conditions of the upper gastrointestinal (GI) tract.\(^{[60]}\) The guideline included the following statements on surveillance of patients with BE:

The cost effectiveness of surveillance in patients without dysplasia is controversial. Surveillance endoscopy is appropriate for patients fit to undergo therapy, should endoscopic/histologic findings dictate. For patients with established Barrett's esophagus of any length and with no dysplasia, after 2 consecutive examinations within 1 year, an acceptable interval for additional surveillance is every 3 years.

Patients with high-grade dysplasia are at significant risk for prevalent or incident cancer. Patients who are surgical candidates may elect to have definitive therapy. Patients who elect surveillance endoscopy should undergo follow-up every 3 months for at least 1 year, with multiple large capacity biopsy specimens obtained at 1 cm intervals. After 1 year of no cancer detection, the interval of surveillance may be lengthened if there are no dysplastic changes on 2 subsequent endoscopies performed at 3-month intervals. High-grade dysplasia should be confirmed by an expert GI pathologist.

Surveillance in patients with low-grade dysplasia is recommended. The significance of low-grade dysplasia as a risk factor for cancer remains poorly defined; therefore, the
optimal interval and biopsy protocol has not been established. A follow-up EGD (screening esophagogastroduodenoscopy) (i.e., at 6 months) should be performed with concentrated biopsies in the area of dysplasia. If low-grade dysplasia is confirmed, then one possible management scheme would be surveillance at 12 months and yearly thereafter as long as dysplasia persists.

The ASGE Technology Committee published a Technology Status Evaluation Report on CLE in 2014. The report concluded that CLE is an emerging technology with the potential to improve patient care. However, before the technology can be widely accepted, further studies are needed in the following areas:

- Use of CLE outside of the academic setting, particularly the applicability of the technology in community settings.
- The learning curve of CLE image interpretation and any additional time needed to perform the procedure.
- The clinical efficacy of the technology compared to other available advanced imaging technologies.
- Approaches to CLE imaging and image interpretation.

In 2016, based on a systematic review of 102 studies conducted between 2004 and 2015, the ASGE concluded additional clinical trials on CLE are still necessary.

### SUMMARY

There is not enough research to know if or how well confocal laser endomicroscopy (CLE) works to improve health outcomes for people with any condition. This does not mean that it does not work, but more research is needed to know. Therefore, use of CLE with endoscopy is considered investigational for all indications.

### REFERENCES


<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0397T</td>
<td>Endoscopic retrograde cholangiopancreatography (ERCP), with optical endomicroscopy (List separately in addition to code for primary procedure)</td>
</tr>
<tr>
<td></td>
<td>43206</td>
<td>Esophagoscopy, flexible, transoral; with optical endomicroscopy</td>
</tr>
<tr>
<td></td>
<td>43252</td>
<td>Esophagogastroduodenoscopy, flexible, transoral; with optical endomicroscopy</td>
</tr>
<tr>
<td></td>
<td>88375</td>
<td>Optical endomicroscopic image(s), interpretation and report, real-time or referred, each endoscopic session.</td>
</tr>
</tbody>
</table>

**Date of Origin:** July 2014
Coverage of Treatments Provided in a Clinical Trial

Effective: December 1, 2017

Next Review: October 2018
Last Review: November 2017

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic.Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Effective January 1, 2014, the Affordable Care Act (ACA) requires group health plans or a health insurance issuer offering group or individual health insurance coverage to provide coverage for routine patient costs associated with participating in an approved clinical trial.[1] This policy is written to assist in applying Sec. 2709 of the ACA, Coverage for Individuals Participating in Approved Clinical Trials.

MEDICAL POLICY CRITERIA

Routine patient costs associated with approved clinical trials may be considered medically necessary for qualified individuals with respect to treatment of cancer or other life threatening disease or condition, when the Affordable Care Act definitions for clinical trial participation are met.

• See Background for definitions.
• See Policy Guidelines for clinical trial registry resource.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

POLICY GUIDELINES

It is critical that the list of information below is submitted for review to determine if the policy

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.
Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.

1. Pertinent History and Physical, including specific diagnosis and treatment history
2. Clinical trial name and the NCT number
3. Phase of the trial
4. Currently planned, requested interventions
5. Anticipated possible interventions

ClinicalTrials.gov includes a registry of publicly and privately supported clinical studies.

CROSS REFERENCES
None

BACKGROUND

DEFINITIONS

• Routine patient costs
  o Routine patient costs include all items and services consistent with the coverage provided in the plan (or coverage) that is typically covered for a qualified individual who is not enrolled in a clinical trial.
  o Routine patient costs do not include the investigational item, device, or service, itself; items and services that are provided solely to satisfy data collection and analysis needs and that are not used in the direct clinical management of the patient; or a service that is clearly inconsistent with widely accepted and established standards of care for a particular diagnosis.

• Approved clinical trial

An approved clinical trial is defined as a phase I, phase II, phase III, or phase IV clinical trial that is conducted in relation to the prevention, detection, or treatment of cancer or other life-threatening disease or condition, and that is described by any of the following:

  o The study or investigation is approved or funded by one or more of the following:
    ▪ The National Institutes of Health
    ▪ The Centers for Disease Control and Prevention
    ▪ The Agency for Health Care Research and Quality
    ▪ The Centers for Medicare & Medicaid Services
    ▪ A cooperative group or center of any of the above four entities or the Department of Defense or the Department of Veterans Affairs
    ▪ A qualified non-governmental research entity identified in the guidelines issued by the National Institutes of Health for center support grants
    ▪ The Department of Veterans Affairs, the Department of Defense or the Department of Energy if the study or investigation has been reviewed and approved through a system of peer review that the Secretary determines to be comparable to the system of peer review of studies and investigations used by the National Institutes of Health, and assures
unbiased review of the highest scientific standards by qualified individuals who have no interest in the outcome of the review; OR

- The study or investigation is conducted under an investigational new drug application reviewed by the Food and Drug Administration; OR
- The study or investigation is a drug trial that is exempt from having such an investigational new drug application.

- Life-threatening condition

A life-threatening condition is defined as any disease or condition from which the likelihood of death is probable unless the course of the disease or condition is interrupted.

- Qualified individual

A participant who is a beneficiary in a health plan who is eligible to participate in an approved clinical trial according to the trial protocol with respect to treatment of cancer or another life threatening disease or condition and either:

- The referring health care professional is a participating health care provider and has concluded that the individual’s participation in such trial would be appropriate based upon the individual meeting the clinical trial eligibility requirements; or
- The participant or beneficiary provides medical and scientific information establishing that the individual’s participation in such trial would be appropriate based upon the individual meeting the clinical trial eligibility requirements.

REFERENCES


CODES

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCPCS</td>
<td>S9988</td>
<td>Services provided as part of a phase I clinical trial</td>
</tr>
<tr>
<td></td>
<td>S9990</td>
<td>Services provided as part of a phase II clinical trial</td>
</tr>
<tr>
<td></td>
<td>S9991</td>
<td>Services provided as part of a phase III clinical trial</td>
</tr>
</tbody>
</table>

Date of Origin: November 2013

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
**Corneal Collagen Cross-Linking**

**Effective:** January 1, 2018

**Next Review:** October 2018  
**Last Review:** November 2017

**IMPORTANT REMINDER**

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

**DESCRIPTION**

Corneal collagen cross-linking (CXL) is a procedure performed in the outpatient setting using topical anesthesia with the photosensitizer riboflavin (vitamin B2) and ultraviolet-A (UVA) irradiation done to increase corneal rigidity and stability for a number of corneal conditions.

**MEDICAL POLICY CRITERIA**

I. Epithelium-off collagen cross-linking using riboflavin and ultraviolet A may be considered **medically necessary** for the treatment of keratoconus and keratectasia (corneal ectasia).

II. Epithelium-on (transepithelial) collagen cross-linking is considered **investigational** for keratoconus, keratectasia, and all other indications.

III. Any type of collagen cross-linking is considered **investigational** in all other situations, including but not limited to treatment of infectious keratitis and in combination with other procedures (e.g., intrastromal corneal ring segments, PRK or phakic intra-ocular lens implantation) (CXL-plus).

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.
**POLICY GUIDELINES**

Epithelium-off (also referred to as “epi-off”) CXL is a procedure in which the epithelium is removed or weakened by various methods to allow penetration of riboflavin into the corneal tissue prior to crosslinking with UV-A light. Currently the only FDA-approved CXL treatment is the KXL® system (Avedro), which was approved using the epithelium-off procedure.

Epithelium-on (also known as transepithelial or “epi-on”) CXL is a procedure in which the corneal epithelial surface is left intact (or may be partially disrupted) and a longer riboflavin loading time is needed.

**CROSS REFERENCES**

None

**BACKGROUND**

Corneal collagen cross-linking (CXL) is performed with the photosensitizer riboflavin (vitamin B2) and ultraviolet A (UV-A) irradiation. A common CXL protocol removes about 8 mm of the central corneal epithelium under topical anesthesia to allow better diffusion of the photosensitizer riboflavin into the stroma. Following de-epithelialization, a solution with riboflavin is applied to the cornea (every 1-3 minutes for 30 minutes) until the stroma is completely penetrated. The cornea is then irradiated for 30 minutes with 370 nm ultraviolet A, a maximal wavelength for absorption by riboflavin, while the riboflavin continues to be applied. The interaction of riboflavin and UV-A causes the formation of reactive oxygen species, leading to additional covalent bonds (cross-linking) between collagen molecules, resulting in stiffening of the cornea. Theoretically, by using a homogeneous light source and absorption by riboflavin, the structures beyond a 400-micron thick stroma (endothelium, anterior chamber, iris, lens, retina) are not exposed to a UV dose that is above the cytotoxic threshold.

CXL is being evaluated primarily for corneal stabilization in patients with progressive corneal thinning, such as keratoconus. CXL may also have anti-edematous and antimicrobial properties.

Keratoconus is a bilateral dystrophy that is characterized by progressive ectasia (paracentral steepening and stromal thinning) that impairs visual acuity. While frequently diagnosed at a young age, the progression of keratoconus is variable. Initial treatment often consists of hard contact lenses. A variety of keratorefractive procedures have also been attempted, broadly divided into subtractive and additive techniques. Subtractive techniques include photorefractive keratectomy or laser in situ keratomileusis (LASIK), although generally, results of these techniques have been poor. Implantation of intrastromal corneal ring segments is an additive technique in which the implants are intended to reinforce the cornea, prevent further deterioration, and potentially obviate the need for penetrating keratoplasty. Penetrating keratoplasty (i.e., corneal grafting) is the last line of treatment. About 20% of patients with keratoconus will require corneal transplantation. All of these treatments attempt to improve the refractive errors, but are not disease-modifying. In contrast, corneal CXL has the potential to slow the progression of disease.

Ectasia (also known as keratectasia, iatrogenic keratoconus or secondary keratoconus) is a serious long-term complication of LASIK surgery. Reported treatments for the management of
post-LASIK ectasia include hard contact lenses, intraocular pressure-lowering drugs, and intracorneal ring segments. Frequently, a penetrating keratoplasty is required.

There are two different methods of cross-linking the collagen in the cornea. According to the National Institute for Health and Care Excellence (NICE) regarding the different variations of the collagen corneal cross-linking (CXL) procedure:[1]

1. Epithelium-off CXL (also known as “epi-off”): the epithelium is first removed or weakened, typically by abrasion, to allow penetration of riboflavin into the corneal tissue. Riboflavin eye drops are applied to the corneal surface before the procedure and intermittently during the procedure. The corneal surface is exposed to UVA radiation: precise timings and treatment protocols vary. Postoperatively, topical antibiotics and anti-inflammatory drops are normally prescribed, with topical steroids if necessary. In some cases, a bandage contact lens may also be used for a few days. The procedure is done on 1 eye at a time and may also be repeated if needed.

2. Epithelium-on (also known as “epi-on” or transepithelial) CXL: the corneal epithelial surface is left intact (or may be partially disrupted) and a longer riboflavin loading time is needed.

Currently the only FDA-approved CXL treatment is the KXL® system (Avedro), which is only indicated using the epithelium-off method. There are no FDA approved CXL treatments using the epithelium-on method of CXL.

Other procedures performed to address various corneal conditions, including keratoconus, include:

1. Photorefractive keratectomy (PRK) is a refractive surgical procedure involving the reshaping of the surface of the cornea with an excimer laser for correction of refractive errors (e.g., myopia, hyperopia, astigmatism, and presbyopia) in persons with otherwise non-diseased corneas.

2. Intrastromal Corneal Ring Segments (INTACS) are flexible rings that come in different sizes that are inserted beneath the surface of the cornea to elevate the edge of the cornea. This procedure flattens the front of the eye, decreasing nearsightedness.

3. Phakic Implantable Contact Lenses (IOLs) are thin lenses implanted permanently into the eye to help reduce the need for glasses or contact lenses. Phakic refers to the fact that the lens is implanted into the eye without removing the eye's natural lens. During phakic lens implantation surgery, a small incision is made in the front of the eye and the lens is inserted through the incision and placed just in front of or just behind the iris.

REGULATORY STATUS

In April 2016 the US Food and Drug Administration (FDA) has approved a riboflavin ophthalmic solution (Photrex and Photrex Viscous, Avedro) in combination with the company's particular UVA irradiation device, marketed as the KXL® system (Avedro), for the treatment of progressive keratoconus. The KXL® system is used to perform corneal collagen cross-linking using Photrex (riboflavin 5’-phosphate ophthalmic solution) 0.146% for topical ophthalmic use or Photrex Viscous (riboflavin 5’-phosphate in 20% dextran ophthalmic solution) 0.146% for topical ophthalmic use with UV-A (NDA 203324). The FDA clinical trials that served as a basis for the FDA approval of the KXL System indicated that all 640 patients...
were treated with the "epithelium-off" CXL method. Also, as indicated in the "Dosage and Administration" section of the NDA document: the first step in the protocol is to debride the epithelium, which is the epithelium-off method.

In July 2016 the NDA was amended and the approval was extended for the use of the KXL System for the treatment of corneal ectasia following refractive surgery.

The KXL® system has not been approved for the use of infectious keratitis, corneal ulcers, or any other indication.

In addition, there have been FDA trials for the VEGA system (TopCon Medical Systems) that have been terminated based on administrative reasons before the data was analyzed. The status of the VEGA system is currently unknown.

**EVIDENCE SUMMARY**

Evidence on whether corneal collagen cross-linking (CXL) improves health outcomes for patients with progressive keratoconus includes systematic reviews and six randomized controlled trials (RCTs), three of which were regulated by the U.S. Food and Drug Administration (FDA) under a new drug application (NDA), one of which is unpublished. In addition, there are a number of prospective controlled studies as well as uncontrolled trials that report on longer term outcomes of the procedure. The main health outcome for corneal CXL treatment is improvement, or stabilization, of visual acuity. Other outcomes commonly reported in trials of CXL include physiologic measures, such as the steepness of the corneal curvature measured by maximum keratometry (K-max) and/or the manifest refraction spherical equivalent. These are intermediate outcomes that may corroborate whether improvements in visual acuity correlate with physiologic changes.

**SYSTEMATIC REVIEWS**

In 2016 a systematic review assessed the efficacy and safety of CXL in pediatric patients with keratoconus, including 17 unique articles: 10 articles on epithelium-off cross-linking, 2 on accelerated cross-linking, 2 on transepithelial cross-linking, 1 on both epithelium-off and transepithelial cross-linking, and 2 on transepithelial cross-linking with iontophoresis. The reviewers concluded that epithelium-off cross-linking is both apparently safe and effective when used to prevent keratoconus progression in pediatric patients. However, disease progression occurred in 22% of the treated eyes.

In 2016 a systematic review assessed the efficacy of CXL for the treatment of keratoconus (KCN). A modest, but not significant improvement in visual acuity of 1 to 2 Snellen lines was found three months or more after undergoing CXL. Changes were more pronounced in uncorrected visual acuity. Some secondary outcomes were found to be improved (0.6-1 diopters) 12 to 24 months after CXL, but others were not. The reviewers concluded that although CXL appears to be effective for halting the deterioration of KCN it was only slightly effective at improving visual function.

In 2016 a systematic review assessed the efficacy of CXL in the management of infectious keratitis. Twenty-five studies were included (2 randomized controlled trials, 13 case series, and 10 case reports) with a total of 210 eyes of 209 patients, of which 175 eyes underwent CXL. Proportion of eyes healed with CXL was 87.2% (95% confidence interval (CI), 81.9%,...
91.8%). The reviewers concluded that although CXL seems promising in the management of infectious keratitis, more randomized controlled trials are required to assess its efficacy.

A Cochrane review on the use of corneal CXL for the treatment of keratoconus was published in 2015.[6] The literature search for this systematic review was conducted in August 2014 and does not include the three unpublished phase 3 trials that were submitted to FDA (described below). The review included three small RCTs conducted in Australia, the United Kingdom, and the United States that enrolled a total of 225 eyes and analyzed 219 eyes.[7-9] The total number of people enrolled was not clear in two of the studies. Only adults were enrolled into these studies. Out of the eyes analyzed, 119 were treated with CXL (all using the epithelium-off technique) and 100 served as controls. Only one study had sham treatments for controls. All three studies were at high risk for performance bias (lack of masking), detection bias (only one trial attempted to mask outcome assessment), and attrition bias (incomplete follow-up). It was not possible to pool data due to differences in measuring and reporting outcomes. The overall quality of the evidence was judged to be very low primarily due to downgrading the evidence due to risk of bias in the included studies, imprecision, indirectness and publication bias.

RANDOMIZED CONTROLLED TRIALS

In 2017, Hersh published results from a multi-center clinical trial of corneal collagen cross-linking for the treatment of corneal ectasia after refractive surgery.[10] The treatment group underwent CXL whereas the sham group received riboflavin alone without removal of the epithelium. A total of 179 subjects were enrolled and the primary outcome was the one-year change in topography-derived maximum keratometry (K). Secondary outcomes included corrected distance visual acuity (CDVA), uncorrected distance visual acuity (UDVA), manifest refraction spherical equivalent, endothelial cell count, and adverse events. The authors concluded that significant improvements occurred in the treatment group for K value, CDVA, and UDVA after one year of treatment.

A 2016 study by Rush compared epithelium-off versus transepithelial CXL in 144 eyes with progressive corneal ectasia.[11] The primary outcome was change in maximum simulated keratometry value (Ksteep) after 24 months and the secondary outcome was change in best spectacle-correct visual acuity (BSCVA). The authors reported a significant improvement in Ksteep in the epithelium-off group compared to the transepithelial group (p=0.032) after 24 months of follow up. There was no significant difference between groups for BSCVA.

In 2016, Bikbova published results from a randomized trial with 24 months of follow-up comparing standard CXL to transepithelial iontophoresis-assisted CXL in 149 eyes (119 subjects).[12] The authors reported a statistically significant difference in corrected distance visual acuity (CDVA) between the two groups, with a better outcome in the transepithelial group after 6 months (p = 0.037); however, no significant difference was found 24 months after treatment (p = 0.829). Stabilization and regression of keratometry values were achieved in both groups, but standard CXL was more effective.

In 2014 Wittig-Silva reported three-year results from the first RCT of corneal epithelium-off CXL in 2008.[9,13] Recruitment for the trial was completed in 2009 with 50 eyes randomized to CXL treatment and 50 randomized to untreated control. To be eligible for enrollment, clear evidence of progression of ectasia over the preceding 6 to 12 months was required. Progression was confirmed if at least one of the following criteria were met: an increase of at least 1 D in the steepest simulated keratometry reading (K-max); an increase in astigmatism

---

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
determined by manifest subjective refraction of at least 1 D; an increase of 0.50 D in manifest refraction spherical equivalent; or a 0.1 mm or more decrease in back optic zone radius of the best-fitting contact lens.

At the time of analysis for the 2008 report, 20 eyes had reached 1-year follow-up. The 3-year results included 46 CXL-treated and 48 control eyes. LOCF was used for 26 eyes, including 17 eyes from the control group with progressive disease that underwent compassionate-use CXL or corneal transplantation. In the CXL group, there was a flattening of K-max by -1.03 D, compared with an increase in K-max of 1.75 in the control group. One eye in the CXL group progressed by more than 2 D, compared with 19 eyes in the control group. Uncorrected visual acuity (UCVA) and best-spectacle corrected visual acuity (BSCVA) improved in the CXL-treated eyes at 1, 2, and 3 years. In control eyes, UCVA was significantly reduced at 36 months (p=0.034) and there was a trend of a decrease in BSCVA (p=0.10). The difference between groups in UCVA was significant (p<0.001), but there was no between group difference in BSCVA. At three-year follow-up the authors concluded that “despite the growing body of literature and continuing efforts to optimize the treatment protocol, there remains a lack of randomized controlled studies with longer-term follow-up to support the widespread clinical use of CXL for keratoconus”.

In 2013 a small randomized trial was published which assessed CXL treatment for pseudophakic bullous keratopathy; however CXL followed by keratoplasty was performed in all 24 patients within the study, limiting any conclusion regarding the safety and efficacy of CXL treatment compared to other methods.[14]

In 2012, Renesto reported 2-year results of a randomized trial that compared CXL versus 1 month of riboflavin eye drops in 39 eyes of 31 patients with keratoconus.[15] After 3 months, all patients received intrastromal corneal ring segments (ICRS; see evidence review 9.03.14). Patients were evaluated at 1 and 3 months after treatment with CXL or riboflavin, and then at 1, 3, 6, 12, and 24 months after ICRS insertion. There was no significant difference between the 2 groups for UCVA, BCVA, or in 3 topographic parameters (flattest-K, steepest K, and average keratometry) throughout the 24-month follow-up.

Data submitted to FDA under the NDA for riboflavin ophthalmic solution/KXL System came from three RCTs with a total anticipated sample size of 640 patients.[16] Results from the first of the trials were published in 2011 and 2012 (100 eyes in 76 patients) are described below.[7,17] Each of the phase 3 trials was a parallel group, open-label trial in patients with keratoconus or corneal ectasia due to laser in situ keratomileusis (LASIK) or photorefractive keratectomy. Sham-control eyes were treated with a topical anesthetic and riboflavin solution (1 drop every 2 minutes for 30 minutes) but did not undergo epithelial débridement or have the ultraviolet A (UVA) light source turned on. The primary outcome was a difference of 1 D or more in the mean change in K-max (progression of steepening) between the CXL group and control group at 12 months. Control patients could cross over to CXL at 3 months; by 12 months, 99% had done so. Missing data were analyzed by last observation carried forward (LOCF), which is a conservative method of analysis in this situation, because it reduces the expected worsening over time in untreated patients. In the pooled analysis of patients with keratoconus, steepening worsened by 1 D in the control group and improved by 1.6 D in the CXL group, for a total difference between groups of 2.6 D. Pooled analysis showed that CXL resulted in either stabilization or improvement in K-max in 72% of keratoconus patients. In the sham control group, there was no statistically significant change in K-max. The mean improvement in BCVA was 5.6 letters following corneal CXL compared with 2 letters for sham treatment (p=0.009).
Although this difference is not typically considered clinically significant, it is limited by the use of 3-month data for many of the patients in the control group, which would minimize between-group differences over time. The proportion of patients who had a clinically significant 3-line or greater improvement in BCVA was 19.4% for the CXL-treated patients and 8.1% for controls. Treatment-related adverse events were generally transient, mild, and expected, based on the epithelial débridement and corneal remodeling. Although these RCTs had sham-controlled groups for comparison, these sham patients were allowed to crossover into the treated group as early as three months into the study.

In 2011, Hersh and colleagues published one-year results of FDA-approved unblinded sham-controlled clinical of CXL for keratoconus or ectasia in 71 eyes of 58 patients (active treatment), 41 eyes received sham treatment and an additional 30 eyes were included as a “fellow-eye” control group.[7] The control group consisted of a heterogenous group of eyes from patients who had unilateral CXL treatment (eyes with some evidence of disease, but which may or may not have met study inclusion criteria and eyes without evidence of disease). Patients in the sham control group received riboflavin 0.1% ophthalmic solution without UVA treatment and at three months were given CXL treatment, thus ending the comparative portion of the trial. Visual acuity, refraction, astigmatism and maximum and average keratometry (k) values were primary outcomes. There were no statistically significant differences between treatment and control groups (results for sham and “fellow eye” groups were aggregated) at three months. However, interpretation of these results is limited by the lack of clear target population (results from patients with either keratoconus or ectasia were aggregated into the same treatment group, thus limiting the population to which these results can be generalized). Also, isolation of the impact of CXL on corneal disease requires that any control treatment group be identical to the treatment group in as many aspects as possible with the exception of the treatment itself. Such a heterogenous control group as included here (the “fellow eye” group) may not allow for the isolation of this treatment effect from normal disease course or other components of care. These factors, along with the lack of comparative study beyond 3 months suggest that these results are inconclusive regarding the impact of CXL on corneal disease and that longer randomized controlled trials are needed to clearly evaluate the impact of this treatment.

NONRANDOMIZED STUDIES

A 2017 study by Aixinjueluo reported results from 30 eyes (19 subjects) with progressive keratoconus.[18] Outcomes of interest included uncorrected visual acuity, best corrected visual acuity (BCVA), average keratometry (AveK), maximum keratometry (Kmax), central corneal thickness, thinnest corneal thickness (TCT), endothelial cell density, intraocular pressure and non-mydriatic indirect fundus examination. After 12 months of follow-up, there was a significant decrease in Kmax (p<0.0001), AveK (p=0.003) and TCT (p=0.002), and a significant improvement in BCVA (p=0.001).

Longer term follow-up is being reported from Europe, where corneal CXL has been performed for a greater number of years. Indications for treatment typically include progression of steepening (increase in K-max by at least 1 D in 1 year), deteriorating visual acuity, or the need to be fitted for new contact lenses more than once in two years. The largest and longest series to date are described next.

In 2015, Raiskup published 10-year follow-up of CXL treatment in 34 eyes (24 patients) with progressive keratoconus.[19] Mean patient age at the time of treatment was 28 years (range,
Corneal steepening improved slightly between baseline and 10-year follow-up (p<0.001), while corrected distance visual acuity improved by 0.14 logMAR (p=0.002). Two eyes had repeat CXL, one after 5 years and one after 10 years, without adverse sequelae. One of the 34 eyes treated developed a permanent corneal. The original study, published in 2008 by the same group, reported outcomes of 241 eyes (130 patients) treated with CXL, with a minimum of 6 months of follow-up. This was of a total of 488 eyes (272 patients) with progressive keratoconus and a corneal thickness of at least 400 μm treated at their center in Germany. Follow-up examinations were performed at 1, 6, and 12 months, and then annually. Mean follow-up was 26 months, with a range of 12 months (n=142) to 6 years (n=5). In the first year (n=142), steepening (K-max) improved or remained stable in 86% of eyes, and BCVA improved by at least 1 line in 53% of the eyes. Three years after treatment (n=33), K-max improved by a mean of 2.57 D in 67% of eyes while BCVA improved by at least 1 line in 58% of eyes.

In 2014, Said published a small (n=40) study which compared corneal collagen cross-linking (CXL) with photoactivated riboflavin to antimicrobial therapy as a treatment of infectious keratitis with corneal melting. Authors reported comparable duration until healing; however the complication rate was 21% (3 patients) in the control group. No incidence of corneal perforation or recurrence of the infection in the CXL group was reported. This study was limited by lack of randomization, small sample size and relatively short-term follow-up.

A 2012 publication from the Siena CXL Pediatrics trial reported 12- to 36-month follow-up after CXL in 152 patients aged 18 years or younger with keratoconus progression. Visual acuity increased by an average of 0.15 Snellen lines, whereas a clinically relevant change is generally considered to be 2 Snellen lines.

The French National Reference Center for Keratoconus published their findings in 2011. Of 142 eyes enrolled in the study, 6-month follow-up was available for 104 (73%), and 12-month follow-up was available for 64 (45%). At 12 months after treatment, the BCVA had stabilized in 48% of eyes, improved in 40%, and decreased in 12%. Keratoconus progression had stopped in 69%, and K-max had decreased by more than 2 D in 21% of eyes. There was a 7% complication rate in the total sample, with 5 eyes (3.5% of 142 or 7.8% of 64) losing two or more Snellen lines of visual acuity. This retrospective study is limited by the low proportion of patients available at 12-month follow-up.

A 2010 publication from the Siena Eye Cross Study reported a 52-month mean follow-up (range, 48-60 months) on their first 44 keratoconic eyes treated with CXL. Follow-up evaluations were performed at 1, 2, 3, 6, 12, 24, 36, 48, and 60 months after CXL. Topographic analysis showed the following mean K reading reductions: -1.96 D after 1 year, -2.12 D after 2 years, -2.24 D after 3 years, and -2.26 D after 4 years of follow-up. By comparison, in fellow eyes untreated for the first 24 months, the mean K value increased by 1.2 D at 1 year and 2.2 D at 2 years. In treated eyes, UCVA improved by a mean of 2.41 lines after 12 months, 2.75 lines after 24 months, 2.80 lines after 36 months, and 2.85 lines after 48 months. There was no significant decrease in endothelial cell density, central corneal thickness, or intraocular pressure over follow-up. Temporary adverse effects included stromal edema in the first 30 days (70% of patients) and temporary haze (9.8% of patients). No persistent adverse effects were observed.

**ADVERSE EVENTS**
Reported adverse events are relatively uncommon, but precise rates of adverse events are not available because of the lack of large studies with long-term follow-up. Treatment-related adverse events are generally transient, mild, and expected, based on the epithelial débridement and corneal remodeling. Persistent adverse events have been rarely observed. Adverse events reported to date include corneal endothelial damage, stromal haze, corneal melt, keratitis, gaping of corneal incisions, and corneal scarring.[25-27]

SUMMARY OF EVIDENCE

The evidence for corneal CXL in individuals who have keratoconus includes randomized controlled trials (RCTs) and systematic reviews. Relevant outcomes are change in disease status, functional outcomes, and treatment-related morbidity. There is evidence from RCTs, including several pivotal trials that CXL leads to short-term improvements in corneal steepening and visual acuity compared with untreated eyes, and results from one trial have reported that these benefits are maintained at two to three years. In addition, retrospective studies have reported positive outcomes up to 10 years, although these reports have small sample sizes at long-term follow-up and limited information on the entire population of patients treated with corneal CXL during the same time period. The available evidence reports that the procedure is generally safe.

PRACTICE GUIDELINE SUMMARY

In 2013, the National Institute for Health and Care Excellence (NICE) issued an Interventional Procedure Guideline (IPG 466) that replaced the 2009 IPG 320 and were based on a systematic review of the evidence.[1] The new IPG now stratifies their recommendations for corneal CXL as follows:

“Most of the published evidence on photochemical corneal collagen cross linkage (CXL) using riboflavin and ultraviolet A (UVA) for keratoconus and keratectasia relates to the technique known as ‘epithelium-off’ CXL. ‘Epithelium on (transepithelial) CXL’ is a more recent technique and less evidence is available on its safety and efficacy. Either procedure (epithelium off or epithelium on CXL) can be combined with other interventions, and the evidence base for these combination procedures (known as ‘CXL-plus’) is also limited. Therefore, different recommendations apply to the variants of this procedure, as follows.

1.1 Current evidence on the safety and efficacy of epithelium off CXL for keratoconus and keratectasia is adequate in quality and quantity. Therefore, this procedure can be used provided that normal arrangements are in place for clinical governance, consent and audit.

1.2 Current evidence on the safety and efficacy of epithelium on (transepithelial) CXL, and the combination (CXL-plus) procedures for keratoconus and keratectasia is inadequate in quantity and quality. Therefore, these procedures should only be used with special arrangements for clinical governance, consent and audit or research….”

SUMMARY

There is enough research to show that epithelium-off collagen cross-linking using riboflavin and ultraviolet A improves health outcomes for patients with keratoconus and keratectasia (corneal ectasia). In addition, there are good quality evidence-based clinical practice guidelines that recommend the use of epithelium-off collagen cross-linking using riboflavin...
and ultraviolet A for the treatment of keratoconus and keratectasia. Therefore, epithelium-off collagen cross-linking using riboflavin and ultraviolet A may be considered medically necessary for the treatment of keratoconus and keratectasia.

There is not enough research to show that epithelium-on collagen cross-linking improves health outcomes for patients with keratoconus, keratectasia (corneal ectasia), or any other condition. In addition, evidence-based clinical practice guidelines recommend against the use of epithelium-on collagen cross-linking for the treatment of keratoconus and keratectasia. Therefore, epithelium-on collagen cross-linking is considered investigational for the treatment of any condition, including but not limited to keratoconus and keratectasia.

There is not enough research to show that collagen cross-linking of any type improves health outcomes in patients with infectious keratitis. In addition, there is not enough research to show that collagen cross-linking, when done in combination with other procedures (e.g., intrastromal corneal ring segments, PRK or phakic intra-ocular lens implantation) (known as CXL-plus), improves health outcomes in patients with any condition. Therefore, any type of collagen cross-linking, is considered investigational in all other situations, including but not limited to treatment of infectious keratitis and in combination with other procedures (e.g., intrastromal corneal ring segments, PRK or phakic intra-ocular lens implantation) (CXL-plus).

REFERENCES


10. Hersh, PS, Stulting, RD, Muller, D, Durrie, DS, Rajpal, RK. United States Multicenter Clinical Trial of Corneal Collagen Crosslinking for Keratoconus Treatment. *Ophthalmology*. 2017 Sep;124(9):1259-70. PMID: 28495149


November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


**CODES**

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0402T</td>
<td>Collagen cross-linking of cornea (including removal of the corneal epithelium and intraoperative pachymetry when performed)</td>
</tr>
<tr>
<td></td>
<td>66999</td>
<td>Unlisted procedure, anterior segment of the eye</td>
</tr>
<tr>
<td>HCPCS</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>

*Date of Origin: October 2016*
Gait Analysis

Effective: April 1, 2018

Next Review: March 2019
Last Review: March 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Gait analysis (GA) is the quantitative assessment of coordinated muscle function; evaluation is conducted in a laboratory and typically involves a dedicated facility and staff. A visual assessment of walking is supplemented by video recording. Videos can be observed from several visual planes at slow speed, allowing detection of movements not observable at normal speed. Joint angles and various time-distance variables, including step length, stride length, cadence, and cycle time, can be measured. Electromyography (EMG), assessed during walking, may be an included component of gait analysis and measures timing and intensity of muscle contractions. This calculation allows determination of whether a certain muscle’s activity is normal, out of phase, continuous, or clonic.

MEDICAL POLICY CRITERIA

Note: Surface electromyography (SEMG) may be included as a component of gait analysis. See Medical Policy, Medicine No. 73, Surface Electromyography (SEMG) Including Paraspinal SEMG for specific criteria regarding SEMG.

I. Gait analysis may be considered medically necessary in children and adolescents with cerebral palsy to select surgical or other therapeutic interventions for gait improvement.

II. All other indications for gait analysis are considered investigational.
NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

**POLICY GUIDELINES**

It is critical that the list of information below is submitted for review to determine if the policy criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.

- History and physical/chart notes
- Indication for Gait Analysis

**CROSS REFERENCES**

1. [Surface Electromyography (SEMG) Including Paraspinal SEMG](#), Medicine, Policy No. 73

**BACKGROUND**

Gait analysis has been proposed as an aid in surgical planning, primarily for cerebral palsy (CP), but also for other conditions such as clubfoot. In addition, gait analysis is being investigated as a means to plan rehabilitative strategies (i.e., orthotic-prosthetic devices) for ambulatory problems related to cerebral palsy, aging, stroke, spinal cord injury, and other conditions.

Kinematics is the term used to describe movements of joints and limbs such as angular displacement of joints and angular velocities and accelerations of limb segments. The central element of kinematic assessment is some type of marker system that is used to represent anatomic landmarks, which are then visualized and quantitatively assessed by videotaped observations or optoelectronic data. Movement data are compiled by computer from cameras oriented in several planes, and the movement data are processed so that the motion of joints and limbs can be assessed in three dimensions. The range and direction of motion of a particular joint can be isolated from all the other simultaneous motions that are occurring during walking. Graphic plots of individual joint and limb motion as a function of gait phase can be generated.

Inertial and magnetic measurement systems (IMMSs) are under investigation for the assessment of joints and limbs in 3-dimensions.[1,2] Rather than videotaped or optoelectronic calibration of markers placed on anatomic landmarks, IMMS systems involve sensor units that are comprised of miniaturized 3-dimensional accelerometers, gyroscopes, and magnetometers that are attached to body segments. The 3-dimensional orientation of each sensor is measured in relationship to an earth-based coordinate system through the use of computerized algorithms. One protocol, the “Outwalk” protocol, has been developed to allow the use of an IMMS system for gait analysis.

A non-profit organization established in 1997, the Commission for Motion Laboratory Accreditation evaluates and accredits motion laboratories within clinical facilities. A multidisciplinary team uses a set of criteria to evaluate laboratories in the areas of administration (e.g., staffing, policies, and procedures), equipment (e.g., accuracy and precision), and data management and reporting (e.g., control and clinical data sets).

**REGULATORY STATUS**

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.

Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Gait analysis devices, approved through the FDA 510(k) process include but are not limited to the Peak Motus Motion Measurement System, the Coda cx1 Motion Analysis System, KneeKG and Smart.[3]

**EVIDENCE SUMMARY**

Assessment of a diagnostic technology typically focuses on three parameters: 1) technical feasibility; 2) diagnostic performance (sensitivity, specificity, and positive [PPV] and negative predictive value [NPV]) in appropriate populations of patients; and 3) demonstration that the diagnostic information can be used to improve patient outcomes (clinical utility).

Technical feasibility of a device is typically assessed with two types of studies, those that compare test measurements with a gold standard and those that compare results taken with the same device on different occasions (test-retest). Demonstration of technical feasibility should include an assessment of its reproducibility and precision.

Diagnostic performance is evaluated by the ability of a test to accurately diagnose a clinical condition in comparison with the gold standard. The sensitivity of a test is the ability to detect a disease when the condition is present (true-positive), while specificity indicates the ability to detect patients who are suspected of disease but who do not have the condition (true-negative). Evaluation of diagnostic performance, therefore, requires independent assessment by the two methods in a population of patients who are suspected of disease but who do not all have the disease.

Evidence related to improvement of clinical outcomes with use of this testing assesses the data linking use of a test to changes in health outcomes (clinical utility). While in some cases, tests can be evaluated adequately using technical and diagnostic performance, when a test identifies a new or different group of patients with a disease; randomized trials are needed to demonstrate impact of the test on the net health outcome.

**ACCURACY/RELIABILITY**

Rathinam (2014) published a systematic review of studies of the reliability and validity of pediatric gait analysis tools.[4] Five observational gait tools were identified in nine studies of children with CP and one for children with Down’s syndrome. None of these observational gait tools the level of consistency found in instrumented gait analysis (IGA). While the Edinburgh Visual Gait Score (EVGS) was found to have better reliability and validity than the other observational tools, the limited studies available were insufficient to determine their impact on clinical outcomes.

A systematic review of 18 studies on gait classification systems was published in 2007.[5] The review included studies that involved classification of gait impairment based on kinematic, temporal-spatial kinetic, or electromyographic (EMG) data. Fifteen studies used three-dimensional gait analysis, one study used video observation analysis and 6 studies used EMG data. The authors assessed the overall methodological quality of the studies as low. Many studies appeared to classify patients arbitrarily rather than using clear clinical decision-making principles. Only two studies evaluated the reliability of classification, and the methods for determining the validity of classification systems were inadequate.

McGinley (2009) published a systematic review of studies of intersession and interassessor reliability of 3-dimensional kinematic gait analysis that included 15 full manuscripts and eight prepublication manuscripts.
abstracts. Similar to the 2007 systematic review summarized above, the authors noted variability in methodologic quality across the studies, but concluded that most studies demonstrated interassessor error of between 2 and 5 degrees of measurement, which the authors considered was “reasonable but may require consideration in data interpretation.”

Benedetti (2013) conducted an analysis of between-site consistency in gait analysis measurements of one healthy subject at seven different laboratories. The authors concluded that there was generally high concordance of segment and joint kinematics, except in the knee and the hip.

**IMPACT ON HEALTH OUTCOMES**

The ideal study design to demonstrate the clinical utility of gait analysis is a randomized controlled trial (RCT) comparing treatment decisions and health outcomes in patients managed with and without gait analysis.

Wren (2011) published a systematic review of literature on the efficacy of GA. The authors identified seven studies evaluating the effect of GA on patients’ health outcomes; none were RCTs. The studies addressed a variety of clinical conditions, so the authors were not able to pool findings. The systematic review also identified studies evaluating other aspects of GA including technical accuracy, diagnostic accuracy, and societal efficacy (i.e., impact on number and cost of procedures). The authors concluded that, although there is lower-level evidence (e.g., case series, case-control studies) supporting GA, there is a lack of evidence from RCTs on the effect of GA on health outcomes.

**SPECIFIC APPLICATIONS OF GAIT ANALYSIS**

In addition to the literature addressing gait analysis in general, several studies evaluate specific indications for GA.

**Pre- and/or Post-Surgical Evaluation for Children with Cerebral Palsy**

Two reports from one randomized controlled trial were published after the 2011 systematic review summarized above.

Wren and colleagues compared post-surgery health outcomes in children with cerebral palsy who were managed with and without gait analysis. This was a single-center, single-blind study. The trial included 186 ambulatory children with cerebral palsy who were candidates for lower extremity surgery to improve their gait. All participants underwent gait analysis at a gait laboratory. Patients were randomized to a treatment group in which the surgeon received the gait analysis report or a control group in which the surgeon did not receive the report. The reports included a summary of test results and treatment recommendations from the gait laboratory physician. The same surgeons treated the intervention and control patients i.e., they received gait reports for half of the patients. Patients were re-examined the day before surgery (i.e., following gait analysis) for pre-operative treatment planning. Outcomes were assessed pre-operatively and approximately 1 year post-surgery. There were three primary outcomes: 1) pre- to post-surgical change between groups in the walking scale of the Gillete Functional Assessment Questionnaire (FAQ), 2) the Gait Deviation Index (GDI), and 3) the oxygen cost of walking, a measure of the energy expended while walking.
A total of 156 of 186 (84%) participants returned for the follow-up examination; analysis was not intention to treat. There was no statistically significant difference between groups in any of the three primary outcomes. For example, the proportion of patients improved according to the FAQ was 31% in the intervention group and 25% in the control group (p=0.38). There were significant differences between groups at the p=0.05 level for two of 19 secondary outcome variables; p values were not adjusted for multiple comparisons. The authors noted that physicians followed only 42% of recommendations in the gait analysis report for patients in the treatment group, which may partially explain the lack of significant differences between groups in the primary outcomes and most of the secondary outcomes. They further noted that there was a positive relationship between gait outcomes and following gait analysis recommendations. Wren (2013) published a secondary analysis of data from the RCT previously described to evaluate the impact of gait analysis on the correction of excessive internal hip rotation among ambulatory children with cerebral palsy.\[10\] In the secondary analysis, the authors included the subset of children for whom the gait laboratory recommended external femoral derotation osteotomy (FDRO) to correct excessive passive and active internal hip rotation and who had both pre- and postoperative data available. As in the primary study, the intervention was receipt of the gait analysis report by the treating orthopedic surgeon for participants in the intervention group; in this subset of patients, all patients had had FDRO recommended by the gait analysis report, but the decision to actually perform surgery was up to the treating surgeon. Physical measurements for this subanalysis included femoral anteversion, maximum hip internal and external rotation range of motion, and rotational alignment during gait. The primary outcome variables included femoral anteversion and mean hip rotation and foot progression in the stance phase of gait. Outcomes postsurgery and change in variables pre- to postsurgery were compared between intervention and control groups, with additional analyses based on whether patients in the gait report (intervention) group had had the gait report recommendations followed.

This subanalysis included 44 children (65 limbs) in whom FDRO was recommended. FDRO was performed in 7/39 limbs in which it was recommended in the gait report (intervention group); it is not clear how many children in the control group for whom FDRO was recommended received surgery. There were no significant differences in outcomes between the gait report and control groups on intent-to-treat analysis. However, among children in the intervention group who had FDRO done (n=7 limbs), the limbs demonstrated greater improvements in femoral anteversion (-32.9° vs -12.2°; p=0.01), dynamic hip rotation (-25.5° vs -7.6°; p=0.001), and foot progression (-36.2° vs -12.4°; p=0.02) than limbs in the control group. The discrepancy between the intent-to-treat and per-protocol results may be related to generally poor compliance with the gait report recommendations, as only seven of 39 recommended FDROs performed in the gait analysis group. Interpretation of this study’s significance is limited by its subgroup analysis design and the small number of patients who received gait analysis and FDRO.

In 2013\[11\] and 2014\[12\], Schwartz et al. published two retrospective analyses to evaluate the role of a random forest algorithm in children with cerebral palsy using data from a motion analysis center database. The random forest algorithm was a statistical method used to predict an outcome for a particular observation based on a series of predictor values. The algorithm included gait analysis to predict outcomes after single-event, multilevel surgery for children with ambulatory cerebral palsy. The study authors reported that their random forest algorithm was able to generate criteria that were predictive of good outcomes for patients undergoing a single-event, multilevel orthopedic surgery. However, methodological limitations, such as the potential bias inherent in studies based on retrospective analysis of a motion analysis center...
database, make interpretation difficult. In addition, the complexity of the random forest decision algorithm makes it is difficult to determine the degree to which gait analysis independently predicts outcomes.

Pre- and/or Post-Surgical Evaluation for Conditions Other Than Cerebral Palsy

In a study by Suda (2002), gait analysis recommendations in 60 patients with neurogenic intermittent claudication were evaluated and compared with 50 healthy controls.[13] The authors concluded that gait analysis provided useful quantitative and objective information to evaluate postsurgical treatment. However, the study does not address how the gait analysis influenced treatment decisions or affected health outcomes.

Sankar (2009) reviewed the records of 35 children (56 feet) who had recurrent deformity after treatment of idiopathic clubfoot.[14] Gait lab recommendations were compared to surgical plans prior to gait analysis, and then to the actual surgery received. Thirty of 35 (86%) of children underwent surgery. GA resulted in changed procedures in 19 of 30 (63%) patients. GA was found to influence clinical decisions, but, like the study by Suda et al, this study did not evaluate whether these changes resulted in improved health outcomes.

Gait analysis has been used in the assessment of multiple other conditions (e.g., knee pain in older patients with osteoarthritis[15], gait after acute stroke[16], and of frailty in older patients[17]); however, the evidence linking the use of gait analysis to outcomes in these conditions is limited.

PRACTICE GUIDELINE SUMMARY

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE (NICE)

The National Institute for Health and Care Excellence (NICE) provided guidance in July 2012 for children and young people with spasticity who plan to have orthopedic surgery.[18] The NICE clinical guideline CG145 revised in 2016 states that “the decision to perform orthopaedic surgery to improve gait should be informed by a thorough pre-operative functional assessment, preferably including gait analysis”.

SUMMARY

Despite the lack of research, gait analysis has evolved to a standard of care for select surgical or other therapeutic interventions for children and adolescents with gait disorders associated with cerebral palsy. Therefore, gait analysis may be considered medically necessary in this population.

There is not enough research to show that gait analysis improves health outcomes for indications other than cerebral palsy. No clinical guidelines based on research recommend gait analysis for any other indication. Therefore, gait analysis is considered investigational for indications other than cerebral palsy.

REFERENCES


### CODES

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>96000</td>
<td>Comprehensive computer-based motion analysis by video-taping and 3-D kinematics</td>
</tr>
<tr>
<td></td>
<td>96001</td>
<td>Comprehensive computer-based motion analysis by video-taping and 3-D kinematics; with dynamic plantar pressure measurements during walking</td>
</tr>
<tr>
<td></td>
<td>96002</td>
<td>Dynamic surface electromyography, during walking or other functional activities, 1 to 12 muscles</td>
</tr>
<tr>
<td></td>
<td>96003</td>
<td>Dynamic fine wire electromyography, during walking or other functional activities, 1 muscle</td>
</tr>
<tr>
<td></td>
<td>96004</td>
<td>Review and interpretation by physician or other qualified health care professional of comprehensive computer-based motion analysis, dynamic plantar pressure measurements, dynamic surface electromyography during walking or other functional activities, and dynamic fine wire electromyography, with written report</td>
</tr>
<tr>
<td>HCPCS</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>

*Date of Origin: July 1998*
Hyperbaric Oxygen Pressurization (HBOT)

Effective: November 1, 2018

Next Review: September 2019
Last Review: September 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Hyperbaric oxygen therapy (HBOT) is a technique of delivering higher pressures of oxygen to the tissues. Two methods of administration are available, systemic and topical.

MEDICAL POLICY CRITERIA

I. Topical hyperbaric and topical normobaric oxygen therapies are considered investigational.

II. Systemic hyperbaric oxygen therapy

A. Systemic hyperbaric oxygen therapy (HBOT) services must comply with the following guidelines which are consistent with the Undersea and Hyperbaric Medical Society criteria:
   1. Patient must breathe 100% oxygen intermittently or continuously while the pressure of the treatment chamber is increased above one atmosphere absolute
   2. Systemic hyperbaric oxygen pressurization should be at least 1.4 atmospheres absolute (atm abs) (20.5 psi)
   3. Treatment is provided in a hospital or clinic setting.

B. Systemic hyperbaric oxygen pressurization (i.e., 100% oxygen delivered within a...
chamber at a pressure of at least 1.4 atm abs) may be considered medically necessary in the treatment of any of the following conditions:

1. Acute carbon monoxide poisoning (Recommended treatment review threshold: 5 treatments)
2. Acute traumatic ischemia (Recommended treatment review threshold: Reperfusion injury – 1 treatment; Crush injury – 12 treatments (3 times per day for 2 days, then twice a day for 2 days, then daily for 2 days); Compartment syndrome – 3 treatments (twice a day for 1 day, then 1 treatment on day 2)
3. Chronic refractory osteomyelitis (Recommended treatment review threshold: 40 treatments)
4. Cyanide poisoning, acute (Recommended treatment review threshold: 5 treatments)
5. Decompression sickness (Recommended treatment review threshold: 10 treatments)
6. Gas or air embolism, acute (Recommended treatment review threshold: 10 treatments)
7. Gas gangrene (i.e., clostridial myositis and myonecrosis; *Recommended treatment review threshold: 10 treatments)
8. Non-healing diabetic wounds of the lower extremities as an adjunct to ongoing conventional wound care in patients who meet all of the following 3 criteria (Recommended treatment review threshold: 30 treatments (one or two treatments daily):
   a. Patient has type I or type II diabetes and has a lower extremity wound that is due to diabetes
   b. Patient has a wound classified as Wagner grade 3 or higher (see Policy Guidelines)
   c. Patient has no measurable signs of healing after 30 days of an adequate course of standard wound therapy including all of the following:
      i. Assessment of vascular status and correction of any vascular problems in the affected limb if possible
      ii. Optimal glycemic control
      iii. Optimal nutritional status
      iv. Topical wound treatment (eg, saline, hydrogels, hydrocolloids, alginates) with maintenance of a clean, moist bed of granulation tissue
      v. Debridement to remove devitalized tissue, any technique
      vi. Pressure reduction or offloading
      vii. Treatment to resolve infection (e.g., antibiotics)

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
9. Pre- and post-treatment for patients undergoing dental surgery (non-implant-related) of an irradiated jaw

10. Profound anemia with exceptional blood loss: only when blood transfusion is impossible or must be delayed (Recommended treatment review threshold: HBOT should be continued with taper of both time and frequency until red blood cells have been satisfactorily replaced by patient regeneration or the patient can undergo transfusion.)

11. Soft-tissue radiation necrosis (e.g., radiation enteritis, cystitis, proctitis) and osteoradionecrosis (Recommended treatment review threshold for mandibular osteoradionecrosis: 60 treatments)

12. Idiopathic Sudden Sensorineural Hearing Loss of greater than or equal to 41 decibels and an onset of treatment within 14 days (Recommended treatment review threshold: 20 treatments.)

13. Necrotizing soft tissue infections
14. Actinomycosis
15. Central retinal artery occlusion

C. Hyperbaric oxygen pressurization is considered investigational for all other indications including but not limited to other ophthalmologic conditions, non-diabetic wounds, diabetic wounds with Wagner classification of grade 0-2, and acute thermal burns.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

POLICY GUIDELINES

It is critical that the list of information below is submitted for review to determine if the policy criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.

- History and physical/chart notes
- Indication for the requested service including type of HBOT planned
- Treatment plan including the following:
  - Percent of oxygen that the patient will breathe while receiving therapy
  - Pressurization (atm abs, psi)
  - Treatment setting
- Condition being treated including how many treatments being requested
  - If a diabetic wound is being treated then the request must include the following:
    - Type of diabetes
    - Location of wound
    - Wagner Classification
    - Measurable signs of healing following standard wound therapy including therapy length of time with documentation of the following:
      - Vascular assessment and correction, if possible, of vascular problems to affected area
      - Glycemic data for patient (e.g., A1C)
      - Nutritional status
— Topical wound treatments utilized including wound bed description
— Debridement
— Pressure reduction or offloading
— Any infection treatment utilized
  o If dental surgery, include description and diagnosis
  o If anemia, include blood loss and ability to transfuse patient
  o If necrosis, include type
  o If idiopathic sudden sensorineural hearing loss, include decibels of loss and onset of treatment

WAGNER CLASSIFICATION

- Grade 0: No open lesion
- Grade 1: Superficial ulcer without penetration to deeper layers
- Grade 2: Ulcer penetrates to tendon, bone, or joint
- Grade 3: Lesion has penetrated deeper than grade 2 and there is abscess, osteomyelitis, pyarthrosis, plantar space abscess, or infection of the tendon and tendon sheaths
- Grade 4: Wet or dry gangrene in the toes or forefoot
- Grade 5: Gangrene involves the whole foot or such a percentage that no local procedures are possible and amputation (at least at the below the knee level) is indicated

CROSS REFERENCES

None

BACKGROUND

SYSTEMIC HBOT

In systemic or large chamber hyperbaric oxygen, the patient is entirely enclosed in a pressure chamber and breathes oxygen at a pressure greater than 1 atmosphere (atm, the pressure of oxygen at sea level). Thus, this technique relies on systemic circulation to deliver highly oxygenated blood to the target site, typically a wound. In addition, systemic hyperbaric oxygen therapy can be used to treat systemic illness, such as air or gas embolism, carbon monoxide poisoning, clostridial gas gangrene, etc. Treatment may be carried out either in a monoplace chamber pressurized with pure oxygen or in a larger, multiplace chamber pressurized with compressed air, in which case the patient receives pure oxygen by mask, head tent, or endotracheal tube.

Mild Hyperbaric Oxygen Therapy

Oxygen therapy delivered via soft-sided chambers is referred to as mild hyperbaric oxygen therapy. While this implies that these chambers provide HBOT, the therapy is not considered hyperbaric as they provide pressurization of only about 4.5 psi, compared with true HBOT which is defined as pressurization of 20.5 psi or higher.

TOPICAL OXYGEN THERAPY

Topical Hyperbaric Oxygen Therapy

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.
Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Topical hyperbaric oxygen therapy is a technique of delivering 100% oxygen directly to an open, moist wound at a pressure slightly higher than atmospheric pressure. It is hypothesized that the high concentrations of oxygen diffuse directly into the wound to increase the local cellular oxygen tension, which in turn promotes wound healing. This therapy has been investigated as a treatment of skin ulcerations resulting from diabetes, venous stasis, postsurgical infection, gangrenous lesion, decubitus ulcers, amputations, skin graft, burns, or frostbite.

Topical hyperbaric oxygen devices consist of an appliance to enclose the wound area (frequently an extremity) and a source of oxygen; conventional oxygen tanks may be used. Topical hyperbaric oxygen therapy may be performed in the office, clinic, or may be self-administered by well-trained patients in the home. Typically, the therapy is offered for 90 minutes per day for 4 consecutive days. After a 3-day break, the cycle may be repeated. The regimen may last for 8 to 10 weeks.

**Topical Normobaric Oxygen Therapy**

Devices that deliver topical oxygen to a wound at normal atmospheric pressure (normobaric) are not considered hyperbaric oxygen therapy. These devices may also be called low dose tissue oxygenation systems. An example of a normobaric oxygen delivery system is the TransCu O2™, a small handheld device with an attached cannula. According to the manufacturer, the TransCu O2 is “intended for use with wound dressings to treat the following: skin ulcerations due to diabetes, venous stasis, post-surgical infections and gangrenous lesions; pressure ulcers; infected residual limbs; skin grafts; burns; and frostbite.” The device concentrates room air to 99.9% oxygen which is delivered via the cannula which is placed under the wound dressing.

**REGULATORY STATUS**

In 2013, U.S. Food and Drug Administration (FDA) published a statement warning that non-FDA approved uses of HBOT may endanger the health of patients.[1] “Patients may incorrectly believe that these devices have been proven safe and effective for uses not cleared by FDA, which may cause them to delay or forgo proven medical therapies. In doing so, they may experience a lack of improvement and/or worsening of their existing condition(s).”

The following are examples of oxygen therapy devices:

In February 1999, the Numobag™ Kit (Numotech, Inc) for application of topical hyperbaric therapy was cleared for marketing by the FDA through the 510(k) process. The FDA determined that this device was substantially equivalent to existing devices. Another example is the AOTI Hyper-Box™ (AOTI Ltd., Galway, Ireland) which was cleared by FDA in 2008.

In August 2009, the TransCu O2 (Electrochemical Oxygen Concepts, Inc.) was cleared for marketing by the FDA through the 510(k) process as substantially equivalent to existing devices.

There are numerous FDA-approved hyperbaric oxygen chambers. In May 2005, the ATA Monoplace Hyperbaric System (ATA Hyperbaric Chamber Manufacturing, Inc.) was cleared for marketing by the FDA through the 510(k) process. The FDA determined that this device was substantially equivalent to existing hyperbaric devices.

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
EVIDENCE SUMMARY

Current evidence is sufficient to determine the effectiveness of hyperbaric oxygen therapy (HBOT) for the indications that meet the above medical necessity criteria. Assessing the effectiveness and safety of HBOT for the investigational indications requires randomized controlled trials comparing HBOT with the conventional treatments for each indication. Therefore, the following literature review on HBOT focuses on randomized controlled trials (RCTs) and systematic reviews of RCTs for the investigational indications.

Assessment of efficacy for therapeutic interventions involves a determination of whether the intervention improves health outcomes. The optimal study design for a therapeutic intervention is a randomized controlled trial (RCT) that includes clinically relevant measures of health outcomes. Intermediate outcome measures, also known as surrogate outcome measures, may also be adequate if there is an established link between the intermediate outcome and true health outcomes. When the primary outcomes are subjective (e.g., pain, depression), sham-controlled RCTs are needed to assess the effect of the intervention beyond that of a placebo effect.

TOPICAL HYPERBARIC OXYGEN

Due to their different methods of delivery, topical and systemic hyperbaric oxygen are distinct technologies such that they must be examined separately. There is minimal published literature regarding topical hyperbaric oxygen therapy. A 2015 Cochrane review of interventions for treating gas gangrene evaluated the safety and efficacy topical HBOT and Chinese herbs as treatments options. Re-analysis if cure rate did not show beneficial effects from either treatment. In 1984, Heng and colleagues published a controlled study of topical hyperbaric oxygen therapy in 6 patients with 27 ulcers compared to no treatment in 5 patients with 10 ulcers. Although a greater improvement was noted in the treated group, the results were calculated according to the number of ulcers rather than based on individual patients. Leslie and colleagues reported on a trial that randomly assigned 18 patients with diabetic foot ulcers to receive either topical hyperbaric oxygen therapy plus standard wound care or standard wound care alone. Changes in ulcer size and depth did not differ between the 2 groups. Other studies consist of anecdotal reports or uncontrolled case series.

SYSTEMIC HYPERBARIC OXYGEN THERAPY (HBOT)

In-home Hyperbaric Oxygen

A position statement from the National Board of Diving & Hyperbaric Medical Technology on in-home HBOT has been published on the Web site for The Undersea and Hyperbaric Medicine Society (UHMS). The statement indicates that in-home HBOT “is inherently unsafe and cannot be condoned.” This position is based on concern for the safety and well-being of patients as well as those people in proximity to the HBOT delivery system because in-home provision of HBOT is likely to:

- Bypass otherwise mandatory federal, state, and local codes related to design, construction, installation, and operation of these devices; and
- Occur without adequate physician oversight and the operational support of appropriately qualified HBOT providers.

Acute Coronary Syndromes
Systematic Reviews

A 2012 Cochrane review by Bennett and colleagues identified 6 trials with a total of 665 patients evaluating HBO for acute coronary syndrome.[8] All of the studies included patients with acute myocardial infarction (MI); one study also included individuals presenting with unstable angina. Additionally, all trials used HBOT as an adjunct to standard care. Control interventions varied; only 1 trial described using a sham therapy to blind participants to treatment group allocation. In a pooled analysis of data from 5 trials, there was a significantly lower rate of death in patients who received HBOT compared to a control intervention (RR: 0.58; 0.36 to 0.92). Due to variability of outcome reporting in the studies, few other pooled analyses could be conducted. A pooled analysis of data from 3 trials on improvements in left ventricular function did not find a statistically significant benefit of HBOT (RR: 0.09; 95% CI: 0.01 to 1.4). The authors noted that, although there is some evidence from small trials that HBOT is associated with a lower risk of death, larger trials with high methodologic quality are needed in order to determine which patients, if any, can be expected to derive benefit from HBOT. Therefore, HBOT is considered investigational in the treatment of acute coronary syndromes.

Autism Spectrum Disorders (ASD)

Systematic Reviews

A 2016 systematic review on hyperbaric oxygen therapy for treatment of children with autism identified one RCT[9] with a total of 60 children. The study quality was rated as low using GRADE criteria with small sample size and wide confidence intervals. The results indicated no improvement in social interaction and communication, behavioral problems, communication and linguistic abilities, or cognitive function. The authors reported minor-grade ear barotrauma as adverse events.

A 2012 systematic review[10] of RCTs on hyperbaric oxygen therapy for treatment of children with autism identified two RCTs[11,12] with a total of 89 participants. In both RCTs the active hyperbaric treatment was 24% oxygen delivered at an atmospheric pressure of 1.3 atmospheres (atm). Although this regimen was referred to as HBOT in the article, it differed from standard HBOT which uses 100% oxygen and a pressure of at least 1.4 atm. A detailed analysis of these RCTs is provided below. Briefly, one of the two RCTs found better outcomes after hyperbaric oxygen compared with placebo treatment, and the other did not find significant differences in outcomes. The author concluded that additional sham-controlled trials with rigorous methodology are needed in order to draw conclusions about the efficacy of HBOT for treating autism.

Randomized Controlled Trials (RCTs)

The following is a summary of the 2 RCTs reported in the above systematic review:

- One of the above two RCTs was by Rossignol and colleagues.[11] This study was a double-blind RCT that included 62 children, ages 2-7, meeting DSM-IV criteria for autistic disorder. The active treatment was hyperbaric treatment at 1.3 atmospheres (atm) and 24% oxygen in a hyperbaric chamber. (This regimen differs from standard HBOT which uses 100% oxygen and a pressure of at least 1.4 atm). The other group received a sham treatment consisting of 1.03 atm and ambient air (21% oxygen). Both groups received 40 sessions of active or sham treatment lasting 60 minutes each over a...
period of 4 weeks. The equipment, procedures, etc. in the two groups were as similar as possible to maintain blinding. The investigators, participants, parents, and clinic staff were blinded to treatment group. Only the hyperbaric technician, who had no role in outcome assessment, was aware of group assignment. After completion of the 4-week study, families with children in the control group were offered the active intervention. When asked at the end of the study, there was no significant difference in the ability of parents to correctly guess the group assignment of their child.

The outcomes were change compared to baseline after 4 weeks on the following scales: Aberrant Behavior Checklist (ABC) total score and 5 subscales; Autism Treatment Evaluation Checklist (ATEC) total score and 4 subscales; and Clinical Global Impression-Improvement (CGI) overall functioning score and 18 subscales. P values of <0.05 were considered statistically significant; there was no adjustment for multiple comparisons. The analysis included all children who completed at least one complete session. Of the 33 children assigned to active treatment, 30 were included in the analysis and 29 completed all 40 treatments. Of the 29 children assigned to the control treatment, 26 completed all 40 sessions and were included in the analysis.

There was no significant between-group improvement on the ABC total score, any of the ABC subscales, or on the ATEC total score. Compared to the control group, the treatment group had a significant improvement in 1 of 4 subscales of the ATEC, the sensory/cognitive awareness subscale. The change from baseline on this subscale was a mean of 16.5 in the treatment group and a mean of 5.4 in the control group, a difference of 11.1 (p=0.037). (Note: due to an administrative error, baseline ATEC was not collected at one site, and thus data were not available for 23 children in the treatment group and 21 children in the control group). On the physician-rated CGI total score, 9/30 (30%) children in the treatment group had a score of 1 (very much improved) or 2 (much improved) compared to 2/26 (8%) in the control group (p=0.047). On the parental-rated CGI total score, 9/30 (30%) children in the treatment group had a score of 1 or 2 compared to 4/26 (15%) in the control group (p=0.22, not statistically significant). (The exact numbers receiving scores of 1 vs. 2 were not reported). Change in mean CGI scores were also reported, but this may be a less appropriate way to analyze these data. Among the parental-rated CGI subscales, significantly more children were rated as improved in the treatment group compared to control on 2 out of 18 subscales, receptive language (p=0.017) and eye contact (p=0.032).

A key limitation of this study was that the authors reported only outcomes at 4 weeks, directly after completion of the intervention. It is not known whether there are any long-term effects. Additional follow-up data cannot be obtained because members of the control group crossed over to the intervention after 4 weeks. Other limitations included lack of adjustment for multiple comparisons and unclear clinical significance of the statistically significant outcomes. The Undersea and Hyperbaric Medical Society (UHMS) issued a position paper after publication of the Rossignol et al. study stating that they still did not recommend routine treatment of autism with HBOT.\[13\]

- The other RCT included in the systematic review was a double-blind RCT that began with 46 children with autism, ages 2-14 years, who were matched in pairs according to age and the number of hours of Applied Behavior Analysis (ABA) treatment they were receiving at the start of the study. Randomized\[12\] treatment allocation of the matched
pairs was by coin toss. Both groups received 80 1-hour sessions of active treatment (24% oxygen at 1.3 atm) or sham treatment (room air at ambient pressure) for up to 15 weeks. Participants were allowed to undergo ABA, take any supplements, pharmacological interventions, and dietary modifications. Twelve patients withdrew from the trial, leaving 18 patients in the treatment group and 16 in the control group.

The primary outcome of change in symptoms was based on direct observation and the scales noted in the Rossignol et al. study above in addition to the Behavior Rating Inventory of Executive Functioning (BRIEF), Parent Stress Index (PSI), Peabody Picture Vocabulary Test (PPVT-III), Repetitive Behavior Scale (RBS), and the Vineland Adaptive Behavior Scales (VABS-II). Direct observation and intention to treat analysis of test scores found no significant difference on any outcome measures between the treatment and sham groups. No participants experienced adverse effects attributable to barotrauma (e.g., pressure injury to tympanic membranes or sinuses).

A limitation of this study was the small sample size which was determined to be adequate to detect only large effects, which were not present in this study. In addition, since some patients in both groups received intensive ABA interventions during the study period, any potential effects of HBOT could not be isolated. The authors concluded that the active treatment had no significant beneficial effect on ASD and was not recommended for the treatment of ASD symptoms.

One additional RCT not included in the systematic review above was identified:

A 2012 RCT published after the systematic review randomly assigned 60 children with autism to receive 20 one-hour sessions with HBOT or sham air treatment (n=30 per group). The primary outcome measures were change in the ATEC and CGI, evaluated separately by clinicians and parents. There were no statistically significant differences between groups on any of the primary outcomes. For example, post-treatment clinician-assessed mean scores on the ATEC were 52.4 in the HBOT group and 52.9 in the sham air group.

Conclusion

There is insufficient evidence from well-designed RCTs that HBOT improves health outcomes for patients with autism spectrum disorder; therefore, HBOT therapy for this indication is considered investigational.

Bell’s Palsy

Systematic Review

In 2012, Holland and colleagues published a Cochrane review evaluating HBOT in adults with Bell’s palsy. The authors identified one RCT with 79 participants, and this study did not meet the Cochrane review methodologic standards because the outcome assessor was not blinded to treatment allocation. Therefore, the evidence is insufficient to permit conclusions and HBOT is considered investigational for the treatment of Bell’s palsy.

Randomized Controlled Trials (RCTs)

No RCTs have been published since the 2012 Cochrane review.
Bisphosphonate-related Osteonecrosis of the Jaw (BRONJ)

Randomized Controlled Trials (RCTs)

An unblinded RCT was published by Freiberger and colleagues in 2012 on use of HBOT as an adjunct therapy for patients with bisphosphonate-related osteonecrosis of the jaw.\[16\] Forty-nine patients were randomly assigned to HBOT in addition to standard care (n=22) or standard care alone (n=27). Five patients in the standard care group received HBOT and 1 patient assigned to the HBOT group declined HBOT. The investigators decided to do a per protocol (PP) analysis (actual treatment received) because of the relatively large degree of crossover. Participants were evaluated at 3, 6 12 and 18 months. Data were available on 46 patients, 25 received HBOT in addition to standard care and 21 received standard care alone. The primary outcome measure was change in oral lesion size or number. When change from baseline to last available follow-up was examined, 17 of 25 (68%) of HBO-treated patients had improvement in oral lesion size or number compared to 8 of 21 (38%) in the standard care group, p=0.043. When change from baseline to 6, 12 or 18 months was examined, there was not a statistically significant difference between groups in the proportion of patients with improvement. In addition, the proportion of patients who healed completely did not differ significantly between groups at any time point. This single trial does not report consistent findings of benefit across outcome measures. It also has a number of methodologic limitations, e.g., unblinded, cross-over, and analysis performed on a per-protocol basis rather than intention to treat. A disadvantage of the per-protocol analysis is that randomization is not preserved, and the two groups may differ on characteristics that affect outcomes. As a result, this trial is insufficient to conclude that HBOT improves health outcomes for patients with bisphosphonate-related osteonecrosis of the jaw.

Conclusion

Current evidence is insufficient to determine the safety and efficacy of HBOT in the treatment of bisphosphonate-related osteonecrosis of the jaw. Therefore, HBOT is considered investigational for this indication.

Cancer Treatment

Randomized Controlled Trials (RCTs)

In an RCT of 32 patients, Heys and colleagues found no increase in 5-year survival in patients treated with HBOT prior to chemotherapy for locally advanced breast carcinoma to increase tumor vascularity.\[17\] This approach is being studied since studies in animal models have suggested that HBOT increases tumor vascularity and thus may make chemotherapy more effective. In a Cochrane review, Bennett and colleagues concluded that HBOT given with radiotherapy may be useful in tumor control; however, the authors expressed caution since significant adverse effects were common with HBOT and indicated further study would be useful.\[18\]

Conclusion

Current evidence is insufficient to determine the safety and efficacy of HBOT in the treatment of cancer of any type and location. Therefore, HBOT is considered investigational for this indication.

Cerebral Palsy
Randomized Controlled Trials (RCTs)

- In 2012, Lacey and colleagues published a double-blind RCT that included 49 children age 3-8 years with spastic cerebral palsy. Participants were randomized to receive 40 treatments with either HBOT (n=25) or hyperbaric air to simulate 21% oxygen at room air (n=24). The primary efficacy outcome was change in the Gross Motor Function Measure (GMFM-88) global score after the 8-week treatment period. The study was stopped early due to futility, when an interim analysis indicated that there was less than a 2% likelihood that a statistically significant difference between groups would be found. At the time of the interim analysis, there was no significant between-group difference in the post-treatment GMFM-88 global score (p=0.54).

- In the largest RCT to date, Collet et al. randomly assigned 111 children with cerebral palsy to 40 treatments over a 2-month period of either HBOT (n=57) or slightly pressurized room air (n=54). The authors found HBOT and slightly pressurized air produced similar improvements in both groups for outcomes such as gross motor function and activities of daily living.

Conclusion

HBOT is considered investigational as a treatment for cerebral palsy because it has not been shown to provide additional health benefits in this patient population.

Compromised Skin Grafts and Flaps

Systematic Reviews

- In a 2010 Cochrane review, Estes and colleagues found a lack of high quality evidence regarding HBOT in the treatment of skin grafts and flaps. The authors found one randomized controlled trial (RCT) on skin grafts for burn wounds (n=48) which reported significantly higher graft survival with HBOT, and one RCT on flap grafting (n=135) which reported no significant differences in graft survival with HBOT compared with dexamethasone or heparin. However, these data are unreliable due to various methodologic limitations such as biased analysis, omitted data, and small size.

- In 2006, Friedman and colleagues published a systematic review of literature on use of HBOT for treating skin flaps and grafts. No RCTs were found. The authors identified 2 retrospective case series on use of HBOT for clinically compromised skin grafts and flaps. The series had sample sizes of 65 and 26, respectively; both were published in the 1980s based on treatment provided in the 1970s and 1980s.

Randomized Controlled Trials (RCTs)

No RCTs have been published since the above systematic reviews.

Conclusion

Although the study of HBOT for compromised skin grafts and flaps goes back several decades, the clinical trial data is limited to noncomparative case series and a single randomized controlled trial. This evidence is insufficient to determine the safety and efficacy of HBOT in the treatment of compromised skin grafts and flaps. Therefore, HBOT is considered investigational for these indications.
Carbon Monoxide Poisoning

A 2011 Cochrane review of 7 RCTs concluded that the available evidence is insufficient to determine whether adverse neurologic outcomes in patients with carbon monoxide poisoning are reduced with HBOT.[24] In 2008, the American College of Emergency Physicians (ACEP) published a clinical policy on critical issues in carbon monoxide poisoning.[25] Their literature review indicated there was only level C evidence (preliminary, inconclusive, or conflicting evidence) for treatment of acute carbon monoxide poisoning. The 2008 UHMS guidelines, however, list carbon monoxide poisoning as an indication for HBOT.

Two blinded randomized trials were discussed in both the Cochrane and ACEP reviews. One is a study by Scheinkestel et al, a double-blind, RCT comparing HBOT with normobaric oxygen in patients with carbon monoxide poisoning.[26] The authors reported that HBOT did not benefit patient outcomes of neuropsychologic performance when HBOT was completed and at 1-month follow-up. This study was limited, however, by a high rate (46%) of patients who were lost to follow-up. Moreover, the trial has been criticized for administering 100% normobaric oxygen for at least 72 hours between treatments, which has been called a toxic dose of oxygen.[27] The critiques also mention that there was an unusually high rate of neurologic sequelae after the treatment period, which could be due in part to the high dose of oxygen and/or the high rate of cognitive dysfunction in the study population (69% were poisoned by carbon monoxide through suicide attempts).

The other blinded trial, by Weaver et al, also compared hyperbaric and normobaric oxygen.[28] Patients received either 3 sessions of HBOT or 1 session of normobaric oxygen plus 2 sessions of exposure to normobaric room air. The primary outcome was the rate of cognitive sequelae at 6 weeks. Cognitive function was assessed using a battery of neuropsychological tests. At the 6-week follow-up, the intention-to-treat analysis found that 19 of 76 (25.0%) in the HBOT group and 35 of 76 (46.1%) in the control group had cognitive sequelae; the difference was statistically significant (p=0.007). There was a high rate of follow-up at 6 weeks, 147 of 152 (97%) of randomized patients. Enrollment in the study was stopped early because an interim analysis found HBOT to be effective. A follow-up study, which included 147 patients from the randomized trial and 75 who had been eligible for the trial but had not enrolled, was published in 2007.[29] Of the group treated with HBOT (n=75), cognitive sequelae were identified in 10 of 58 (17%) at 6 months and 9 of 62 (14%) at 12 months. Of the group not treated with HBOT (n=163), 44 of 146 (30%) at 6 months and 27 of 149 (18%) at 12 months had cognitive sequelae. (The follow-up rate was higher at 12 months because the investigators received additional funding for data collection.)

Delayed-Onset Muscle Soreness

Systematic Review

In a 2005 Cochrane review, Bennett and colleagues concluded that available evidence is insufficient to demonstrate beneficial outcomes with HBOT for delayed-onset muscle soreness and closed soft-tissue injury.[30] It was noted that HBOT possibly even increases pain initially and further studies are needed. Therefore, use of HBOT for this indication is considered investigational.

Randomized Controlled Trials (RCTs)

No RCTs have been published since the 2005 Cochrane review.
Dementia

Systematic Review

A 2012 Cochrane review identified 1 RCT evaluating HBOT for the treatment of vascular dementia.[31] The 2009 study compared HBOT plus donepezil to donepezil-only in 64 patients. The HBOT and donepezil group had significantly better cognitive function after 12 weeks of treatment, as assessed by the Mini-Mental State Examination. However, the Cochrane investigators judged the trial to be of poor methodologic quality because it was not blinded and the methods of randomization and allocation concealment were not discussed.

Randomized Controlled Trials (RCTs)

No RCTs have been published since the 2012 Cochrane review.

Conclusion

The current evidence for HBOT as a treatment of dementias of any cause is limited to a single short-term clinical trial on vascular dementia. This evidence is insufficient to permit conclusions about the safety and efficacy of HBOT on vascular dementia. No other randomized controlled trials were found for HBOT as a treatment of demential from any cause. Due to the lack of sufficient evidence, HBOT is considered investigational for treatment of dementias.

Femoral Neck Necrosis, Idiopathic

Randomized Controlled Trials (RCTs)

In 2010, Camporesi and colleagues published the results of a double-blind RCT that evaluated HBOT in 20 adult patients with idiopathic unilateral femoral head necrosis.[32] Patients received 30 treatments over 6 weeks with either HBOT at 2.5 ATA (n=10) or a sham treatment consisting of hyperbaric air (n=10). The mean severity of pain on a 0-to-10 scale was significantly lower in the HBOT group than the control group after 30 sessions (p<0.001) but not after 10 or 20 sessions. (The article did not report exact pain scores). Several range-of-motion outcomes were also reported. At the end of the initial treatment period, extension, abduction and adduction, but not flexion, were significantly greater in the HBOT group compared to the control group. Longer-term comparative data were not available because the control group was offered HBOT at the end of the initial 6-week treatment period.

Conclusion

The current evidence is limited to a single, small short-term RCT. Thus, there is insufficient data on which to draw conclusions about the efficacy of HBOT for treating femoral head necrosis, and it is considered investigational for this indication.

Fibromyalgia

One quasi-randomized trial and 1 delayed-treatment RCT on HBOT for fibromyalgia were identified. In 2004, a study by Yildiz et al included 50 patients with fibromyalgia who had ongoing symptoms despite medical and physical therapy.[33] On an alternating basis, patients were assigned to HBOT or a control group. The HBOT consisted of fifteen 90-minute sessions at 2.4 ata (1 session per day, 5 d/wk). The control group breathed room air at 1 ata on the same schedule. Baseline values on the 3 outcomes were similar in the 2 groups. After the
course of HBOT treatment, the mean (SD) number of tender points were 6.04 (1.18) in the HBO group and 12.54 (1.10) in the control group. The mean (SD) pain threshold was 1.33 kg (0.12) and 0.84 kg (0.12), respectively, and the mean VAS was 31.54 (8.34) and 55.42 (6.58), respectively. In the study abstract, the authors stated that there were statistically significant differences between the HBO and control groups after 15 therapy sessions, but the table presenting outcomes lacked the notation used to indicate between-group statistical significance. It is not clear whether the control group actually received a sham intervention that would minimize any placebo effect ie whether or not the control intervention was delivered in a hyperbaric chamber. The authors stated that the study was double-blind but did not specify any details of patient blinding.

In 2015, Efrati et al published an RCT that included 60 female patients who had fibromyalgia for at least 2 years and were symptomatic. Patients were randomized to an immediate 2 month course of HBOT or delayed HBOT after 2 months. The HBOT protocol was forty 90-minute sessions of 100% oxygen at 2 ata (1 session per day, 5 d/wk). Forty-eight of 60 patients (80%) completed the study and were included in the analysis. After the initial 2 months, outcomes including number of tender points, pain threshold, and quality of life (SF-36) were significantly better in the immediate treatment group compared with the delayed treatment group (which received no specific intervention during this time). After the delayed treatment group had undergone HBOT, outcomes were significantly improved compared with scores prior to HBOT treatment. These findings are consistent with a clinical benefit of HBOT, but also with a placebo effect. A sham-control is needed to confirm the efficacy of HBOT in the treatment of fibromyalgia and other conditions where primary end points are pain and other subjective outcomes.

The above studies were few in number with relatively small sample sizes and had methodological limitations, e.g., quasi-randomization and no or uncertain sham control for a condition with subjective outcomes susceptible to a placebo effect. Moreover, the HBO protocol varied (e.g., 15 HBOT sessions vs 40 HBOT sessions). Thus, the evidence is insufficient to draw conclusions about the impact of HBOT on health outcomes for patients with fibromyalgia.

**Fracture Healing**

**Systematic Review**

In 2012, Bennett and colleagues published a Cochrane review on HBOT to promote fracture healing and treat non-union fractures. The investigators did not identify any published RCTs on this topic that compared HBOT to no treatment, sham treatment, or another intervention and reported bony union as an outcome.

**Randomized Controlled Trials (RCTs)**

No RCTs have been published since the 2012 Cochrane review.

**Conclusion**

Due to the lack of RCTs, it is not possible to conclude whether the use HBOT to promote fracture healing improves outcomes; therefore, the use of HBOT for this indication is considered investigational.

**Headaches**

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
When assessing any treatment focused on pain relief, randomized, placebo-controlled trials are necessary to investigate the extent of any placebo effect and to determine whether any improvement with the treatment exceeds that associated with a placebo.

The following is a summary of the available evidence:

**Migraine headaches**

- **Systematic Review**

  A 2008 Cochrane review by Bennett and colleagues identified RCTs that evaluated the effectiveness of systemic hyperbaric oxygen therapy for preventing or treating migraine headache compared to another treatment or a sham control.\(^{[36]}\) Five trials with a total of 103 patients were identified that addressed treatment of acute migraine with HBOT. A pooled analysis of 3 trials (total of 43 patients) found a statistically significant increase in the proportion of patients with substantial relief of migraine within 45 minutes of HBOT (relative risk [RR] 5.97, 95% confidence interval [CI]1.46-24.38, p=0.001). No other pooled analyses were conducted due to variability in the outcomes reported in the trials. The meta-analysis did not report data on treatment effectiveness beyond the immediate post-treatment period, and the methodologic quality of trials was moderate to low, e.g., randomization was not well-described in any trial. There was no evidence that HBOT could prevent episodes of migraine headache.

- **Randomized Controlled Trials (RCTs)**

  In 2004 Eftedal and colleagues reported the results of a randomized, double-blind, placebo-controlled trial to assess whether HBOT had a prophylactic effect on migraine headache.\(^{[37]}\) Forty patients were randomly assigned to either a treatment group receiving 3 sessions of HBOT or a control group receiving 3 hyperbaric treatments with room air. Thirty-four patients completed the study. Efficacy was measured as the difference between pre- and post-treatment hours of headache per week. There was no significant reduction in hours of headache with HBOT compared with hyperbaric air treatments. Nor was there a significant difference in either group in pre- and post-treatment levels of endothelin-1 in venous blood. The authors concluded that that HBOT had no significant prophylactic effect on migraine headache or on the endothelin-1 level in venous blood.

**Cluster headaches**

- **Systematic Reviews**

  Two 2008 systematic reviews, including the Cochrane review noted above, reported few studies comparing HBOT with sham treatment for cluster headaches.\(^{[36,38]}\) Available randomized, placebo-controlled trials measuring effect on symptoms are unreliable due to very small size.\(^{[39,40]}\)

- **Randomized Controlled Trials (RCTs)**

  No RCTs have been published since the 2008 systematic reviews.
• Conclusion

Due to the lack of sufficient evidence from well-designed clinical trial, HBOT for the treatment of headaches from any cause is considered investigational.

**Herpes Zoster**

**Randomized Controlled Trial (RCT)**

In 2012, Peng and colleagues published an RCT evaluating HBOT as a treatment of herpes zoster.[41] Sixty-eight patients with herpes zoster diagnosed within the previous 2 weeks were randomized to 30 sessions of HBOT (n=36) or medication treatment (n=32). Pharmacotherapy included antiviral, pain, nerve nutritive and antidepressive medication. Therapeutic efficacy was calculated at the end of the 3-week treatment period and included the proportion of patients who were healed (i.e., complete subsidence of pain and rash) or improved (i.e., significant pain relief and rash subsistence). Rates of therapeutic efficacy were 97.2% in the HBOT group and 81.3% in the medication group (p<0.05). Limitations of the study included a lack of blinding and lack of long-term follow-up.

**Conclusion**

The evidence from the single randomized controlled trial is insufficient to permit conclusions about the effect of HBOT on health outcomes for patients with herpes zoster; therefore, HBOT is considered investigational for this indication.

**Inflammatory Bowel Disease**

**Systematic Reviews**

A 2014 systematic review by Dulai et al examined the evidence on HBOT for inflammatory bowel disease (Crohn disease and ulcerative colitis).[42] The review was not limited by study design. The authors included 17 studies: 1 RCT, 2 case-control studies, 3 case series, and 11 case reports. The studies reported on a total of 613 patients, 286 with Crohn disease and 327 with ulcerative colitis. The only RCT identified was published in 2013; it was open-label and included 18 patients with ulcerative colitis.[43] Patients were randomized to treatment with standard medical therapy only (n=8) or medical therapy plus HBOT (n=10) consisting of 90-minute treatments at 2.4 atm, 5 days a week for 6 weeks (total of 30 sessions). The primary outcome was the self-reported Mayo score which has a potential range of 0 to 12.[44] Patients with a score of 6 or more are considered to have moderate to severe active disease. At 6 months follow-up there was no significant difference between groups in the Mayo score, with a median score of 0.5 in the HBOT group and 3 in the control group (exact p value not reported). In addition, there were no significant differences in any of the secondary outcomes including laboratory tests and fecal weight. Overall, the authors found that the studies had a high risk of bias, particularly in the areas of attrition and reporting bias, and further study in well-controlled, blinded RCTs was recommended.

**Randomized Controlled Trials (RCTs)**

No RCTs have been published since the 2014 systematic review.

**Conclusion**
There is insufficient evidence that HBOT is effective for treating inflammatory bowel disease. Only 1 small RCT has been published, and this study did not find a significant improvement in health outcomes when HBOT was added to standard medical therapy.

**In Vitro Fertilization**

In a 2005 nonrandomized pilot study, Van Voorhis and colleagues reported that HBOT was well tolerated in women undergoing ovarian follicular stimulation for in vitro fertilization; however no outcomes were reported.[45] Therefore, current evidence is insufficient to permit conclusions and HBOT is considered investigational for this indication.

**Mental Illness**

A Rapid Response Report from the Canadian Agency for Drugs and Technologies in Health (CADTH) searched the literature through July 2014 on the clinical effectiveness of hyperbaric oxygen therapy for treatment of adults with posttraumatic stress disorder, generalized anxiety disorder, and/or depression.[46]

The review’s inclusion criteria were health technology assessments, systematic reviews, meta-analyses, RCTs or nonrandomized studies comparing HBOT to any active treatment and reporting clinical outcomes. No eligible studies were identified.

**Multiple Sclerosis**

A Cochrane review of RCTs on HBOT for multiple sclerosis was published by Bennett et al in 2004.[47] The authors identified 9 RCTs, with a total of 504 participants that compared the effects of HBOT with placebo or no treatment. The primary outcome of the review was score on the Expanded Disability Status Scale (EDSS). A pooled analysis of data from 5 trials (N=271) did not find a significant difference in change in the mean EDSS after 20 HBOT treatments versus control (mean difference [MD], -0.07; 95% CI, -0.23 to 0.09). Moreover, a pooled analysis of data from 3 trials (n=163) comparing HBOT and placebo did not find a significant difference in mean EDSS after 6 months of follow-up (MD = -0.22; 95% CI, -0.54 to 0.09).

**Osteomyelitis**

No prospective clinical trials on chronic refractory osteomyelitis or acute refractory osteomyelitis were identified in updated searches. The justification for the use of HBOT in chronic osteomyelitis has been primarily based on case series. Among the larger case series, Maynor et al reviewed the records of all patients with chronic osteomyelitis of the tibia seen at one institution.[48] Follow-up data were available on 34 patients who had received a mean of 35 adjunctive HBO treatments (range, 6-99). Of the 26 patients with at least 2 years of follow-up after treatment, 21 (81%) remained drainage-free. Twelve of 15 (80%) with follow-up data at 60 months had remained drainage-free. A study by Davis et al reviewed outcomes for 38 patients with chronic refractory osteomyelitis treated at another U.S. institution.[49] Patients received HBOT until the bone was fully recovered with healthy vascular tissue; this resulted in a mean of 48 daily treatments (range, 8-103). After a mean posttreatment follow-up of 34 months, 34 of 38 (89%) patients remained clinically free of infection (i.e., drainage-free and no tenderness, pain, or cellulitis). Success rates from several smaller case series, all conducted in Taiwan, are 12 of 13 (92%) patients, 11 of 14 (79%) patients, and 13 of 15 (86%) patients.[50-52] A high percentage of refractory patients in these series had successful outcomes.

**Radiotherapy Adverse Effects**
Systematic Review

- A 2017 systematic review on the effectiveness of HBOT for the treatment of radiation-induced skin necrosis included eight articles with five case series studies, two case reports, and one observational cohort. The authors investigated the change in symptoms and alteration in wound healing and reported that HBOT was a safe intervention with promising outcomes. However, the authors recommended additional high quality evidence in order for HBOT to be considered as a relevant treatment for this indication.

- A 2014 systematic review on the safety and effectiveness of HBOT for the treatment of non-neurological soft tissue radiation-related injuries (STRI) included 41 articles, 11 of which compared regimens with and without HBOT. Serious adverse effects were rare and the more common adverse effects were minor and self-limiting. Evidence of a beneficial effect of HBOT was reported radiation proctitis and STRI of the head and neck, but not for post-radiation soft tissue edema or radiation cystitis. The authors recommended further studies to validate the use of HBOT as both a definitive and adjunctive treatment for individual STRI.

- In 2010, Spiegelberg and colleagues conducted a systematic review of studies on HBOT to prevent or treat radiotherapy-induced head and neck injuries associated with treatment of malignant tumors. The authors identified 20 studies. Eight of the studies included control groups; their sample sizes ranged from 19 to 78 individuals. Four (50%) of the studies with a control group concluded that HBOT was effective, and the other 4 did not conclude that the HBOT was effective. The authors noted a paucity of RCTs but did not state the number of RCTs identified in their review.

Randomized Controlled Trials

- Teguh and colleagues reported on 17 patients with oropharyngeal or nasopharyngeal cancer who were treated with radiation therapy. Eight patients were randomly assigned to receive 30 sessions of HBOT, beginning within 2 days of completing radiation therapy, and 9 patients received no additional treatment. All patients were included in the analysis. Quality of life outcomes were assessed and the primary outcome was specified as xerostomia at 1 year. Quality of life measures did not differ significantly between groups in the acute phase (first 3 months). For example, 1 month after treatment, the mean visual analog scale (VAS) score for xerostomia (0-to-10 scale) was 5 in the HBOT group and 6 in the control group. However, at 1 year, there was a statistically significant difference between groups; the mean VAS score for xerostomia was 4 in the HBOT group and 6 in the control group (p=0.002). Also at 1 year, the mean quality of life score for swallowing (0-to-100 scale) was 7 in the HBOT group and 40 in the control group (p=0.0001). The study is limited by the small sample size and the wide fluctuation over the follow-up period in quality-of-life ratings.

- In 2010, Gothard et al. randomized 58 patients with arm lymphedema (at least 15% increase in arm volume) following cancer treatment in a 2:1 ratio to receive HBOT (n=38) or usual care without HBOT (n=20). Fifty-three patients had baseline assessments and 46/58 (79%) had 12-month assessments. No statistically significant difference was found in the change in arm volume from baseline to 12-month follow-up. The median change from baseline was -2.9% in the treatment group and -0.3% in the
control group. The study protocol defined response as at least an 8% reduction in arm volume relative to the contralateral arm. According to this definition, 9 of 30 (30%) patients in the HBOT group were considered responders compared with 3 of 16 (19%) in the control group; the difference between groups was not statistically significant. Other outcomes, e.g., quality-of-life scores on the Short-Form (SF)-36, were also similar between groups.

Conclusion

Due to the lack of sufficient evidence from well-designed clinical trial, HBOT for the treatment of adverse effects related to radiation therapy is considered investigational.

Radionecrosis and Osteoradionecrosis

Several systematic reviews of RCTs have been published. A 2008 Cochrane review by Esposito et al reviewed the use of HBOT in patients requiring dental implants.[29] The authors identified 1 randomized trial involving 26 patients. The authors concluded that despite the limited amount of clinical research available, it appears that HBOT in irradiated patients requiring dental implants may not offer any appreciable clinical benefits. They indicated that there is a need for more RCTs to ascertain the effectiveness of HBOT in irradiated patients requiring dental implants.

In 2012, Bennett et al published a Cochrane review on HBOT for late radiation tissue injury.[58] The authors identified 11 RCTs; there was variability among trials, and study findings were not pooled for the primary outcomes of survival, complete resolution of necrosis or tissue damage, and improvement in a late effects symptom scale. In a pooled analysis of 3 studies, a significantly higher proportion of patients with osteoradionecrosis achieved complete mucosal cover after HBOT compared with controls (RR=1.30; 95% CI, 1.09 to 1.55). From their review of the literature, the authors concluded that data from small trials “suggest that for people with LRTI [late radiation tissue injury] affecting the head, neck, anus, and rectum, [HBOT] is associated with improved outcome. HBOT also appears to reduce the chance of ORN [osteoradionecrosis] following tooth extraction in an irradiated field. There was no such evidence of any important clinical effect on neurological tissues. The application of HBOT to selected patients and tissues may be justified.”

Stroke

Current evidence is insufficient to permit conclusions about whether HBOT improves health outcomes in the treatment of stroke or stoke-related functional limitations. The following is a summary of the available evidence:

Acute Stroke

• Systematic Reviews
  • In a 2005 Cochrane systematic review, Bennett and colleagues evaluated HBOT for acute stroke.[59] The investigators identified 6 RCTs with a total of 283 participants that compared HBOT to sham HBOT or no treatment. The authors were only able to pool study findings for 1 outcome, the mortality rate at 3-6 months. A pooled analysis of 3 trials found no significant benefit of HBOT compared to the control for this outcome. Based on the available evidence, acute ischemic stroke is considered investigational.
In a 2005 systematic review, Carson and colleagues concluded that current evidence did not demonstrate any benefit with the use of HBOT for the treatment of stroke. The authors noted it was undetermined whether there were any benefits with HBOT that would outweigh potential harms, and further study was required.

In a 2014 update of a Cochrane systematic review, Bennett et al evaluated HBOT for acute ischemic stroke. The investigators identified 11 RCTs with a total of 705 participants that compared HBOT with sham HBOT or no treatment. The authors were only able to pool study findings for 1 outcome; mortality at 3 to 6 months. A pooled analysis of data from 4 trials with a total of 106 participants did not find a significant benefit of HBOT compared with a control condition for this outcome (RR=0.97; 95% CI, 0.34 to 2.75).

Randomized Controlled Trials (RCTs)

No RCTs have been published since the 2005 systematic reviews.

Stroke-related motor dysfunction

In 2013, Efrati and colleagues published an RCT evaluating HBOT for treatment of neurologic deficiencies associated with a history of stroke. The study included 74 patients with at least one motor dysfunction who had an ischemic or hemorrhagic stroke 6-36 months prior to study participation. Participants were randomly assigned to receive 2 months of HBOT (40 daily sessions, 5 days per week, n=30) or delayed treatment (n=32). Patients were evaluated at baseline and 2 months. For patients in the delayed treatment control group, outcomes were evaluated at 4 months after crossing over and receiving HBOT. Twenty-nine of 32 patients (91%) in the delayed treatment group crossed over to the active intervention. Outcome measures included the National Institutes of Health Stroke Scale (NIHSS), which was measured by physicians blinded to treatment group, and several patient-reported quality-of-life and functional status measures.

At 2 months’ follow-up, there was statistically significantly greater improvement in function in the HBOT group compared to the control group as measured by the NIHSS, quality-of-life scales and the ability to perform activities of daily living (ADLs). These differences in outcome measures were accompanied by improvements in single-photon emission computed tomography (SPECT) imaging in the regions affected by stroke. For the delayed treatment control group, there was a statistically significant improvement in function after HBOT compared to before treatment. This RCT raises the possibility that HBOT may induce improvements in function and quality of life for post-stroke patients with motor deficits. However, the results are not definitive for a number of reasons. This RCT is small and enrolled a heterogeneous group of post-stroke patients. The study was not double-blind and the majority of outcome measures, except for the NIHSS, were patient reported and thus prone to the placebo effect. Also, there was a high total dropout rate of 20% at the 2-month follow-up point. Therefore, larger, double-blind studies with longer follow-up are needed to corroborate these results. Because of these limitations in the evidence, HBOT is considered investigational for treating motor dysfunction associated with stroke.

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Traumatic Brain Injury

Systematic Review

A 2012 Cochrane systematic review addressed HBOT as adjunctive treatment for traumatic brain injury.[62] The investigators identified 7 RCTs with a total of 571 participants comparing a standard intensive treatment regimen to the same treatment regimen with the addition of HBOT. The review did not include studies in which interventions occurred in a specialized acute care setting. The HBOT regimens varied among studies; for example, the total number of individual sessions varied from 3 to 30-40. No trial used sham treatment or blinded the staff members who were treating the patients, and only 1 had blinding of outcome assessment. Allocation concealment was inadequate in all of the studies. The primary outcomes of the review were mortality and functional outcomes. A pooled analysis of data from 4 trials that reported this outcome found a statistically significantly greater reduction in mortality when HBOT was added to a standard treatment regimen. However, when data from the 4 trials were pooled, the difference in the proportion of patients with an unfavorable functional outcome at final follow-up did not reach statistical significance. Unfavorable outcome was commonly defined as a Glasgow Outcome Score (GOS) of 1, 2 or 3, which are described as ‘dead’, ‘vegetative state’ or ‘severely disabled’. Studies were generally small and were judged to have substantial risk of bias.

Randomized Controlled Trials

A 2012 sham-controlled double-blind trial evaluating HBOT was published after the 2012 Cochrane review.[63] The study included 50 military service members, 48 of whom were male, with combat-related mild traumatic brain injury. Participants were randomized to 30 sessions of HBOT over 8 weeks (n=25) or a sham intervention (room air at 1.3 ATA) (n=25). The primary outcome measures were scores on the Immediate Post-Concussive Assessment and Cognitive Testing (ImPACT) and Post-Traumatic Disorder Check List–Military Version (PCL-M) instruments. Patients were evaluated after every 5 treatment sessions and at 6 weeks post-exposure. Forty-eight of 50 participants (96%) completed the study. There were no statistically significant differences on the ImPACT total mean score or the PCL-M composite score at any time point. While the sample size was relatively small, the study was powered to detect clinically significant differences among groups on the cognitive tests.

Several trials on mild traumatic brain injury in military populations have been published and these did not find significant benefits of HBOT compared with sham treatment. The first trial, published by Wolf et al in 2012, included 50 military service members, 48 of whom were male, with combat-related mild traumatic brain injury. Participants were randomized to 30 sessions of HBOT over 8 weeks (n=25) or a sham intervention (room air at 1.3 atmosphere, absolute [ata]) (n=25). The primary outcome measures were scores on the Immediate Post-Concussive Assessment and Cognitive Testing (ImPACT) and Post-Traumatic Disorder Check List–Military Version (PCL-M) instruments. Patients were evaluated after every 5 treatment sessions and at 6 weeks post-exposure. Forty-eight of 50 participants (96%) completed the study. There were no statistically significant differences on the ImPACT total mean score or the PCL-M composite score at any time point. For example, at the 6-week follow-up, mean composite PCL-M scores were 41.6 in the HBOT group and 40.6 in the sham-control group (p=0.28). While the sample size was relatively small, the study was powered to detect clinically significant differences among groups on the cognitive tests.
A 2014 double-blind sham-controlled trial 2014 RCT by Cifu et al included 61 male Marines who had a history of mild traumatic brain injury and postconcussive syndrome. To maintain blinding, all patients were pressured inside a hyperbaric chamber to 2.0 ata. They were randomized to breathe 1 of 3 oxygen p[nitrogen gas mixes equivalent to: (1) 75% oxygen at 1.5 ata (n=21); (2) 100% oxygen at 2.0 ata (n=19); and (3) sham treatment with surface room air (n=21). Patients underwent 40 once daily 60-minute sessions. Outcomes were assessed 3 months after the last exposure. The primary outcome was a clinically meaningful improvement, defined as a 10% difference between groups in the score on the Rivermead Post-Concussion Questionnaire (RPQ)–16 (scale range, 50-84; higher values indicate more severe symptoms). At follow-up, there was no statistically significant difference among groups on RPQ-16 score (p=0.41). A variety of secondary outcomes were also assessed. None of these, including measures of attention, cognition, or depression, differed significantly among groups at follow-up.

Also in 2014, Miller et al evaluated HBOT in 72 military service members with continuing symptoms at least 4 months after mild traumatic brain injury. Patients were randomized to receive 40 daily HBO sessions at 1.5 ata, 40 sham sessions consisting of room air at 1.2 ata or standard care with no hyperbaric chamber sessions. The primary outcome was change in the RPQ. A cutoff of 15% improvement was deemed clinically important, which translates to a change score of at least 2 points on the RPQ-3 subscale. The proportion of patients who met the prespecified change of at least 2 points on the RPQ-3 was 52% in the HBOT group, 33% in the sham group and 25% in the standard care-only group. The difference between rates in the HBOT and sham groups was not statistically significant (p=0.24). None of the secondary outcomes significantly favored the HBOT group. A criticism of this study, as well as the other military population studies, was that the response in the sham group was not due to a placebo effect but to an intervention effect of slightly increased atmospheric pressure (1.2 ata). Other researchers have noted that room air delivered at 1.2 ata would not be considered an acceptable therapeutic dose for any indication, and especially for a condition with persistent symptoms like postconcussive syndrome.

Conclusion

A systematic review of small trials with limitations found a mortality reduction with HBOT but no significant improvement in patient function among survivors of traumatic brain injury. Two double-blind, sham-controlled RCTs of HBO treatment in a military population with mild traumatic brain injury did not find a statistically significant benefit with HBOT. Thus, the evidence is insufficient that HBOT improves health outcomes in patients with traumatic brain injury, and this indication is considered investigational.

Wounds Unrelated to Diabetes

Acute Surgical and Traumatic Wounds

- Systematic Reviews

  - A 2013 updated Cochrane review analyzed randomized controlled trials comparing either HBOT with a different intervention, or two HBOT regimens for acute wounds (e.g., surgical wounds, lacerations, traumatic wounds, and animal bites).[64] The four studies that met inclusion criteria ranged in size from 10 to 135 subjects. Reported outcomes were mixed. Meta-analysis of pooled data was not possible due to differences among studies with respect to patient characteristics, interventions...
studied, and outcome measures. Also identified was a high risk of bias due to insufficient disclosure of randomization methods and selective reporting of outcome data. Findings of individual studies were mixed.

The primary outcome examined by Cochrane reviewers, wound healing was not reported in either of the 2 trials comparing HBOT with usual care\[65,66\] or in the 1 trial comparing HBOT with dexamethasone or heparin.\[67\] Complete wound healing was reported in the 1 RCT comparing active HBOT with sham HBOT.\[68\] In this small study (n=36), there was a statistically higher rate of wound healing in the active HBOT group. The time point for outcome measurement in this study was unclear, but there was no statistically significant difference between groups in the meantime to wound healing. Adverse effects included 2 additional surgical procedures in 1 patient in the HBOT group compared with 8 in 6 patients in the sham group. The HBOT group had significantly fewer patients who developed necrotic tissue (1 and 8, respectively). There were no amputations in the HBOT group compared with 2 amputations in the sham group, but this difference did not reach statistical significance. The authors concluded that evidence remains insufficient to support the routine use of HBOT for acute surgical or traumatic wounds. They recommended further evaluation in high quality RCTs that include outcomes measures of complete wound closure and accelerated wound closure.

○ In 2014 Dauwe et al. published a systematic review that included 8 studies with sample sizes ranging from 5 to 125 patients. Four studies were randomized, 3 were prospective non-RCTs, and 1 was a retrospective non- RCT. Data were not pooled due to the heterogeneity described below. The authors noted that 7 of the 8 studies reported achieving statistical significance in their primary end points, but the end points differed among studies (eg, graft survival, length of hospital stay, wound size). Moreover, the studies were heterogeneous in terms of treatment regimens, patient indications (eg, burns, face lifts), and study designs, making it difficult to draw conclusions about the effect of HBOT on acute wound treatment.

• Randomized Controlled Trials (RCTs)

No RCTs have been published since those included in the systematic reviews summarized above.

Chronic Non-diabetic Wounds

• Systematic Reviews

Several systematic reviews of RCTs have been published. An updated 2007 Cochrane review of randomized controlled trials (RCTs) on HBOT for chronic wounds was published by Kranke and colleagues in 2012.\[69\] The authors identified 9 RCTs with a total of 471 participants that compared the effect of HBOT on chronic wound healing compared with an alternative treatment approach that did not use HBOT. Eight of the 9 trials included in the review evaluated HBOT in patients with diabetes. The remaining trial addressed HBOT for patients with venous ulcers; that study had only 16 participants and the comparator treatment was not specified. In a pooled analysis of data from 3 trials, a significantly higher proportion of ulcers had healed at the end of the treatment period (6 weeks) in the group receiving HBOT compared to the group not receiving HBOT (RR: 5.20: 95% CI: 1.25 to 21.7). Pooled analyses, however, did not
find significant differences between groups in the proportion of ulcers healed in the HBOT versus non-HBO-treated groups at 6 months (2 trials) or 12 months (3 trials). There were insufficient data to conduct pooled analyses of studies evaluating HBOT for treating patients with chronic wounds who did not have diabetes.

In 2013, O’Reilly et al. published a systematic review of studies on HBOT for treatment of diabetic ulcers. The authors identified 6 RCTs and 6 non-RCTs that compared HBOT with standard wound care or sham therapy in patients with diabetes who had nonhealing lower-limb ulcers. Pooled analyses of observational studies found statistically significant benefits of HBOT on rates of major amputation, minor amputation and the proportion of wounds healed at the end of the study period. However, in pooled analyses of RCT data, the stronger study design, there were no statistically significant differences between groups on key outcomes. This included the rate of major amputation (RR=0.40; 95% CI, 0.07 to 2.23; p=0.29), minor amputation (RR=0.79; 95% CI, 0.19 to 3.30, p=0.75), and the proportion of unhealed wounds at the end of the study period (RR=0.54, 95% CI, 0.26 to 1.13, p=0.1).

Systematic reviews have had mixed findings on the impact of HBOT on diabetic ulcers. A Cochrane review found short-term, but not long-term benefit on wound healing, and a 2013 meta-analysis did not find significant benefits of HBOT on outcomes in RCTs, but did find an effect in non-RCTs. There is insufficient evidence on HBOT for treatment of chronic wounds in patients without diabetes.

- **Randomized Controlled Trials (RCTs)**

No RCTs have been published since those included in the systematic reviews summarized above.

- **Conclusion**

Published clinical trial data is insufficient to determine the effectiveness of HBOT for wounds that are not related to diabetes. The UHMS does not include these wounds in their list of indications for HBOT, noting the lack of available evidence. As shown in studies of adjunctive HBOT for treatment of severe diabetic lower extremity ulcers, this treatment is well suited to randomized, controlled comparative trials. In spite of this, only 1 small (n=16) randomized, controlled trial was found for non-diabetic wounds. This trial is too small and short-term to be reliable.

**Other Indications**

No data from well-designed randomized, controlled clinical trials were found that supported HBOT for any other investigational indication, including but not limited to refractory mycoses and acute peripheral arterial insufficiency.

**Other indications**

For the indications listed below, insufficient evidence to support the use of HBOT was identified. Since 2000, there have been no published controlled trials or large case series (i.e., >25 patients):

- bone grafts;
• carbon tetrachloride poisoning, acute;
• cerebrovascular disease, acute (thrombotic or embolic) or chronic;
• fracture healing;
• hydrogen sulfide poisoning;
• intra-abdominal and intracranial abscesses;
• lepromatous leprosy;
• meningitis;
• pseudomembranous colitis (antimicrobial agent-induced colitis);
• radiation myelitis;
• sickle cell crisis and/or hematuria;
• amyotrophic lateral sclerosis;
• retinopathy, adjunct to scleral buckling procedures in patients with sickle cell peripheral retinopathy and retinal detachment;
• pyoderma gangrenosum;
• tumor sensitization for cancer treatments, including but not limited to, radiotherapy or chemotherapy;

SUMMARY OF EVIDENCE

There is sufficient published evidence to determine that use of hyperbaric oxygen therapy (HBOT) in selected patients with nonhealing diabetic wounds of the lower extremities, acute traumatic ischemia, soft-tissue radiation necrosis (eg, radiation enteritis, cystitis, proctitis), osteoradionecrosis (ie, pre- and posttreatment) for patients undergoing dental surgery (non-implant-related) of an irradiated jaw, gas gangrene, and profound anemia with exceptional blood loss when blood transfusion is impossible or must be delayed improves the net health outcome. There is insufficient evidence for patients all other indications included in the Rationale section that HBOT improves the net health outcome.

PRACTICE GUIDELINE SUMMARY

U.S. FOOD AND DRUG ADMINISTRATION (FDA)

In 2013, the FDA published a position statement with a warning that HBOT has not been proven safe and effective for uses not cleared by the agency.\(^1\) This statement was developed due to numerous complaints from consumers and health care professionals that unproven claims made by some HBOT centers may mislead consumers and ultimately endanger their health. The statement included the following conditions for which patients may be unaware that safety and effectiveness of HBOT have not been established:

• AIDS/HIV
• Alzheimer's Disease
• Asthma
• Bell's Palsy
• Brain Injury
• Cerebral Palsy
• Depression
• Heart Disease
• Hepatitis
• Migraine
• Multiple Sclerosis
• Parkinson's Disease
• Spinal Cord Injury
• Sport's Injury
• Stroke

UNDERSEA AND HYPERBARIC MEDICAL SOCIETY (UHMS)

In 2015, the Undersea and Hyperbaric Medical Society (UHMS) published a guideline on the use of HBOT for treatment diabetic foot ulcers. Recommendations are as follows:

• Suggest against using HBOT in patients with Wagner Grade 2 or lower diabetic foot ulcers
• Suggest adding HBOT in patients with Wagner Grade 3 or higher diabetic foot ulcers that have now shown significant improvement after 30 days of standard of care therapy
• Suggest adding acute post-operative HBOT to the standard of care in patients with Wagner Grade 3 or higher diabetic foot ulcers who have just had foot surgery related to their diabetic ulcers.

• Appropriate Indications for HBOT

In 2014, the UHMS updated their list of indications considered appropriate for hyperbaric oxygen therapy. These indications are as follows:

- Acute thermal burn injury
- Air or gas embolism
- Arterial insufficiencies (central retinal artery occlusion; enhancement of healing in selected problem wounds)
- Carbon monoxide poisoning and carbon monoxide poisoning complicated by cyanide poisoning
- Clostridial myositis and myonecrosis (gas gangrene)
- Compromised grafts and flaps
- Crush injury, compartment syndrome, and other acute traumatic ischemias
- Decompression sickness
- Delayed radiation injury (soft tissue and bony necrosis)
- Intracranial abscess
- Idiopathic sudden sensorineural hearing loss (ISSNHL) (patients with moderate to profound ISSNHL who present within 14 days of symptom onset)
- Necrotizing soft tissue infections
- Osteomyelitis (refractory)
- Severe anemia

- Autism Spectrum Disorder (ASD)\(^\text{[13]}\)

The 2009 UHMS position paper included a critical appraisal of the available literature, in particular the 2009 Rossignol et al. RCT\(^\text{[11]}\) which was the only RCT available at that time. The paper concluded that “the UHMS cannot recommend the routine treatment of ASD with HBO\(_2\)T outside appropriate comparative research protocols.”

- Chronic Brain Injury\(^\text{[76]}\)

The most recent UHMS position statement on chronic brain injury (e.g., traumatic brain injury, cerebral palsy, stroke) is from 2003. The statement considered the evidence to be insufficient to support a recommendation for HBOT for the chronic sequelae of traumatic or non-traumatic brain injury, but noted that continued monitoring of data is warranted.

- Idiopathic Sudden Sensorineural Hearing Loss (ISSNHL)\(^\text{[77]}\)

In October 2011, the UHMS Executive Board approved ISSNHL as an additional indication. According to treatment guidelines, patients with moderate to profound ISSNHL who present within 14 days of symptom onset should be considered for HBOT treatment.

- Multiple Sclerosis\(^\text{[47]}\)

A 2010 UHMS position paper reported that most RCTs have failed to show clinical benefit for HBOT therapy for multiple sclerosis. “We conclude that, while there is some case for further investigation of possible therapeutic effects in selected sub-groups of patients (well-characterized and preferably early in the disease course) and for the response to prolonged courses of HBOT, this case is not strong. At this time, the UHMS cannot recommend the routine treatment of MS with HBOT outside appropriate comparative research protocols.”

- Topical Oxygen for Chronic Wounds\(^\text{[78]}\)

A 2005 UHMS position statement reported that, “to date, mechanisms of action whereby topical oxygen might be effective have not been defined or substantiated. Conversely, cellular toxicities due to extended courses of topical oxygen have been reported, although, again these data are not conclusive, and no mechanism for toxicity has been examined scientifically...The only randomized trial for topical oxygen in diabetic foot ulcers actually showed a tendency toward impaired wound healing in the topical oxygen group. Contentions that topical oxygen is superior to hyperbaric oxygen are not proven.” Therefore, the UHMS recommends against application of topical oxygen outside a clinical trial setting, noting that topical oxygen “should be subjected to the same intense scientific scrutiny to which systemic hyperbaric oxygen has been held.”

**NATIONAL BOARD OF DIVING & HYPERBARIC MEDICAL TECHNOLOGY**\(^\text{[7]}\)

As noted above, the current position statement concluded that “the installation and provision of in-home hyperbaric oxygen therapy is inherently unsafe and cannot be condoned.” This position is based on concern for the safety and well-being of patients as well as those people in proximity to the HBOT delivery system because in-home provision of HBOT is likely to:
1. Bypass otherwise mandatory federal, state, and local codes related to design, construction, installation, and operation of these devices; and
2. Occur without adequate physician oversight and the operational support of appropriately qualified HBOT providers.

AMERICAN ACADEMY OF OTOLARYNGOLOGY-HEAD AND NECK SURGERY (AAO-HNS)[79]

In 2012, the AAO-HNS published an evidence-based clinical practice guideline on treatment of sudden hearing loss. The guideline includes a statement that HBOT may be considered a treatment option for patients who present within 3 months of a diagnosis of idiopathic sudden sensorineural hearing loss. The document states, “Although HBOT is not widely available in the United States and is not recognized by many U.S. clinicians as an intervention for ISSNHL, the panel felt that the level of evidence for hearing improvement, albeit modest and imprecise, was sufficient to promote greater awareness of HBOT as an intervention for [this condition]” The strength of the recommendation was rated “Option” defined in this case as based on systematic reviews of RCTs with a balance between benefit and harm.

SUMMARY

Hyperbaric oxygen therapy (HBOT) has been studied for a wide variety of clinical indications. There is enough evidence to show that HBOT is safe and effective for a variety of indications. There are guidelines based on research that recommend the use of HBOT for a variety of indications. Therefore, the use of HBOT may be considered medically necessary when policy criteria are met.

For the investigational indications discussed in the policy, the evidence is not sufficient to permit conclusions concerning the effects of HBOT on final health outcomes. Therefore, these indications are considered investigational.

REFERENCES


November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


44. Ioppolo, F, Tattoli, M, Di Sante, L, et al. Clinical improvement and resorption of calcifications in calcific tendinitis of the shoulder after shock wave therapy at 6 months' follow-up: a systematic review and meta-analysis. *Archives of physical medicine and rehabilitation*. 2013 Sep;94(9):1699-706. PMID: 23499780


46. (CADTH), CAfDaTiH. Hyperbaric Oxygen Therapy for Adults with Mental Illness: A Review of the Clinical Effectiveness. 2014. PMID: 28081957


73. LBE., G. Hyperbaric oxygen therapy indications, 12th Edition. The Hyperbaric Oxygen Therapy Committee Report.: Undersea and Hyperbaric Medical Society. 2008. PMID:


<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>99183</td>
<td>Physician or other qualified health care professional attendance and supervision of hyperbaric oxygen therapy, per session</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Note: This code is not intended for reporting systemic oxygen therapy in chambers that provide oxygen at less than hyperbaric pressure (e.g., “mild hyperbaric” oxygen therapy) which should be reported using code 99199.</td>
</tr>
<tr>
<td></td>
<td>99199</td>
<td>Unlisted special service, procedure or report</td>
</tr>
<tr>
<td>HCPCS</td>
<td>A4575</td>
<td>Topical hyperbaric oxygen chamber, disposable</td>
</tr>
<tr>
<td>Codes</td>
<td>Number</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>E0446</td>
<td></td>
<td>Topical oxygen delivery system, not otherwise specified, includes all supplies and accessories</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NOTE: This code is intended for devices such as the TransCu 02 that deliver oxygen at normal atmospheric pressure under wound dressings; it should not be used to report topical hyperbaric oxygen therapy devices.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1399</td>
<td></td>
<td>Durable medical equipment, miscellaneous</td>
</tr>
<tr>
<td>G0277</td>
<td></td>
<td>Hyperbaric oxygen under pressure, full body chamber, per 30 minute interval</td>
</tr>
</tbody>
</table>

*Date of Origin: January 1996*
Surface Electromyography (SEMG) Including Paraspinal SEMG

Effective: June 1, 2018

Next Review: April 2019
Last Review: April 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Surface electromyography (SEMG) is a non-invasive, computer-based procedure, most commonly used in an office setting to assess muscle function by recording muscle activity from above the muscle on the skin surface.

MEDICAL POLICY CRITERIA

Note: This policy addresses only the use of surface electromyography alone or in combination with other services. See the Cross References below for additional gait analysis criteria not specifically addressed in this policy.

Dynamic surface electromyography (SEMG), including paraspinal SEMG, is considered investigational for all indications, including but not limited to any of the following:

A. Diagnosing and monitoring of back pain
B. Evaluation of myoclonus
C. Component of gait analysis

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.
SEMГ includes a scanner with surface electrodes that record electrical impulses of nerves at rest (i.e. static) and during activity (i.e. dynamic) in order to characterize the electrical potential of a specific muscle or group of muscles. Electrical activity can be assessed by computer analysis of the frequency spectrum (i.e., spectral analysis), amplitude, or root mean square of the electrical action potentials.

Unlike needle electromyography (NEMГ), SEMГ utilizes electrodes that record from a wide muscle area, have a relatively low frequency band, low signal resolution, and are highly susceptible to movement. SEMГ has been proposed as a diagnostic tool in patients with various degenerative, neuromuscular or motor control disorders such as: back pain, intervertebral disc disease, soft tissue injury, temporomandibular joint dysfunction (TMJ), bruxism (teeth grinding), nerve root irritation, and scoliosis.

PARASPINAL SEMГ

Like SEMГ, paraspinal SEMГ is performed using a single or multiple electrodes placed on the skin surface, with recordings made at rest, in various positions, or after a series of exercises. Recordings can also be made by using a handheld device, which is applied to the skin at different sites. Spectral analysis focusing on the median frequency has been used to assess paraspinal muscle fatigue during isometric endurance exercises.

Paraspinal SEMГ is typically performed by physiatrists or chiropractors as a technique to evaluate the physiological functioning of the back, specifically the function of the paraspinal muscles. This technique has been intended for use in patients with back pain symptoms such as spasm, tenderness, limited range of motion, or postural disorders, particularly as it relates to assessing the patient’s capacity to lift heavy objects, or the ability to return to work.

The following clinical applications of paraspinal SEMГ have been proposed:

- Clarification of a diagnosis (i.e., muscle, joint, or disc disease)
- Selection of a medical therapy course
- Selection of a physical therapy plan
- Pre-operative evaluation
- Post-operative rehabilitation
- Follow-up evaluation of acute low back pain
- Evaluation of exacerbation of chronic low back pain
- Evaluation of pain management treatment techniques

REGULATORY STATUS

SEMГ devices approved by the U.S. Food and Drug Administration (FDA) include those that use a single electrode or a fixed array of multiple surface electrodes. Several FDA-approved devices combine SEMГ with other types of monitors.

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Surface and paraspinal surface electromyography (SEMG) have been proposed as a research tool to evaluate the performance of nerves and muscles in patients with neuromuscular disorders, as a component of gait analysis, and to further understand the etiology of the resulting symptomatology, such as pain. However, validation of its use as a clinical diagnostic technique involves a sequential three-step procedure as follows:

1. **Analytical Validity**- of a device is typically assessed by studies that compare test measurements with a gold standard, and those that compare results taken with the same device on different occasions (“test-retest”).

2. **Clinical Validity**- is evaluated by the ability of a test to accurately diagnose a clinical condition in comparison with the gold standard. The sensitivity of a test is the ability to detect a disease when the condition is present (true positive), while specificity indicates the ability to detect patients who are suspected of disease but who do not have the condition (true negative). Therefore, evaluation of diagnostic performance requires independent assessment by the two methods in a population of patients who are suspected of disease but who do not all have the disease.

3. **Clinical Utility**- is established when the evidence demonstrates that the diagnostic information obtained from a test can be used to benefit patient management and improve health outcomes. Typically, randomized trials are needed to demonstrate the impact of the test on net health outcomes.

The following discussion focuses on these three steps as they apply to surface EMG, including paraspinal SEMG.

**ANALYTICAL VALIDITY**

Several studies using different SEMG devices have suggested that paraspinal SEMG, in general, is a reliable technique, based on coefficients of variation or test-retest studies,[2-7] or ability to differentiate healthy test subjects from those with back pain.[8-10] These studies use a range of different methodologies and SEMG parameters, and do not address the accuracy or validity of the test. No studies were identified that compared the performance of SEMG to a gold standard reference test.

**CLINICAL VALIDITY**

No articles were identified directly comparing the results of SEMG (which tests groups of muscles) with needle electromyography (which tests individual muscles) for diagnosing any specific muscle pathology. It is recognized that the pathology of individual muscles (i.e., radiculopathy, neuropathy, etc.) may represent a different process than the pathology of muscle groups (i.e., muscle strain, spasm, etc.); thus, SEMG may be considered by its advocates as a unique test for which there is currently no gold standard. Even if one accepts this premise, there are inadequate data to evaluate the diagnostic performance of SEMG. No articles were identified in the published peer-reviewed literature that established definitions of normal or abnormal SEMG. In some instances, asymmetrical electrical activity may have been used to define abnormality; results may be compared to a “normative data base.” However, there is a lack of published literature defining what degree of asymmetry would constitute abnormality, or how a normative database was established.[11]
In the absence of a gold standard diagnostic test, correlation with the clinical symptoms and physical examination is critical.

Audag (2017) published a systematic review (SR) comparing tools for screening for dysphagia and evaluation in neuromuscular diseases.[12] Four studies including four evaluation tools for Duchenne muscular dystrophy met inclusion criteria. Evaluation tools included were the Sydney Swallow Questionnaire, surface electromyography, Neuromuscular Disease Swallowing Status Scale, and videofluoroscopic swallow study. Three studies were assessed as fair quality and one as good quality. Two studies compared between different evaluation tools and two compared between groups of subjects. The only study that assessed SEMG compared results from patients and healthy controls. Greater intrasubject variability was observed for Duchenne muscular dystrophy patients than healthy controls, but there were no differences within patients between those with and without dysphagia. The SR concluded that more research was needed to identify the best assessment method.

In 2016, Villafane conducted a systematic review of studies testing the validity and clinical applicability of SEMG among patients with chronic non-specific low back pain (CNSLBP).[13] The literature review, conducted through September 2014, identified 24 studies for inclusion. Quality of the studies was assessed using a modification of the checklist for cohort, case-control, and cross-sectional studies from Strengthening the Reporting of Observational Studies in Epidemiology. The checklist has 22 items, and the authors used the 15 items that related to methods and results. Out of a possible total 15 points, the studies’ scores ranged from 6 to 12. The review focused on the 10 studies with scores from 10 to 12. One study was large (N=349), the second largest had 67 patients, while the remaining studies had less than 40 patients. While SEMG recordings were taken, patient position (upright, seated) and type of test (for example, isometric trunk extension, semi-crouched lifting, Roman Chair endurance, etc.) varied among the studies. Villafane report inconsistent findings of validity and reliability for SEMG in discriminating between patients with CNSLBP and healthy controls. Conclusions were limited due to the heterogeneity in methods across the studies.

Wang (2016) published an SR including eleven case-control, cohort, and cross-sectional studies that evaluated the benefit of trunk muscle activity for patients with spinal cord injury (SCI), using SEMG.[14] The studies methodology varied; thus, could not be evaluated together. For example, two studies compared trunk muscles in SCI patients versus those in a normal healthy control group and three studies compared trunk muscle activity in SCI patients with different levels of trunk muscle impairment. The authors concluded that because trunk muscle activity can increase independence and quality of life, SEMG is a useful objective tool for measuring muscle activity for patients with SCI, but more larger studies are needed with attention to comparison of trunk muscle activity in different SCI populations and to further define SEMG protocols.

Azola (2017) published a study comparing submental SEMG (sSEMG) with videofluoroscopy (VF) biofeedback on hyo-laryngeal accuracy when training on a swallowing maneuver.[15] The first stage of the study involved accurate demonstration of the volitional laryngeal vestibule closure maneuver (vLVC) and the second stage involved 20 vLVC training swallows. Thirty healthy adults were randomized into three groups. One group received sSEMG biofeedback only, one group received VF feedback only, and one group received VF for the first stage and sSEMG for the second stage of the study (mixed feedback). The participants and clinicians viewed the biofeedback in real time during the procedure and the clinician provided guidance...
based on the biofeedback. The accuracy of the vLVC performance and the clinician cues was greater (p<0.001) when biofeedback was provided with VF as compared to sSEMG or mixed biofeedback.

A 2016 study by du Rose and Breen looked into the relationship between lumbar intervertebral range of motion and paraspinal muscle activity in healthy adults, as measured by SEMG and quantitative fluoroscopy, in order to establish “normal” measurements.[16] Fluoroscopic images and SEMG measurements were taken on 20 males with no history of low back pain. What would be considered normal intervertebral ranges of motion were related to a diverse set of muscle activation patterns as measured by SEMG. The authors concluded that larger sample sizes and measurements from patients with low back pain are needed to establish standard criterion.

Earp (2016) published a study that compared vastus lateralis muscle activity during heavy squat (HS) and unloaded jump squat (JS) activities for 10 patients using SEMG.[17] Testing occurred over two days to determine if a hypothesis that regional hypertrophy occurred during heavy squat and unloaded squat activities. The authors concluded that SEMG showed more hypertrophy in HS versus JS, which was opposite of previous research outcomes. They concluded SEMG is not a good tool for this type of assessment.

Chmielewska (2016) published a six-week biofeedback training for 21 continent women who had never been pregnant beyond 20 weeks, using SEMG as a measurement tool.[18] The goal was to determine if SEMG-biofeedback training could assist in pelvic floor muscle relaxation; thus, decreasing involuntary urine leakage. Training occurred three times a week for six weeks. SEMG evaluation occurred at baseline, three weeks, six weeks and one month following training. The results showed an increase in pelvic floor relation. The authors concluded that additional research is needed.

De Luca published a series of studies investigating a type of SEMG called the Back Analysis System (BAS), consisting of surface electrodes and other components to measure the electrical activity of muscles during isometric exercises designed to produce muscle fatigue.[19] Using physical examination and clinical history as a gold standard, the author found that BAS was able to accurately identify “control” and “back pain” patients 84% and 91% of the time, respectively, with the values increasing to 100% in some populations of patients. (Accuracy is the sum of true positive and true negative results.) However, these studies were not designed as a clinical diagnostic tool, but were intended to investigate the etiology of back pain and to investigate muscular fatigue patterns in patients with and without back pain.

Hu in Hong Kong published two articles on dynamic topography, an approach to analyzing SEMG findings.[20,21] The studies had similar protocols. Both included low back pain patients and healthy controls; all participants underwent SEMG at study enrollment and then back pain patients participated in a rehabilitation program. The first study[21] found different dynamic topography at baseline between healthy people and people with back pain, e.g., a more symmetric pattern in healthy controls. After physical therapy, the dynamic topography images of back pain patients were more similar to the healthy controls on some of the parameters that were assessed. In the second study, following rehabilitation, back pain patients were classified as responders or nonresponders based on changes in back pain severity.[20] Some associations were found between baseline SEMG parameters and response to rehabilitation. SEMG was not repeated following the rehabilitation program, and thus it is not clear whether there are any significant associations between continued symptoms and SEMG abnormalities.
Moreover, it is not clear how SEMG analysis would affect treatment decisions for low back pain patients.

**CLINICAL UTILITY**

**SEMG**

Numerous studies were identified which incorporated the use of SEMG as an assessment tool to evaluate muscle strength and movement,[22-27] temporomandibular joint dysfunction and disorders,[28-30] and various causes of muscle pain.[31-34] Several studies have proposed using SEMG results to inform treatment decisions; however, none of these studies provided data to validate that treatment based on SEMG results improved outcomes.

- In a 2016 study of patients with chronic LBP (N=216), SEMG showed potential to discriminate between impaired and unimpaired neuromuscular regulation of back extensors, which would provide useful information for designing individualized exercise programs.[35]
- In a 2015 study of patients with LBP (n=27) and pain-free controls (n=23), SEMG detected a loss of discrete motor cortical organization of the paraspinal muscles among those with LBP.[36] The invasive technique of needle electromyography is usually performed to detect this pathology. Patients with cortical reorganization may benefit from motor skill training.
- In two studies (1988, 1992), SEMG was shown to differentiate muscle spasm from muscle contracture. Muscle spasm would be treated with relaxation therapy, and contracture would be treated with stretching exercises.[37,38]

A 2000 SR by Pullman, indicated that SEMG was not found to be better or equivalent to NEMG in diagnosing neuromuscular disease due to electrical cross-talk of muscles, intervening soft tissues, and poor fidelity recordings as a result of limited spatial resolution.[1]

In 2008, Meekins conducted a SR of studies published from 1994-2006 which evaluated SEMG in the diagnosis and treatment of nerve and muscle disorders.[39] Authors concluded that:

1. SEMG may be useful in adding information in the study of fatigue in post-poliomyelitis syndrome and electromechanical coupling dysfunction in myotonic dystrophy.” However, this recommendation was based upon Class III, Level C data indicating studies were retrospective in nature, focused on SEMG for a specific condition and that data indicated SEMG may be possibly effective, ineffective, or harmful for the given condition in the specific population.

2. On the basis of two class III studies, sEMG may be useful to detect the presence of neuromuscular disease (Level C rating).

3. Data were deemed insufficient to determine the ability of SEMG in distinguishing between neuropathic and myopathic disorders, disease severity, to compare the utility of SEMG with NEMG, or as a study of fatigue in myophosphorylase deficiency, muscle fiber and motor unit propagation in myotonia congenita and hypokalemic periodic paralysis, or in evaluation of disease progression in myotonic dystrophy and Charcot–Marie–Tooth disease.
Included studies were small in nature and differed in the utilization of SEMG techniques, diagnostic reference standard and outcome measures. Authors indicated that additional studies were needed that compare SEMG to a carefully selected gold standard, in studies with adequate blinding which address a broad spectrum of subjects. The authors also noted that the lack of standardization of SEMG protocols and lack of methodological documentation prohibited pooled analysis. Well-designed, randomized controlled trials (RCTs) which evaluate SEMG compared to standard assessment measures are required in order to assess the efficacy of SEMG as a diagnostic tool for any condition.

Paraspinal SEMG to Diagnose Back Pain

Several articles described the use of SEMG as an aid in classifying low back pain.[40-49] The articles focused on the use of spectral analysis to assess muscle fatigability. However, it is unclear how this information may be used in the management of the patient. For example, while the innovators of the BAS system indicated that SEMG can suggest potential therapies by distinguishing deconditioning from muscle inhibition secondary to pain-related behavior, no clinical studies described the use of SEMG in suggesting therapy.[40]

In another application of SEMG, Arena assessed the amplitude of SEMG recordings as a measure of paraspinal muscle tension in 66 patients and reported that the degree of muscle tension did not correlate with pain levels.[50] These findings raised questions about the role of biofeedback, muscle relaxants, or other therapies designed to reduce muscle tension.

While SEMG may be used to objectively document muscle spasm or other muscular abnormalities, it is unclear how such objective documentation would supplant or enhance clinical evaluation, or how this information would be used to alter the treatment plan. For example, SEMG has been proposed as a technique to differentiate muscle spasm from muscle contracture, with muscle spasm treated with relaxation therapy, and contracture treated with stretching exercises. However, there are no data to validate that such treatment suggested by SEMG resulted in improved outcomes.[37,51] Part of the difficulty in clinical interpretation is understanding, to what extent, the SEMG abnormalities are primary or secondary. In addition, no specific workup is recommended for acute low back pain without warning signs.

A review of spinal muscle evaluation in low-back pain patients indicated that the validity of SEMG remains controversial.[52] The authors noted that although many studies showed increased fatigability of the paraspinal muscles in patients with low back pain, it is not known whether these changes are causes or consequences of the low back pain. Also, “the considerable inter-individual variability and the absence of normative data complicate the description of normal or abnormal profiles, thereby limiting the diagnostic usefulness of SEMG.”

Gait Analysis

The ideal study design to demonstrate the clinical utility of gait analysis would be a RCT comparing treatment decisions and health outcomes in patients managed with and without SEMG as a component of gait analysis. Although numerous studies were identified in which SEMG was used as a component of gait analysis to evaluate a specific treatment, no RCT were identified which evaluated the contribution of SEMG as a component of gait analysis to diagnose or treat any condition.

Myoclonus
The evidence regarding the use of SEMG to diagnose or treat myoclonus associated with any condition is limited to small case series and case reports.

**PRACTICE GUIDELINE SUMMARY**

The American Pain Society issued guidelines on the evaluation and management of low back pain that were released in two phases in 2007 and 2009.[53] When discussing the diagnostic accuracy of nonimaging tests, the guidelines stated that “There is no evidence supporting the use of thermography or surface electromyography for diagnosis of low back pain (level of evidence: fair).”

**SUMMARY**

There is not enough research to show that surface electromyography (SEMG), including paraspinal SEMG improves health outcomes for any indication, including but not limited to the diagnosis and monitoring of back pain, evaluation of myoclonus or as a component of gait analysis. No clinical guidelines based on research recommend SEMG for any indication. Therefore, the use of the use of SEMG, including paraspinal SEMG, is considered investigational for all indications.

**REFERENCES**


### CODES

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>95999</td>
<td>Unlisted neurological or neuromuscular diagnostic procedure</td>
</tr>
<tr>
<td></td>
<td>96002</td>
<td>Dynamic surface electromyography, during walking or other functional activities, 1 to 12 muscles</td>
</tr>
<tr>
<td></td>
<td>96004</td>
<td>Review and interpretation by physician or other qualified health care professional of comprehensive computer-based motion analysis, dynamic plantar pressure measurements, dynamic surface electromyography during walking or other functional activities, and dynamic fine wire electromyography, with written report</td>
</tr>
<tr>
<td></td>
<td>97799</td>
<td>Unlisted physical medicine/rehabilitation service or procedure</td>
</tr>
<tr>
<td></td>
<td>99199</td>
<td>Unlisted special service, procedure or report</td>
</tr>
<tr>
<td>HCPCS</td>
<td>S3900</td>
<td>Surface electromyography (EMG)</td>
</tr>
</tbody>
</table>

*Date of Origin: April 1999*
In Vivo Analysis of Colorectal Polyps

Effective: November 1, 2018

Next Review: October 2019
Last Review: October 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Several adjunct techniques of in vivo analysis of polyps are being researched for the purpose of improving the analysis of lesions and detection of changes in walls of colon. Use of these devices is proposed to increase the rate of polyp detection and/or to distinguish premalignant from benign lesions for removal.

MEDICAL POLICY CRITERIA

In vivo analysis of colorectal lesions, including but not limited to polyps, is considered investigational.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES

1. Confocal Laser Endomicroscopy, Medicine, Policy No. 151

BACKGROUND

During a colonoscopy or sigmoidoscopy as a screening test for colorectal cancer, the physician must often decide which polyp should be removed for histologic diagnosis. While
hyperplastic polyps are considered benign without malignant potential, adenomatous polyps are thought to represent one of the earliest stages in the progression to a malignancy. Identification of these premalignant lesions is considered one of the cornerstones of colorectal cancer prevention. The physician must thus balance the time and potential morbidity of removing all polyps, many of which will be benign, versus removal of those polyps most likely to be adenomatous.

Several techniques of in vivo analysis of polyps are being researched for the purpose of improving the analysis of lesions and detection of changes in walls of colon. These methods are intended to be used as an adjunct to colonoscopy. Some of these methods include autofluorescence, narrow band imaging (NBI), multi-band imaging, chroendoendoscopy, third eye retroscop and firoboptic analysis. It is proposed that these technologies may allow for in vivo analysis of the polyps, possibly avoiding unnecessary biopsies and increasing detection of difficult to visualize lesions (e.g., flat lesions).

The first system developed was based on the observation that benign and malignant tissues emit different patterns and wavelengths of fluorescence after exposure to a laser light. This system consists of an optical fiber, emitting a laser that is directed against three different regions of the same polyp. The subsequent fluorescent signal is collected, measured, and analyzed by a proprietary software system, which classifies a polyp as "suspicious" (i.e., adenomatous) or "not suspicious" (i.e., hyperplastic). There are several different types of spectroscopy-based in vivo techniques that rely on autofluorescence, emitting light at different frequencies in an attempt to distinguish between hyperplastic and adenomatous lesions.

Narrow band imaging (NBI) is another new technique that allows visualization of the mucosal surface and capillary vessels and thus may assist in the differentiation of abnormal from normal mucosa during colonoscopy. Two NBI systems are available. The NBI color chip system is used in the United States; in this system a single filter with a two-band pass characteristic is used to generate central wavelengths at 415 nm (blue) and 540 nm (green and red). The NBI red-green-blue sequential illumination system uses narrow spectra of red, green, and blue light and a video endoscopic system with a frame sequential lighting method. The light source unit consists of a xenon lamp and a rotation disk with three optical filters. The rotation disk and monochrome charge-coupled device are synchronized and sequentially generate images in three optical filter bands. By use of all three band images, a single color endoscopic image is synthesized by the video processor. NBI has limited penetration into the mucosal surface and has enhanced visualization of capillary vessels and their fine structure on the surface layer of colonic tissue.

Chromoendoscopy, also known as chromoscopy and chromocolonoscopy, refers to the application of topical stains or dyes during endoscopy to enhance tissue differentiation or characterization and facilitate identification of mucosal abnormalities. Chromoendoscopy may be particularly useful for detecting flat or depressed lesions. Standard colonoscopy uses white light to view the colon. In chromoendoscopy, stains are applied, resulting in color highlighting of areas of surface morphology of epithelial tissue. The dyes or stains are applied via a spray catheter that is inserted down the working channel of the endoscope. Chromoendoscopy can be used in the whole colon (pancolonic chromoendoscopy) on an untargeted basis or can be directed to a specific lesion or lesions (targeted chromoendoscopy). Chromoendoscopy differs from endoscopic tattooing in that the former uses transient stains, whereas tattooing involves the use of a long-lasting pigment for future localization of lesions.
Virtual chromoendoscopy (also called electronic chromoendoscopy) involves imaging enhancements with endoscopy systems that could be an alternative to dye spraying. One system is the Fujinon® Intelligent Color Enhancement (FICE®) feature (Fujinon Inc.). This technology uses postprocessing computer algorithms to modify the light reflected from the mucosa from conventional white light to various other wavelengths.

REGULATORY STATUS

Auto-fluorescence

In 2000, the Optical Biopsy™ System (SpectraScience™, Inc.) was approved by the Food and Drug Administration (FDA). The FDA-labeled indication for the Optical Biopsy™ System reads as follows:[1]

"The SpectraScience™ Optical Biopsy™ System is indicated for use as an adjunct to lower gastrointestinal endoscopy. The device is intended for the evaluation of polyps less than 1 cm in diameter that the physician has not already elected to remove. The device is only to be used in deciding whether such polyps should be removed (which includes submission for histological examination)."

NBI

NBI received FDA clearance through the 510K process in 2005. This clearance (K051645) added NBI with the EVIS EXERA 160A System (Olympus Medical Systems Corp.) to existing endoscopic equipment. FDA indications are for endoscopic diagnosis, treatment, and video observation. In addition, in 2012, the EVIS EXERA III System, which has dual focus (DF) capabilities received FDA approval.[2]

Chromoendoscopy

In August of 2016, the Fuse Colonoscope with FuseBox Processor was cleared for marketing by the U.S. Food and Drug Administration (FDA) through the 510(k) process.[3] This system is indicated for use within the lower digestive tract for adult patients. This system includes Lumos and is intended to be used as an optional adjunct following white light endoscopy and is not intended to replace histopathological sampling as a means of diagnosis.

In August 2014, the Fujifilm EPX-4440HD Digital Video Processor with FICE and Light Source was cleared for marketing by the U.S. Food and Drug Administration (FDA) through the 510(k) process. In October of 2015, the PMA was extended to include and additional digital video processor, EPX-4440. FDA documents state that FICE can be used to supplement white-light endoscopy but is not intended to replace histopathologic sampling as a means of diagnosis. In January 2017, the Fujifilm Processor VP-7000 and Light source BL-7000 was cleared for marketing by the U.S. Food and Drug Administration (FDA) through the 510(k) process with the EPX-4440HD as a predicate device.[4] FDA documents state “BLI (Blue Light Imaging), LCI (Linked Color Imaging) and FICE (Flexible spectral-Imaging Color Enhancement) are adjunctive tools for gastrointestinal endoscopic examination which can be used to supplement Fujifilm white light endoscopy. BLI, LCI and FICE are not intended to replace histopathological sampling as a means of diagnosis.”

In April 2003, the i-scan™ (Pentax), used for virtual chromoendoscopy, was cleared for marketing by FDA through the 510(k) process.[5] This is a digital image enhancement technology and is part of the Pentax EPK-i5010 and EPK-i7010 Video Processors. The i-scan
has several modes that digitally enhance images in real–time during endoscopy. FDA documents state that i-scan is intended as an adjunct following white-light endoscopy and is not intended to replace histopathologic analysis.

No dye or stain product has been specifically approved by FDA for use in chromoendoscopy.

EVIDENCE SUMMARY

MULTIPLE TECHNIQUES

Systematic Reviews

Lord (2018) performed a systematic review of the diagnostic accuracy of several techniques of colonic lesion characterization.[6] A total of 22 studies assessing techniques for in-vivo optical characterization of lesions in patients with colonic IBD during colonoscopy, including 1,491 patients, met inclusion criteria. Techniques examined were virtual chromoendoscopy (VCE), dye-based chromoendoscopy (DBC), magnification endoscopy and confocal laser endomicroscopy (CLE). The quality of included studies was rated and there was mixed quality for all three domains of risk of bias (patient selection, index test, and reference standard). Pooled sensitivities of CLE, magnification endoscopy, DBC, and VCE were 91% (95% CI: 94-98%), 90% (95% CI: 77-96%), 67% (95%CI: 44-84%) and 86% (95%CI: 62-95%), respectively. Pooled specificities of magnification endoscopy, VCE, and DBC were 87% (95%CI: 81-91%), 87% (95%CI: 72-95%), and 86% (95%CI: 72-94%), respectively, and the area under the SROC curve for CLE was 0.98 (95%CI: 0.97-0.99). The authors concluded that real-time CLE is a highly accurate technology while acknowledging that this study is limited by the fact that most CLE studies were performed by single expert users within tertiary centers.

In 2013, Wanders assessed the sensitivity, specificity, and real-time negative predictive value or NBI, image-enhanced endoscopy (i-scan), Fujinon intelligent chromoendoscopy (FICE), CLE, and autofluorescence imaging for differentiating neoplastic from non-neoplastic colon lesions.[7] A total of 91 studies were included in the analysis (NBI=56, i-scan=10, FICE=14, CLE=11 and autofluorescence imaging=11). The authors reported the following for each modality:

- “For NBI, overall sensitivity was 91.0% (95% CI 88.6-93.0), specificity 85.6% (81.3-89.0), and real-time negative predictive value 82.5% (75.4-87.9).
- For i-scan, overall sensitivity was 89.3% (83.3-93.3), specificity 88.2% (80.3-93.2), and real-time negative predictive value 86.5% (78.0-92.1).
- For FICE, overall sensitivity was 91.8% (87.1-94.9), specificity 83.5% (77.2-88.3), and real-time negative predictive value 83.7% (77.5-88.4).
- For autofluorescence imaging, overall sensitivity was 86.7% (79.5-91.6), specificity 65.9% (50.9-78.2), and real-time negative predictive value 81.5% (54.0-94.3).
- For CLE, overall sensitivity was 93.3% (88.4-96.2), specificity 89.9% (81.8-94.6), and real-time negative predictive value 94.8% (86.6-98.1).”

The authors did not recommend autofluorescence imaging as a reliable optical diagnostic option due to low specificity rates. This study did not assess whether any of these optical imaging modalities improved patient management or overall health outcomes.

Randomized Controlled Trials
Iacucci (2018) performed a randomized non-inferiority trial to determine detection rates of neoplastic lesions in IBD patients with longstanding colitis. A total of 270 patients with inactive disease were enrolled and divided evenly to be assessed by high definition (HD), dye spraying chromoendoscopy (DCE), or VCE using i-scan image enhanced colonoscopy. Neoplastic lesions were classified by the Paris classification and Kudo pit pattern followed by histological classification using the Vienna classification. VCE was determined to have non-inferior neoplastic lesion detection rates compared to DCE. HD rates of detection of all neoplastic lesions were non-inferior to DCE and VCE. Kudo pit pattern and location at the right colon were found to predict neoplastic lesions. The authors concluded that HD-WLE alone was sufficient for detection of dysplasia, adenocarcinoma, or all neoplastic lesions.

**AUTO-FLUORESCENCE IMAGING**

**Systematic Reviews**

**Nonrandomized Studies**

In 2013, Inomata conducted a prospective nonrandomized trial to evaluate colorectal lesions using a new auto-fluorescence imaging (AFI) system. A total of 88 patients with 163 lesions greater than 5 mm were evaluated using the novel AFI system which assessed the green/red (G/R) ratio for each lesion using a computer-assisted color analysis system that permits real-time color analysis during endoscopic procedures. Authors reported significant differences in the G/R ratios of hyperplastic polyps, adenoma/intramucosal cancer/submucosal (SM) superficial cancer, and SM deep cancer (p< 0.0001). The mean ± SD G/R ratios were 0.984 ± 0.118 in hyperplastic polyps and 0.827 ± 0.081 in neoplastic lesions. When a cut-off value of >0.89 was applied to non-neoplastic lesions, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were 83.9%, 82.6%, 53.1%, 95.6% and 82.8%, respectively. When a cut-off value of <0.77 was applied to identify SM deep cancers, the sensitivity, specificity, PPV, NPV, and accuracy were 80.0%, 84.4%, 29.6%, 98.1% and 84.1%, respectively. Additional studies are needed to validate these cut-off values and to assess the impact of AFI upon improved health outcomes.

The FDA approval for the SpectraScience™ Optical Biopsy™ System was based on a prospective, nonrandomized phase II study involving 101 subjects from five sites. The data from this trial have not been published in a peer-reviewed journal but are available as an FDA summary of safety and effectiveness. Patients who participated in the study had undergone a prior lower GI endoscopic procedure with at least one polyp identified. They were then referred for an additional colonoscopy exam, in which fiberoptic analysis of the polyps was performed. At the time of the colonoscopy, the physicians documented whether or not the polyp was considered hyperplastic or adenomatous, and whether or not they would remove the polyp. The fiberoptic probe was then applied to three different portions of the polyp and a segment of normal adjacent mucosa. The physician did not know the results of the analysis and thus the test did not affect patient treatment. The effectiveness of the analysis was then calculated as its ability to correctly identify adenomatous polyps (sensitivity) and to correctly identify hyperplastic polyps (specificity), either alone or in conjunction with the physician assessment. The sensitivity and specificity of the physician assessment alone was 82.7% and 50%, respectively, compared to a combined sensitivity and specificity of 96.3% and 33%, respectively. In other words, fiberoptic analysis identified additional adenomatous polyps that the physician had classified as hyperplastic and presumably would not have removed based on visual assessment alone. This increase in sensitivity comes at the price of a decrease in
specificity, as more hyperplastic polyps will undergo biopsy. However, according to the FDA, the risk of taking biopsies of additional hyperplastic polyps is minimal.

The clinical significance of these results and their effect on patient management is difficult to interpret from the data presented. It is not clear how the physician decided to select additional polyps for fiberoptic analysis (it is not entirely clear whether all polyps were analyzed and then underwent biopsy), or whether the same results could be obtained by simply randomly taking a biopsy of a subset of polyps that were considered hyperplastic on visual assessment. While adenomatous polyps are considered premalignant lesions, the evolution to cancer is a slow process requiring seven to eight years, and thus the immediate removal of all adenomatous polyps is not required. In addition, the finding of an adenomatous polyp serves as a marker that the patient should undergo more frequent endoscopic exams. It is well known that the current practice of visual inspection of polyps will certainly miss some adenomatous polyps, but this lack of sensitivity is considered acceptable if at least one adenomatous polyp is identified and the patient undergoes more frequent screening.

Few studies have been published on the SpectraScience™ Optical Biopsy™ System since 2002. A feasibility study of fiberoptic analysis of normal, adenomatous, and cancerous tissue in 11 patients was published by Mayinger in 2003. No additional literature on the Optical Biopsy™ System was found, but a report in 2006 detailed the results of spectral scattering to different colonic lesions in a small series of 45 patients.

**NARROW BAND IMAGING (NBI)**

The following evidence review for the diagnostic utility of NBI will focus on RCTs comparing NBI with white light and standard colonoscopy techniques.

**Systematic Reviews**

Sabbagh (2011) conducted a meta-analysis of studies (regardless of indication) evaluating NBI compared to colonoscopy and did not find any significant differences in the mean number of polyps (five RCT, 2479 participants), the mean number of adenomas (eight RCT, 3517 participants), and the rate of patients with at least one adenoma (eight RCT, 3512 participants). However, individual studies included in the analysis were noted to have heterogeneous populations and indications, as well as diverse findings. Overall, the authors concluded that NBI did not improve detection of colorectal polyps when compared with conventional colonoscopy.

Additional reviews assessing the ability of NBI to differentiate between neoplastic and non-neoplastic polyps have been published; however, these studies are limited due to their inclusion of nonrandomized studies and lack of analysis regarding the impact of NBI upon patient management of overall health outcomes.

**Randomized Controlled Trials**

Data from several randomized trials of NBI versus white-light colonoscopy (WLE) failed to show any advantage in total detection rate for NBI. Published randomized trials differ from the conventional approach to the assessment of diagnostic tests. In these trials patients were randomized to one test or the other (i.e., they received only one test). In general, when comparing diagnostic tests, each patient would receive both tests and the test results would be compared.
In a 2017 RCT, Min reported on 152 patients (142 were included in the analysis) that underwent crossover colonoscopies with white light endoscopy and linked color imaging (LCI), which uses narrow-band short-wavelength light and WL, randomized for order.\textsuperscript{[19]} The sensitivities in the white light and LCI groups were significantly different, at 73\% and 91\%, respectively. Negative predictive value was not reported.

In a 2016 RCT, Klare randomized 380 patients to the NBI arm or the high-definition white light arm.\textsuperscript{[20]} Accuracy was 73.7\% and 79.2\%, sensitivity was 82.4\% and 79.8\%, and negative predictive value was 75.5\% and 73.4\% in the NBI and white light arms, respectively. These values were not significantly different between arms.

Adler published trials in 2008 and 2009. The first trial enrolled 401 participants where the majority of the patients (89\%) were enrolled for a diagnostic colonoscopy and evaluated by expert endoscopists (>500 patients per provider).\textsuperscript{[14]} The second trial enrolled 1,256 participants evaluated with a screening colonoscopy in a private practice setting by six endoscopists with substantial lifetime experience (>10,000 total colonoscopies).\textsuperscript{[15]} Both trials randomized participants to receive NBI or white-light colonoscopy; neither trial showed a benefit of NBI over white-light for overall polyp detection rate.

In a similar study, with the same conclusion, Rex (2007) enrolled 434 participants, in a population split between 60\% screening colonoscopy and 40\% returning for surveillance.\textsuperscript{[17]} Each participant was randomized to either NBI or white-light colonoscopy. No benefit of NBI for the detection of adenomas was observed over white-light colonoscopy.

Kaltenback (2008) randomized 434 participants to receive both NBI and a white-light colonoscopy, or two white-light colonoscopies. Participants were screened by experienced endoscopists. With the first test, all visible polyps were removed, then the second test was performed to pick up any additional “missed” polyps; from this difference, the polyp miss rate was calculated. The major limitation with this method is that removing polyps with the first test eliminates the opportunity for the second test to “miss” any polyps which were already removed. NBI did not improve what was termed the “neoplasm miss rate” compared with white light.\textsuperscript{[16]}

Inoue (2008), in a randomized, controlled trial of 243 patients in Japan, presented data showing that NBI did improve overall adenoma detection rates over conventional colonoscopy, as well as improving the number of small (<5 mm) adenomas detected, while the number of patients with at least one adenoma remained the same.\textsuperscript{[21]} Participants in this trial had a previous positive colonoscopy or positive fecal occult blood test; approximately 80\% were undergoing polyp surveillance. All testing was performed at an endoscopy center by six experienced endoscopists. Differences in results may be attributed to different study populations and/or differences in the version of NBI system used.

In addition to the meta-analysis reviewed above, Sabbagh (2011) randomized 482 patients to NBI colonoscopy or conventional colonoscopy.\textsuperscript{[12]} They reported the overall rate of polyp detection was significantly higher in the conventional group compared with the NBI group; however, no significant differences were found in the mean number of polyps and the mean number of adenomas detected. A noted limitation of this study was the lack of tandem colonoscopy in both groups.

In a randomized controlled trial reported by Gross (2011), 100 patients undergoing routine screening and surveillance were randomized to receive tandem colonoscopies with standard...
definition white light (SDWL) and image-enhanced (HD-NBI) colonoscopy.\[22\] The main outcome measurement was the per-polyp false-negative ("miss") rate. Secondary outcomes were adenoma miss rate, and per-patient polyp and adenoma miss rates. Polyp and adenoma miss rates for SDWL colonoscopy were 57 % (60/105) and 49 % (19/39); those for image-enhanced colonoscopy were 31 % (22/72) and 27 % (9/33) (P = 0.005 and P = 0.036 for polyps and adenomas, respectively). Image-enhanced and SDWL approaches had similar per-patient miss rates for polyps (6/35 vs. 9/32, P = 0.27) and adenomas (4/22 vs. 8/20, P = 0.11). The authors concluded that utilization of multiple recent improvements in image-enhanced colonoscopy was associated with a reduced miss rate for all polyps and for adenomatous polyps. It is not known which individual feature or combination of image-enhancement features led to the improvement.

Kakol (2013) evaluated the usefulness of NBI for detection of missed polyps after colonoscopy comparing white light (WL) to NBI.\[23\] After initial colonoscopy 253 patients were randomized to a second colonoscopy with either NBI or WL. Authors found no significant difference between missed polyps or adenomas between groups.

East (2012) reported on 214 patients who were randomized to examination with either NBI or WL in order to determine whether NBI would enhance adenoma detection in high-risk patients.\[24\] High risk was defined as a patient with a history of three or more adenomas on last colonoscopy, colon cancer, and positive fecal occult blood test. There were no significant differences observed in detection of either polyps or adenomas between groups.

In 2014, Leung evaluated a new generation of NBI (190-NBI), with twice the brightness of previous versions, upon adenoma detection compared to HD-WL\[25\] colonoscopy. A total of 360 patients who were scheduled for colonoscopy for symptoms, screening, or surveillance were recruited to the study. Patients were randomized to receive either NBI or WL upon colonoscopy withdrawal. The primary outcomes were adenoma and polyp detection rates. Significantly higher adenoma and polyp detection rates with 190-NBI were reported compared to HD-WL (adenoma: 48.3% vs. 34.4%, P=0.01; polyp: 61.1% vs. 48.3%, P=0.02). However, there were no differences in adenoma miss rates between groups (21.8% vs. 21.2%).

In 2014, Wallace published results an RCT which compared NBI to standard colonoscopy and found no differences between groups.\[26\] A total of 522 patients were randomized and 927 total polyps were analyzed. No differences were observed in adenoma detection rate or diagnostic accuracy, regardless of polyp size.

Several randomized trials addressed both total detection rate and differentiation of neoplastic from nonneoplastic lesions.

Pohl conducted a randomized multicenter trial in 2009 of virtual chromoendoscopy with the “Fujinon intelligent colour enhancement” system (FICE or NBI) versus standard colonoscopy with targeted indigocarmine chromoscopy.\[27\] This German trial included 764 patients in the final analysis and reported that FICE/NBI was not superior to control for overall adenoma detection rates; it was comparable on the differentiation of neoplastic and non-neoplastic lesions. The sensitivity of FICE/NBI was 92.7% versus 90.4% for the control.

Rastogi (2011) reported on a randomized controlled trial of 630 subjects who were randomized to undergo colonoscopy with standard-definition white-light (SD-WL), HD-WL, or NBI.\[28\] The proportion of subjects with adenomas was 38.6% with SD-WL compared with 45.7% with HD-WL and 46.2% with NBI (P = .17 and P = .14, respectively). Adenomas detected per subject
were 0.69 with SD-WL compared with 1.12 with HD-WL and 1.13 with NBI (P = .016 and P = .014, respectively). HD-WL and NBI detected more subjects with flat and right-sided adenomas compared with SD-WL (all P values <.005). NBI had a superior sensitivity (90%) and accuracy (82%) to predict adenomas compared with SD-WL and HD-WL (all P values <.005). The authors concluded there was no difference in the proportion of subjects with adenomas detected with SD-WL, HD-WL, and NBI. However, HD-WL and NBI detected significantly more adenomas per subject (>60%) compared with SD-WL. NBI had the highest accuracy in predicting adenomas in real time during colonoscopy.

Additional RCTs were identified[29-31]; however, these studies contained several methodological flaws in that they only reported on the accuracy of the NBI system in the in vivo evaluation of colonic polyps. In addition, none of the studies evaluated the impact of this technology on outcomes including whether or not there would be an improvement in the selection of polyps for removal during colonoscopy. Furthermore, subsequent RCTs[32] demonstrate no differences in polyp detection rate of NBI compared to WL.

**CHROMOENDOSCOPY**

**Systematic Reviews**

Azizi (2018) performed a systematic review comparing white light endoscopy and chromoendoscopy for identifying dysplastic or cancerous lesions in patients with ulcerative colitis without primary sclerosing (PSC) or Crohn's disease (CD).[33] Studies were included if they reported on colonoscopy detection rates of dysplasia and cancers in UC without involvement of PSC or CD. Ten studies met inclusion criteria; most were of moderate quality. Publication bias was not assessed due to the low number of publications per incidence outcome. A meta-analysis of the five studies reporting overall pick-up rate of dysplastic/cancerous lesions on WLE random biopsies calculated showed a pooled rate of 5.6%. Only one study reported on the use of chromoendoscopy for ulcerative colitis patients without PSC. The reported pick-up rate of dysplastic lesions in this study was 7%.

In 2016, Brown updated their 2010 Cochrane review that compared chromoendoscopy and conventional colonoscopy for the detection of colorectal lesions in individuals at increased risk of colorectal neoplasia due to family history, previous polyp detection, or previous CRC resection.[34,35] The review excluded studies of individuals with IBD or a known polyposis syndrome. Seven RCTs (2727 participants) were included, five of which were used for a meta-analysis. All of these studies were published prior to 2012. The review found that chromoendoscopy was likely to yield more people with at least one neoplastic lesion (odds ratio (OR) 1.53, 95% confidence interval (CI) 1.31 to 1.79; seven trials; 2727 participants), and significantly more people with three or more neoplastic lesions were also detected, but only when studies that used high-definition colonoscopy in the control group were excluded (OR 4.63, 95% CI 1.99 to 10.80; two trials; 519 participants). None of the included studies reported any adverse events related to the use of the contrast dye. However, all the trials had some methodological drawbacks, and all were graded as low quality. In addition, some of the included studies were underpowered and significant heterogeneity was present between the included studies (variability of the colonoscopes used in the studies and differences in dye-spraying technique). There are also differences in the study inclusion criteria between the included studies).

Representative trials included in the Cochrane review are described below.

**Randomized Controlled Trials**
Vleugels (2018) randomized patients undergoing dysplasia surveillance for longstanding ulcerative colitis at five centers in the Netherlands and the UK to receive autofluorescence imaging or chromoendoscopy. Patients were eligible if they were age 18 years or older and were undergoing dysplasia surveillance after a diagnosis of extensive colitis at least eight years before the study start or left-sided colitis at least 15 years before the study start. Each group contained 105 patients. Primary outcomes were the proportion of patients in whom at least one dysplastic lesion was detected and the mean number of dysplastic lesions per patient. Dysplasia was detected in 12% and 19% of patients in the autofluorescence and chromoendoscopy groups, respectively. The mean number of detected dysplastic lesions per patient was 0.13 (SD 0.37) and 0.37 (SD 1.02) for autofluorescence and chromoendoscopy, respectively. Two and three adverse events were reported in the autofluorescence and chromoendoscopy groups, respectively. Autofluorescence imaging did not meet criteria for proceeding to a large non-inferiority trial.

In 2011, Pohl in Germany published a large RCT comparing pancolonic chromoendoscopy with indigo carmine dye with standard colonoscopy. The study included patients presenting for primary CRC screening (51%) and patients undergoing diagnostic colonoscopy (49%). Patients with known IBD, overt bleeding, polyposis syndromes, or a history of surgical resection were excluded. A total of 1024 patients were randomized, and 16 dropped out, leaving 496 patients in the chromoendoscopy group and 512 patients in the standard colonoscopy (i.e., control) group. The mean extubation time was 11.6 minutes in the chromoendoscopy group and 10.1 minutes in the standard colonoscopy group; the difference between groups was statistically significant (p<0.001). The primary study outcome, the proportion of patients with adenomas, differed significantly between groups (p=0.002). A total of 223 patients (46.2%) in the chromoendoscopy group and 186 (36.3%) in the standard colonoscopy group had at least one adenoma identified.

In 2010, one large randomized trial involving 660 average-risk patients was conducted at four centers in the United States. Those eligible for inclusion had an average risk of CRC, were aged 50 years and older, and were undergoing screening colonoscopy for the first time. Participants were randomized to undergo chromoendoscopy with indigo carmine dye (n=321) or standard colonoscopy (n=339). The primary outcomes were the proportion of patients with at least one adenoma and the mean number of adenomas per patient, which was then compared between groups. No significant between-group differences were noted for either outcome. A total of 178 (55.5%) subjects in the chromoendoscopy group and 164 (48.4%) subjects in the standard colonoscopy group had one or more adenomas (p=0.07). The mean number of adenomas per subject that were less than 5 mm in diameter differed significantly between the two groups, which was 0.8 in the chromoendoscopy group and 0.7 in the standard endoscopy group (p=0.03). However, this difference did not reach statistical significance; nor was there a statistically significant difference between groups in the number of larger adenomas. The mean number of adenomas per subject that were 10 mm or larger was 0.11 in the chromoendoscopy group and 0.12 in the standard colonoscopy group (p=0.70). A total of 39 (12%) subjects in the chromoendoscopy group and 49 (15%) subjects in the standard colonoscopy group had three or more adenomas; the difference between groups was not statistically significant (p=0.40). The authors stated that the high rate of adenoma detection in both groups may have been due to the use of high-definition colonoscopy.

The trial also reported differences in lesion detection rate by size of lesion. For lesions 5 mm or larger, 151 (30.4%) patients in the chromoendoscopy group and 119 (23.2%) patients in the standard colonoscopy group were found to have at least one adenoma; the difference between
groups was statistically significant (p=0.012). For lesions 10 mm or larger, 64 (12.9%) patients in the chromoendoscopy group and 48 (9.4%) patients in the standard colonoscopy group had at least one adenoma (p=0.092). The difference between groups in the detection of adenomas 10 mm or larger did not differ significantly; the study may have been underpowered for this analysis.

In 2008, Stoffel published findings of a study with five sites in the United States, Canada, and Israel.[39] Eligibility criteria included a personal history of CRC or at least three colorectal adenomas. The study involved back-to-back colonoscopies, the first of which was a standard colonoscopy with removal of all visualized polyps. Patients were then randomized to a second standard colonoscopy with intensive inspection (n=23) or chromoendoscopy (n=27). During the first colonoscopy, 17 of 50 (34%) patients had adenomas identified: 11 of 23 (48%) in the intensive inspection group and 6 (27%) in the chromoendoscopy group (p not reported). During the second colonoscopy, additional adenomas were found in 4 of 23 (17%) in the intensive inspection group and 12 of 27 (44%) in the chromoendoscopy group (p not reported). The mean size of adenomas found on the second examination was 3.2 mm in the intensive inspection group and 2.7 mm in the chromoendoscopy group. This compared with a mean size of 3.6 mm in the intensive inspection group and 4.7 mm in the chromoendoscopy group during the first examination. In a multivariate analysis, use of chromoendoscopy was significantly associated with an increased likelihood of finding at least one additional adenoma on the second examination (p=0.04).

Le Rhun published findings of a French study in 2006 involving 203 patients with a history of familial or personal colonic neoplasia or alarm symptoms (e.g., change in bowel habit, abdominal pain) after age 60 years.[40] Patients were randomized to standard colonoscopy (n=100) or high-resolution colonoscopy with chromoendoscopy (n=103). In the chromoendoscopy group, each segment of the colon was examined before and after spraying indigo carmine dye. The primary end point of total number of adenomas per patient did not differ significantly between groups. Mean values (SD) were 0.5 (0.9) in the standard colonoscopy group and 0.6 (1.0) in the chromoendoscopy group. The number of flat adenomas (at least 5 mm) per patient also did not differ significantly between groups; there was a mean (SD) of 0.04 (0.20) in the standard colonoscopy group and 0.10 (0.39) in the chromoendoscopy group (p=0.17).

VIRTUAL CHROMOENDOSCOPY

Systematic Reviews

A meta-analysis by Omata published in 2014 compared the rate of polyp detection by virtual chromoendoscopy (i.e., FICE or i-scan) with white-light colonoscopy.[41] The review included patients of all risk levels and was limited to RCTs. Five trials on FICE/i-scan met eligibility criteria and the analysis did not find a significantly higher detection rate with virtual chromoendoscopy. The pooled relative risk of adenoma/neoplasia detected by virtual chromoendoscopy versus conventional chromoendoscopy was 1.09 (95% CI, 0.97 to 1.23; p>0.05).

Randomized Controlled Trials

Kidambi (2018) randomized 740 patients undergoing screening and surveillance for colorectal neoplasia to receive colonoscopies with i-scan or with standard high-definition white-light.[42] Endoscopists were permitted to switch between i-scan and high-definition white-light imaging
to confirm polyps. Polyps were collected and analyzed by histology. The primary outcome was adenoma detection rate (ADR, proportion of subjects with at least one adenoma of any size). Intention to treat and per-protocol analyses were performed. ADR was significantly higher in the i-scan group for both the intent to treat and per-protocol analyses, with values of 47.2% and 47.6% in the i-scan group and 37.7% and 37.2% in the standard group, respectively. However, there was inconsistency across endoscopists. Secondary analyses showed that increased ADR was associated with improved detection of diminutive flat adenomas in the right colon. The groups had significantly different rates of neoplasia detection (i-scan, 56.4%; standard, 46.1%; p=0.005), but not detection of sessile serrated polyps.

Nonrandomized Studies

In 2016, Albrecht assessed the sensitivity, specificity, and positive and negative predictive values of i-scan. A total of 298 images of colonic lesions were assessed by endoscopists after undergoing a dedicated training. The sensitivity was 94.2% and the specificity was 90.9%. The positive predictive value was 87.5% and the negative predictive value was 95.9%. The intraobserver agreement was 0.9301.

In 2014, a large study using modified back-to-back designs in patients undergoing screening colonoscopy was conducted by Chung in South Korea, and included 1650 adults at average risk of CRC, who were randomly divided across three groups.[43] During the colonoscopy, the endoscope was fully inserted and each of three colonic segments (ascending, transverse, descending) was inspected twice during withdrawal. Participants received first withdrawal with narrow-band imaging (NBI), virtual chromoendoscopy using FICE, or white-light colonoscopy (n=550 each group). White light was used in all groups for the second inspection. Ninety-one patients (5.5%) were excluded from analysis due to inadequate bowel preparation. For the primary outcome of adenoma detection rate, no statistically significant difference was found among the three groups. The percentage of patients with at least one adenoma was 24.5% in the NBI group, 23.6% in the FICE group, and 25.3% in the white-light group (p=0.75). Moreover, the mean number of adenomas per patient was 0.35 in the NBI group, 0.36 in the FICE group, and 0.37 in the white-light group (p=0.59). The adenoma miss rate, defined as an adenoma identified only during the second inspection, was 22.9% in the NBI group, 26.0% in the FICE group, and 20.8% in the white-light–only group; a difference that was not statistically significant (p=0.30). The mean size of the missed adenomas was 3.6 mm, which was smaller than the mean size of adenomas found during the first withdrawal, which was 4.4 mm.

A study using a modified back-to-back colonoscopy design was published in 2012 by Kiriyama in Japan.[44] The study included 102 consecutive patients with increased risk of colon cancer who received virtual chromoendoscopy using FICE and white-light colonoscopy in random order. Patients were eligible for study inclusion if they had been referred for a colonoscopy following sigmoidoscopy or for postoperative surveillance after anterior resection. Those with known IBD, bleeding, and polyposis syndrome were excluded; the right-sided colon was examined in the remaining patients. All lesions identified on either examination were removed, and specimens were sent for evaluation. Two patients were excluded from the analysis because insertion was not possible, leaving 100 patients in the analysis. A total of 110 lesions were detected. Of these, 65 lesions were detected using FICE and 45 with white light; the difference in the number of detected lesions did not differ significantly between groups. Most of the lesions detected were neoplastic; of these, 59 (91%) were found using FICE and 38 (84%) were found with white-light colonoscopy. The miss rate was defined as the proportion of total lesions in that grouping that were detected on the second examination. The miss rate for all
polyps with FICE (12/39 lesions [31%]) was significantly less than that with white light (28/61 lesions [46%]) (p=0.03). Twenty-six of 59 (44%) neoplastic lesions detected by FICE and 14 of 38 (37%) of neoplastic lesions detected by white-light colonoscopy were at least 5 mm in size. For neoplastic lesions larger than 5 mm, there was no statistically significant difference between the FICE and white-light examinations in terms of the number of lesions detected.

In 2010, Cha evaluated South Korean patients at increased risk of CRC due to a personal history of polyps or gastrointestinal symptoms.[45] A total of 135 patients underwent colonoscopy, and seven were excluded due to poor bowel preparation or diagnosis of colon cancer or intestinal disease. Thus, 128 patients were randomized to white-light colonoscopy (n=65) or virtual chromoendoscopy with FICE (n=63). The overall percentage of adenomas and the overall number of polyps did not differ significantly between groups. A total of 31 patients (49.2%) in the FICE group and 23 (35.4%) in the white-light group were found to have one or more adenomas (p=0.12). The mean number of adenomas identified per patient was also similar between groups: 1.39 in the FICE group and 1.96 in the white-light group (p=0.46). The number of adenomas less than 5 mm in size (the primary study outcome) differed significantly between groups. A total of 28 (44.4%) of patients in the FICE group and 14 (21.5%) in the white-light group (p=0.006) were found to have adenomas between 0 and 5 mm. All adenomas identified were low grade and no complications were reported in either group.

A 2010 study by Chung included 359 asymptomatic patients receiving screening colonoscopies.[46] All received back-to-back examinations with white-light colonoscopy or FICE in random order (n=181 received white light first, n=178 received FICE first). In the initial colonoscopy, a total of 60 (33.7%) of patients in the FICE group and 55 (30.4%) in the white-light group were found to have at least one adenoma; the difference between groups was not statistically significant (p=0.74). The adenoma miss rate was 6.6% in the FICE group and 8.3% in the white-light group; the difference in miss rates was not statistically significant (p=0.59). All of the missed adenomas were low grade and nonpedunculated. All but one (which was 6 mm) were 5 mm or less in size. In both Chung studies, virtual chromoendoscopy was not found to improve the rate of adenoma detection compared with white-light endoscopy and did not identify more large adenomas.

A 2009 industry-supported multicenter RCT by Pohl in Germany compared FICE and targeted standard chromoendoscopy using indigo carmine stain.[47] The study enrolled 871 patients presenting for screening (57%) or diagnostic (43%) colonoscopy. All patients were examined using high-resolution zoom endoscopes. Patients in the group receiving standard chromoendoscopy underwent withdrawal using white-light colonoscopy. Indigo carmine was applied using a spray catheter through the working channel of the colonoscope for further assessment of any lesions that were identified. In the FICE group, withdrawal was performed using FICE at the preset for examining colorectal mucosa. Data were available for analysis on a total of 764 patients (368 in the FICE group, 396 in the standard chromoendoscopy group); 107 patients were excluded for poor bowel preparation, incomplete colonoscopy, or incomplete documentation. A total of 131 (35.6%) patients in the FICE group and 140 (35.4%) patients in the standard chromoendoscopy group had at least one adenoma; the difference between groups was not statistically significant (p=1.0). The number of small adenomas (here defined as no more than 10 mm) did not differ significantly between groups (p=0.41). The proportion of large adenomas greater than 10 mm identified in the two groups was not reported. The proportion of patients with carcinoma was small in both groups and did not differ significantly; 12 (3.3%) in the FICE group and 12 (3.0%) in the standard chromoendoscopy group (p=0.85).

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
PRACTICE GUIDELINE SUMMARY

U.S. MULTI-SOCIETY TASK FORCE ON COLORECTAL CANCER

This consensus-based guideline on colonoscopy surveillance after screening and polypectomy, published in 2012, stated that chromoendoscopy and narrow-band imaging may enable endoscopists to accurately determine if lesions are neoplastic, and if there is a need to remove them and send specimens to pathology. The guideline noted that, at this point, these technologies have not been studied in surveillance cohorts and therefore do not have an impact on surveillance interval.\textsuperscript{[48]} The task force published evidence based recommendations for colorectal cancer screening in 2017.\textsuperscript{[49]} These recommendations do not include in vivo analysis of colorectal polyps.

AMERICAN GASTROENTEROLOGICAL ASSOCIATION (AGA)

In 2008, the American Gastroenterological Association (AGA) published a technology assessment of image-enhanced endoscopy, which mentions optical and electronic devices potentially playing a role in colon screening in the future, but currently, more data are needed.\textsuperscript{[50]} In a 2010 position statement regarding diagnosis of colorectal neoplasia in patients with inflammatory bowel disease, the AGA stated, “Additional studies are needed to evaluate the efficiency of other imaging methods, such as narrow band imaging and confocal endomicroscopy, in detecting dysplasia.”\textsuperscript{[51]}

AMERICAN COLLEGE OF GASTROENTEROLOGY

In 2018, the American College of Gastroenterology (ACG) published an evidence based clinical guideline on the management of Crohn’s Disease in adults.\textsuperscript{[52]} The guideline makes the following statements regarding adjunct colonoscopy technologies:

- In patients at particularly high risk for colorectal neoplasia (e.g., personal history of dysplasia, primary sclerosing cholangitis), chromoendoscopy should be used during colonoscopy, as it may increase the diagnostic yield for detection of colorectal dysplasia, especially compared with standard-definition white light endoscopy (conditional recommendation, low level of evidence).
- For patients undergoing surveillance colonoscopy there is insufficient evidence to recommend universal chromoendoscopy for IBD colorectal neoplasia surveillance if the endoscopist has access to high-definition white light endoscopy (conditional recommendation, moderate level of evidence)
- Narrow-band imaging should not be used during colorectal neoplasia surveillance examinations for Crohn’s disease (conditional recommendation, very low level of evidence)

SUMMARY

More research is needed to know whether in vivo assessment of colorectal polyps using various imaging systems as adjuncts to colonoscopy improves health outcomes. There is not enough research to show whether there would be an improvement in the selection of polyps for removal during colonoscopy. Therefore, in vivo analysis of colorectal polyps using any system is considered investigational.
REFERENCES


12. Sabbagh, LC, Reveiz, L, Aponte, D, de Aguiar, S. Narrow-band imaging does not improve detection of colorectal polyps when compared to conventional colonoscopy: a randomized controlled trial and meta-analysis of published studies. BMC gastroenterology. 2011;11:100. PMID: 21943365


November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


44. Kiriyama, S, Matsuda, T, Nakajima, T, Sakamoto, T, Saito, Y, Kuwano, H. Detectability of colon polyp using computed virtual chromoendoscopy with flexible spectral imaging
color enhancement. *Diagnostic and therapeutic endoscopy.* 2012;2012:596303. PMID: 22474404


---

### CODES

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>44799</td>
<td>Unlisted procedure, small intestine</td>
</tr>
<tr>
<td></td>
<td>45399</td>
<td>Unlisted procedure, colon</td>
</tr>
<tr>
<td></td>
<td>88375</td>
<td>Optical endomicroscopic image(s), interpretation and report, real-time or referred, each endoscopic session</td>
</tr>
</tbody>
</table>

**Date of Origin:** June 2002

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Orthopedic Applications of Stem Cell Therapy

Effective: April 1, 2018

Next Review: October 2018
Last Review: March 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Mesenchymal stem cells (MSCs) are multipotent cells (also called “stromal multipotent cells”) that possess the ability to differentiate into various tissues including organs, trabecular bone, tendon, articular cartilage, ligaments, muscle, and fat. Potential uses of MSCs for orthopedic applications include treatment of damaged bone, cartilage, ligaments, tendons, and intervertebral discs.

MEDICAL POLICY CRITERIA

I. Mesenchymal stem cell therapy is considered investigational for all orthopedic applications, including but not limited to use in repair or regeneration of musculoskeletal tissue.

II. Allograft bone products containing viable stem cells are considered investigational for all orthopedic applications, including but not limited to demineralized bone matrix (DBM) with stem cells.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.
BACKGROUND

MSCs are associated with the blood vessels within bone marrow, synovium, fat, and muscle where they can be mobilized for endogenous repair, as occurs with healing of bone fractures. Stimulation of endogenous MSCs is the basis of procedures such as bone marrow stimulation (e.g., microfracture) and harvesting/grafting of autologous bone for fusion. Bone-marrow aspirate is considered to be the most accessible source and, thus, the most common place to isolate MSCs for treatment of musculoskeletal disease. However, harvesting MSCs from bone marrow requires an additional procedure that may result in donor-site morbidity. In addition, the number of MSCs in bone marrow is low, and the number and differentiation capacity of bone marrow-derived MSCs decreases with age, limiting their efficiency when isolated from older patients.

Tissues such as muscle, cartilage, tendon, ligaments, and vertebral discs show limited capacity for endogenous repair. Tissue engineering techniques are being developed to improve the efficiency of repair or regeneration of damaged musculoskeletal tissues. Tissue engineering focuses on the integration of biomaterials with MSCs and/or bioactive molecules such as growth factors. In vivo, the fate of stem cells is regulated by signals in the local 3-dimensional microenvironment from the extracellular matrix and neighboring cells. It is believed that the success of tissue engineering with MSCs will also require an appropriate 3-dimensional scaffold or matrix, culture conditions for tissue-specific induction, and implantation techniques that provide appropriate biomechanical forces and mechanical stimulation. Given that each tissue type requires different culture conditions, induction factors (e.g., signaling proteins, cytokines, growth factors), and implantation techniques, each preparation must be individually examined. The ability to induce cell division and differentiation, without adverse effects such as the formation of neoplasms, remains a significant concern.

The U.S. Food and Drug Administration (FDA) stated:

“Cell-based therapies are one of the most rapidly advancing approaches intended to repair, replace, restore, or regenerate cells, tissues and organs. They can be applied to damage caused by disease, injury, or aging. Many cell-based therapies use immature cells (stem cells) that are expanded outside of the body. The expanded cells are sometimes used in their immature state, but they are often manufactured into more mature cells before they are given to a patient. The resulting cells are intended to repair cell or tissue damage (efficacy) without unintended serious consequences such as tumors, severe immune reactions, or unwanted tissue development (safety). Manufacturing of large numbers of cells outside the natural environment of the human body may lead to ineffective or dangerous cells, so it is important to understand and carefully control the production process and to define measures that reliably predict safety and efficacy of the cell-based products.”[1]

REGULATORY STATUS
Concentrated autologous MSCs do not require approval by the U.S. Food and Drug Administration (FDA).

Demineralized bone matrix (DBM), which is processed allograft bone, is considered minimally processed tissue and does not require FDA approval. At least 4 commercially available DBM products are reported to contain viable stem cells:

- Allostem® (AlloSource) is partially demineralized allograft bone seeded with adipose-derived MSCs
- Map3™ (RTI surgical) contains cortical cancellous bone chips, DBM, and multipotent adult progenitor cells
- Osteocell Plus® (NuVasive): an allograft cellular bone matrix containing native MSCs.
- Trinity Evolution Matrix™ (Orthofix): an allograft that is processed and cryopreserved to maintain viable adult MSCs and osteoprogenitor cells.

Whether these products can be considered minimally manipulated tissue is debated. A product would not meet the criteria for FDA regulation part 1271.10 if it is dependent upon the metabolic activity of living cells for its primary function. Otherwise, a product would be considered a biologic product and would need to demonstrate safety and efficacy for the product’s intended use with an investigational new drug and Biologics License Application (BLA).

Other products contain DBM and are designed to be mixed with bone marrow aspirate. Some of the products that are currently available are:

- Fusion Flex™ (Wright Medical): a dehydrated moldable DBM scaffold that will absorb autologous bone marrow aspirate.
- Ignite® (Wright Medical): an injectable graft with DBM that can be combined with autologous bone marrow aspirate.

Other commercially available products are intended to be mixed with bone marrow aspirate and have received 510(k) clearance, such as:

- CopiOs sponge or paste (Zimmer): synthetic bone graft material consisting of mineralized, lyophilized collagen.
- Collage™ Putty (Orthofix): Composed of type-1 bovine collagen and beta Tri-calcium phosphate.
- Vitoss® (Stryker, developed by Orthovita): composed of beta tricalcium phosphate.
- nanOss® Bioactive (RTI Surgical, developed by Pioneer Surgical): nanostructured hydroxyapatite and an open structured engineered collagen carrier.

No products using engineered MSCs have been approved by the FDA for orthopedic applications.

In 2008, the FDA determined that the mesenchymal stem cells sold by Regenerative Sciences for use in the Regenexx™ procedure would be considered drugs or biological products and thus require submission of a New Drug Application (NDA) or Biologics Licensing Application (BLA) to the FDA. In 2014, a federal appellate court upheld FDA’s power to regulate adult stem cells as drugs and biologics and ruled that the Regenexx cell product fell within FDA’s authority to regulate human cells, tissues, and cellular and tissue-based products (HCT/Ps) (Section 351).[2] To date, no NDA or BLA has been approved by the FDA for this product. As of 2015,
the expanded stem cell procedure is only offered in the Cayman Islands. Regenexx™ network facilities in the U.S. provide same-day stem cell and blood platelet procedures, which do not require FDA approval.[3]

**EVIDENCE SUMMARY**

At this time, the literature consists mainly of articles describing the potential of stem cell therapy for orthopedic applications in humans, along with basic science experiments on sources of mesenchymal stem cells (MSCs), regulation of cell growth and differentiation, and development of scaffolds.[4] Although the evidence base has been steadily increasing, authors indicate that the technology is in an early stage of development. In order to assess the safety and efficacy of orthopedic applications of MSCs and allograft bone products, such as demineralized bone matrix, high quality randomized trials are required that compare health outcomes with versus without the use of these products.

**CARTILAGE DEFECTS**

In 2016, Cui published a systematic review on 18 studies looking at the effect of MSC in treating patients with osteoarthritis.[5] MSC treatment in patients with KOA showed continual efficacy for 24 months compared with their pretreatment condition. Effectiveness of MSCs was improved at 12 and 24 months post-treatment, compared with at 3 and 6 months. There was no dose response association in the MSCs numbers. This review only included four randomized trials while the remaining 14 studies were non-randomized and had methodological limitations.

In 2015, Xu published a meta-analysis on the effect of MSCs for articular cartilage degeneration treatment, including 11 controlled trials (N=558). No critical appraisal of the quality of the included studies was reported. MSC treatment significantly improved the American Orthopedic Foot and Ankle Society Scale (Standard Mean Difference [SMD] 0.91; 95% confidence interval [CI], 0.52 to 1.29) and the Osteo-Arthritis Outcome Score (SMD, 2.81; 95% CI, 2.02 to 3.60).[6] Comprehensive evaluation indexes, such as the American Knee Society Knee Score System (SMD -0.12, 95% CI, -1.02 to 0.78), the Hospital for Special Surgery Knee Rating Scale (SMD, 0.24, 95% CI, -0.56 to 1.05) and the International Knee Documentation Committee (SMD, -0.21; 95% CI, -0.77 to 0.34), were no different between MSC use and other treatments. The reviewers concluded that there was no obvious advantage regarding the application of stem cells to treat cartilage injury, compared with other treatments.

In 2013, Filardo conducted a systematic review of mesenchymal stem cells for the treatment of cartilage lesions.[7] They identified 72 preclinical papers and 18 clinical reports. Of the 18 clinical reports, none were randomized, 5 were comparative, 6 were case series, and 7 were case reports. In 2 clinical studies the source of MSCs was adipose tissue, in 5 it was bone marrow concentrate, and in 11 studies the source of MSCs was bone marrow-derived. The authors reached the following conclusion:

“Despite the growing interest in this biological approach for cartilage regeneration, knowledge on this topic is still preliminary, as shown by the prevalence of preclinical studies and the presence of low-quality clinical studies. Many aspects have to be optimized, and randomized controlled trials are needed to support the potential of this biological treatment for cartilage repair and to evaluate advantages and disadvantages with respect to the available treatments.”
The source of MSCs may have an impact on outcomes, but this is not well understood and the available literature uses multiple different sources of MSC. Because of the uncertainty over whether these products are equivalent, the summary of the key evidence to date is grouped by source of MSC.

**Cartilage Defects: MSCs Expanded From Bone Marrow**

Since the systematic review by Filardo, one RCT was published. Wong, reported on the use of cultured MSCs in 56 patients with osteoarthritis who underwent medial opening-wedge high tibial osteotomy and microfracture of a cartilage lesion. Bone marrow was harvested at the time of microfracture and the MSCs were isolated and cultured. After 3 weeks, the cells were assessed for viability and delivered to the clinic, where patients received an intra-articular injection of MSCs suspended in hyaluronic acid (HA) or, for controls, intra-articular injection of HA alone. The primary outcome was the International Knee Documentation Committee (IKDC) score at 6 months, 1 year, and 2 years. Secondary outcomes were the Tegner and Lysholm scores through 2 years and the Magnetic Resonance Observation of Cartilage Repair Tissue (MOCART) scoring system by MRI at 1 year. All patients completed the 2-year follow-up. After adjusting for age, baseline scores, and time of evaluation, the group treated with MSCs showed significantly better scores on the IKDC (mean difference, 7.65 on 0-100 scale; \( p=0.001 \)), Lysholm (mean difference, 7.61 on 0-100 scale; \( p=0.02 \)), and Tegner (mean difference, 0.64 on a 0-10 scale; \( p=0.02 \)). Blinded analysis of MRI results found higher MOCART scores in the MSC group. The group treated with MSCs had a higher proportion of patients who had complete cartilage coverage of their lesions (32% vs 0%), greater than 50% cartilage cover (36% vs 14%) and complete integration of the regenerated cartilage (61% vs 14%). This study is ongoing and recruiting additional patients.

**Cartilage Defects: MSCs Concentrated From Bone Marrow**

A small RCT was recently published by Vega that assessed the efficacy of bone marrow derived MSCs as a treatment for knee osteoarthritis, randomizing 30 patients with chronic knee pain unresponsive to conservative treatments and showing radiological evidence of osteoarthritis. Fifteen patients were treated with allogeneic bone marrow MSCs by intra-articular injection, while 15 controls received intra-articular hyaluronic acid (HA). Clinical outcomes were followed for 1 year and included evaluations of pain, disability, and quality of life. Articular cartilage quality was assessed by quantitative magnetic resonance imaging T2 mapping. The MSC-treated patients displayed significant improvement in algofunctional indices versus the active controls. Quantification of cartilage quality by T2 relaxation measurements showed a significant decrease in poor cartilage areas, with cartilage quality improvements in MSC-treated patients.

Centeno reported a multicenter registry of patients treated with autologous stem cells, bone marrow concentrate, and platelet-rich plasma. This report focused on 102 patients (115 shoulders) diagnosed with either osteoarthritis of the shoulder or rotator cuff tears. Patients were treated with a protocol that included a hypertonic dextrose solution (prolotherapy) injection to create an inflammatory response several days prior to the bone marrow concentrate injection. The bone marrow concentrate injection included platelet-rich plasma and platelet lysate. Both DASH (Disabilities of the Arm, Shoulder, and Hand) score and numeric pain scores (NPS) decreased by about 50%, although the absolute decrease in the NPS was a very modest 0.9. Interpretation of these results is limited by the lack of a placebo control and blinding, subjective outcome measures, and the multiple treatments used, although it is
acknowledged that neither prolotherapy nor PRP appear to have efficacy on their own. Additional study with randomized and placebo-controlled trials is needed to evaluate this treatment protocol.

**Cartilage Defects: Adipose-Derived MSCs**

The literature on adipose-derived MSCs for articular cartilage repair is very limited, coming from two research groups in Korea. One of the groups appears to have been providing this treatment as an option for patients for a number of years and recently published a RCT that evaluated cartilage healing after high tibial osteotomy (HTO) in 52 patients with osteoarthritis of the medial compartment.\[^{11}\] Patients were randomly assigned to HTO with application of platelet-rich plasma (PRP) or HTO with application of PRP plus MSCs. MSCs from adipose tissue were obtained through liposuction from the buttocks. The tissue was centrifuged and the stromal vascular fraction mixed with PRP for injection. A total of 44 patients completed second look arthroscopy and 1- and 2-year clinical follow-up. There were statistically significant differences for PRP only versus PRP+MSC on the Knee Injury and Osteoarthritis Outcome Score (KOOS) subscales for pain (74±5.7 vs 81.2±6.9, p<0.001) and symptoms (75.4±8.5 vs 82.8±7.2, p=0.006). There were also statistically significant differences on the final pain score for the PRP only versus PRP+MSC groups (16.2±4.6 vs 10.2±5.7, p<0.001), but the Lysholm score, which is more scientifically proven, was not significantly different between the PRP only and PRP+MSC groups (80.6±13.5 vs 84.7±16.2, all respectively, p=0.36). Articular cartilage healing was rated as improved with MSCs following video review of second-look arthroscopy; blinding of this measure is unclear. There are a number of limitations of this study, including the small sample size, short duration of follow-up, and significant improvements on only some of the outcomes. All of the significant differences in outcomes were modest in magnitude, and as a result, there is uncertainty regarding the clinical significance of the findings.

This group also published a trial comparing treatment with adipose-derived MSCs, fibrin glue, and microfracture to microfracture alone.\[^{12}\] A total of 80 patients with a single International Cartilage Repair Society grade III/IV symptomatic cartilage defect on the femoral condyle were randomized to receive one of the treatments. The mean follow-up time was 27.4 months. At follow-up, the MSC + fibrin glue + microfracture group had significantly greater improvements in the Knee Injury and Osteoarthritis Outcome Score pain and symptom subscores than the microfracture alone group (P = .034 and .005, respectively). There were no significant differences between groups for the activities of daily living, sports and recreation, or quality of live subscores. Second-look arthroscopies were performed in 57 of the 80 patients, with no significant differences between groups. The lack of blinding in this study limits the conclusions that can be drawn from its results.

The remaining evidence is limited to reports on 3 small retrospective analyses from the same investigators, two for focal osteochondral lesions of the ankle\[^{13,14}\] and one for post-debridement knee osteoarthritis\[^{15}\], and a small phase I trial for severe osteoarthritis.\[^{16}\] Due to methodological limitations, these studies do not permit conclusions about the effectiveness and safety of adipose-derived MSCs. These limitations include small sample size, the lack of randomized treatment allocation, and the lack of prospective comparison of outcomes.

**Cartilage Defects: MSCs from Peripheral Blood**

A 2013 report described a small randomized controlled trial with autologous peripheral blood MSCs for focal articular cartilage lesions.\[^{17}\] Fifty patients with grade 3 and 4 lesions of the knee joint underwent arthroscopic subchondral drilling followed by 5 weekly injections of
hyaluronic acid (HA). Half of the patients were randomly allocated to receive injections of peripheral blood stem cells or no further treatment. There were baseline differences in age between the groups, with a mean age of 38 for the treatment group compared to 42 for the control group. The peripheral blood stem cells were harvested after stimulation with recombinant human granulocyte colony-stimulating factor, divided in vials, and cryopreserved. At 6 months after surgery, HA and MSC were re-administered over 3 weekly injections. At 18 months after surgery, second look arthroscopy on 16 patients in each group showed significantly (p=.022) higher histological scores (by about 10%) for the MSC group (1,066 vs. 957 by independent observers) while blinded evaluation of MRI showed a statistically significant (p=.013) higher morphologic score (9.9 vs. 8.5). There was no difference in International Knee Documentation Committee (IKDC) scores between the 2 groups at 24 months after surgery. It is uncertain how differences in patient age at baseline may have affected the response to subchondral drilling.

**Cartilage Defects: MSCs from Synovial Tissue**

Akgun reported a small (n=14) investigator-blinded RCT that compared matrix-induced autologous MSCs from synovial tissue versus matrix-induced autologous chondrocyte implantation (MACI).\(^{18}\) Both chondrocytes from cartilage and MSCs from synovia were harvested in an arthroscopic procedure, expanded in culture, and then cultured on a collagen membrane for two days. Implantation was performed with the cells facing the subchondral bone. Follow up evaluations were made through 24 months post-procedure. Outcomes on the KOOS subscales and the VAS pain score were statistically better in the MSC group than the MACI group (p < 0.05) at the six month follow up, although it is not clear if the difference observed would be considered clinically significant. Studies with larger samples sizes and follow-up supported by histological analyses are necessary to determine long-term outcomes of this treatment.

**Section Summary**

The evidence base on MSCs for repair of cartilage defects is limited to four small randomized studies and a number of small case series in which a variety of methods of MSC preparation were used. All four randomized studies reported an improvement in histological and morphologic outcomes, despite being harvested from different sources and compared against different control treatments. Three of these studies also reported an improvement in functional outcomes, although these varied between studies. In the RCT which found no improvement functional outcomes, peripheral blood stem cells were harvested following stimulation with recombinant human granulocyte colony-stimulating factor. The literature on adipose-derived MSCs includes a phase 1/2 study with cultured MSCs and a small RCT from a separate group in Asia that has been using uncultured MSCs as an adjunctive procedure in clinical practice. Comparisons between patients who have and have not received uncultured adipose-derived MSCs shows modest improvement in health outcomes that are of uncertain clinical significance. Potential for bias from non-blinded use of a novel procedure on subjective outcome measures is a major limitation of these studies. The phase I/II study of cultured MSCs from adipose tissue shows promising results for this technology. Additional studies in larger cohorts with longer follow-up is needed to evaluate the long-term efficacy and safety of the procedure.

**FUSION AND NON-UNION**
There is limited evidence on the use of allografts with stem cells for fusion of the extremities or spine or for the treatment of non-union. Eastlack reported outcomes from a series of 182 patients who were treated with anterior cervical discectomy and fusion using Osteocel Plus in a PEEK cage and anterior plating.[19] At 24 months, 74% of patients (180/249 levels treated) were available for follow-up. These patients had significant improvements in clinical outcomes; 87% of levels achieved solid bridging and 92% of levels had range of motion less than 3º. With 26% loss to follow-up at 24 months and lack of a standard of care control group, interpretation of these results is limited. One retrospective series from 2009 was identified on the use of Trinity MSC bone allograft for revision surgery of the foot and ankle.[20] Twenty-three patients were included who had undergone revision foot and/or ankle surgery for residual malunion, non-union, or significant segmental bone loss. Patients were followed to the point of radiographic and clinical union, which occurred at a median of 72.5 days for 21 of the 23 patients (91.3%). However, these outcomes do not permit conclusions because of a lack of a control group for comparison with patients who received stem-cell therapy.

Section Summary

Current evidence is insufficient to determine whether the use of stem cell results in superior outcomes such as higher fusion rates, or lower rates of reoperations and adverse events.

MENISCECTOMY

In 2014, Vangsness reported an industry-sponsored phase 1/2 randomized, double-blind, multicenter study of cultured allogeneic MSCs (Chondrogen™, Osiris Therapeutics) injected into the knee after partial meniscectomy.[21] The 55 patients were randomized to intra-articular injection of either 50´10^6 allogeneic MSCs, 150´10^6 allogeneic MSCs in HA, or HA vehicle control at 7 to 10 days after meniscectomy. The cultured MSCs were derived from bone-marrow aspirates from unrelated donors. At 2-year follow-up, 3 patients in the low-dose MSC group had significantly increased meniscal volume measured by MRI (with an a priori determined threshold of at least 15%) compared to none in the control group and none in the high-dose MSC group. There was no significant difference between the groups in the Lysholm Knee Scale. On subgroup analysis, patients with osteoarthritis who received MSCs had a significantly greater reduction in pain at 2 years compared with patients who received HA alone. This appears to be a post hoc analysis and should be considered preliminary. No serious adverse events were thought to be related to the investigational treatment.

Section Summary

Current evidence for the use of stem cells as an adjunct to meniscectomy is limited to a single preliminary RCT. The outcomes of this study must be validated in large, long-term, randomized controlled trials.

OSTEONECROSIS

Several randomized comparative trials have been identified that evaluated the use of MSCs for osteonecrosis of the femoral head.

Osteonecrosis: MSCs Expanded From Bone Marrow

In 2012, Zhao reported a randomized trial that included 100 patients (104 hips) with early stage femoral head osteonecrosis treated with core decompression and expanded bone marrow MSCs versus core decompression (CD) alone.[22] At 60 months after surgery, 2 of the
53 hips (3.7%) treated with MSCs continued to have progressive disease and underwent vascularized bone grafting, compared with 10 of 44 hips (23%) in the decompression group who had disease progression and underwent either vascularized bone grafting (n=5) or total hip replacement (n=5). In addition, treatment with MSC improved Harris Hip scores compared to CD and decreased the volume of the necrotic lesion of the hips preoperatively classified at stage IC, IIB, and IIC (P<0.05, respectively; stage IIA, P=0.06, respectively).

**Osteonecrosis: MSCs Concentrated From Bone Marrow**

A 2017 randomized, double-blind trial was conducted using autologous bone marrow concentrate in 38 patients with stage three osteonecrosis.[23] A control group of core decompression plus saline injection was compared to patients receiving core decompression plus BMAC implantation. The primary outcome was needing total hip replacement and secondary outcomes were clinical symptoms such as pain and functional ability. There was no difference between groups on any outcomes including total hip replacement requirements, clinical tests, or radiologic evidence. Another small trial randomized 40 patients (51 hips) with early stage femoral head osteonecrosis to core decompression plus concentrated bone marrow MSCs or core decompression alone.[24] Blinding of assessments in this small trial was not described. Harris Hip Score (HHS) was significantly improved in the MSC group (scores of 83.65 and 82.42; p<0.05) compared with core decompression (scores of 76.68 and 77.39). Kaplan-Meier analysis showed improved hip survival in the MSC group (mean of 51.9 weeks) compared with the core decompression group (mean of 46.7 weeks). There were no significant differences between the groups in the radiographic assessment or MRI results. The conflicting report of improvement via HHS compared to no observable improvement via MRI, may point to the need for study blinding to control for confounding bias toward treatment.

**Section Summary**

Two small studies reported improvement in the Harris Hip Score in patients with osteonecrosis of the femoral head treated with core decompression and MSCs, although it was not reported if the patients or investigators were blinded to the treatment group. Hip survival was significantly improved following treatment with either expanded or concentrated MSCs. The effect appears to be larger with expanded MSCs compared with concentrated MSCs. Additional studies with a larger number of patients are needed to permit greater certainty regarding the effect of this treatment on health outcomes.

**PRACTICE GUIDELINE SUMMARY**

Currently, there are no clinical practice guidelines from US professional societies that address the use of stem cells in orthopedics.

**AMERICAN ASSOCIATION OF ORTHOPAEDIC SURGEONS (AAOS)**

An informational statement from the AAOS states that stem cell procedures in orthopedics are still at an experimental stage; most musculoskeletal treatments using stem cells are performed at research centers as part of controlled, clinical trials, and results of studies in animal models provide proof-of-concept that in the future, similar methods could be used to treat osteoarthritis, nonunion of fractures, and bone defects in humans.[25]
There is not enough research to know if or how well mesenchymal stem cells (MSCs) or allograft bone products containing stem cells work to treat people with orthopedic conditions. No clinical guidelines based on research recommend MSC treatment or allograft bone products containing stem cells for people with orthopedic conditions. Therefore, use of stem cells for orthopedic applications is considered investigational.

**REFERENCES**

11. Koh, YG, Kwon, OR, Kim, YS, Choi, YJ. Comparative outcomes of open-wedge high tibial osteotomy with platelet-rich plasma alone or in combination with mesenchymal stem cell treatment: a prospective study. *Arthroscopy: the journal of arthroscopic &


23. Hauzeur, JP, De Maertelaer, V, Baudoux, E, Malaise, M, Beguin, Y, Gangji, V. Inefficacy of autologous bone marrow concentrate in stage three osteonecrosis: a
randomized controlled double-blind trial. *International orthopaedics*. 2017 Oct 07. PMID: 28988340


### CODES

**NOTE:** There are no specific codes for orthopedic applications of stem cell therapy. The appropriate CPT code for reporting this procedure is 20999, or the code for an unlisted procedure of the body area on which the procedure is performed.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>20999</td>
<td>Unlisted procedure, musculoskeletal system, general</td>
</tr>
<tr>
<td></td>
<td>21899</td>
<td>Unlisted procedure, neck or thorax</td>
</tr>
<tr>
<td></td>
<td>22899</td>
<td>Unlisted procedure, spine</td>
</tr>
<tr>
<td></td>
<td>23929</td>
<td>Unlisted procedure, shoulder</td>
</tr>
<tr>
<td></td>
<td>24999</td>
<td>Unlisted procedure, humerus or elbow</td>
</tr>
<tr>
<td></td>
<td>25999</td>
<td>Unlisted procedure, forearm or wrist</td>
</tr>
<tr>
<td></td>
<td>26989</td>
<td>Unlisted procedure, hands or fingers</td>
</tr>
<tr>
<td></td>
<td>27299</td>
<td>Unlisted procedure, pelvis or hip joint</td>
</tr>
<tr>
<td></td>
<td>27599</td>
<td>Unlisted procedure, femur or knee</td>
</tr>
<tr>
<td></td>
<td>27899</td>
<td>Unlisted procedure, leg or ankle</td>
</tr>
<tr>
<td></td>
<td>28899</td>
<td>Unlisted procedure, foot or toes</td>
</tr>
<tr>
<td></td>
<td>29999</td>
<td>Unlisted procedure, arthroscopy</td>
</tr>
<tr>
<td></td>
<td>38206</td>
<td>Blood-derived hematopoietic progenitor cell harvesting for transplantation, per collection; autologous</td>
</tr>
<tr>
<td></td>
<td>38230</td>
<td>Bone marrow harvesting for transplantation; allogeneic</td>
</tr>
<tr>
<td></td>
<td>38232</td>
<td>Bone marrow harvesting for transplantation; autologous</td>
</tr>
<tr>
<td></td>
<td>38241</td>
<td>Bone Marrow or blood-derived peripheral stem cell transplantation; autologous</td>
</tr>
</tbody>
</table>

**Date of Origin:** September 2011

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Radioembolization for Primary and Metastatic Tumors of the Liver

Effective: September 1, 2018

Next Review: July 2019
Last Review: July 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Small radioactive beads are delivered into the hepatic artery for treatment of liver tumors.

MEDICAL POLICY CRITERIA

Note: This policy only addresses radioembolization for the treatment of primary and metastatic tumors of the liver. This policy does not address other tumor locations, transarterial embolization (TAE) with non-radioactive agents, or transarterial chemoembolization (TACE), which may be considered medically necessary for treatment of liver tumors.

I. Radioembolization may be considered medically necessary for treatment of any of the following:

   A. Unresectable primary liver tumors (hepatocellular carcinoma [HCC])
   B. As a bridge to transplantation in primary HCC
   C. Unresectable hepatic metastases from neuroendocrine or colorectal tumors, or melanoma when either criteria 1, 2, or 3 are met:
1. Neuroendocrine tumors (carcinoid and noncarcinoid) when both of the following criteria are met:
   a. The disease is liver-dominant and diffuse (defined as tumor tissue spread throughout the affected organ) and symptomatic
   b. Systemic therapy has failed to control symptoms, or the patient is not a candidate for systemic therapy; or
2. Colorectal tumors, including but not limited to adenocarcinoma when both criteria (a and b) are met:
   a. The disease is liver-dominant, progressive, and diffuse (diffuse is defined as tumor tissue spread throughout the affected organ)
   b. The patient is refractory to or not a candidate for chemotherapy; or
3. Melanoma (ocular/uveal or cutaneous) when the disease is liver-dominant, progressive, and diffuse.

D. Unresectable primary intrahepatic cholangiocarcinoma

II. Radioembolization is considered investigational for all other indications not meeting the criteria above.

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

POLICY GUIDELINES

• It is critical that the list of information below is submitted for review to determine if the policy criteria are met. If any of these items are not submitted, it could impact our review and decision outcome.
  o Description of the planned therapy including the approach and the embolization agent to be used
  o Specific description of the disease including the following:
    ▪ Tumor type (primary vs. metastatic)
    ▪ Extent and location of disease including whether the tumor is liver-dominant, progressive, and diffuse, and the presence or absence of extra-hepatic disease
    ▪ For neuroendocrine metastases, description of the presence or absence of tumor-related symptoms
  o Rationale for the determination that the patient is not a surgical candidate or the tumor is unresectable
  o Prior treatments, if any, and tumor response
  o Rationale for the determination that the patient is not a candidate for initial or continued systemic therapy
  o For treatment of hepatocellular carcinoma, specify if whether treatment is proposed as a bridge to transplantation

• Neuroendocrine tumors are rare, slow-growing, hormone-secreting tumors that may occur in numerous locations in the body.\[1\] Neuroendocrine tumors include the following:
  o Carcinoid Tumors
  o Islet Cell Tumors (also known as Pancreatic Endocrine Tumors)
Neuroendocrine tumors may also be referred to by their location (e.g., pulmonary neuroendocrine tumors; gastroenteropancreatic neuroendocrine tumors)

Some appendiceal carcinoids, also called adeno carcinoids, goblet cell carcinoids or crypt cell carcinoids, have mixed histology, including elements of adenocarcinoma. While these biphasic tumors have both neuroendocrine and adenocarcinoma components, the National Comprehensive Cancer Network (NCCN) recommends they be managed according to colon cancer guidelines.

CROSS REFERENCES

1. Charged-Particle (Proton) Radiotherapy, Medicine, Policy No. 49
2. Intensity Modulated Radiation Therapy (IMRT) of the Abdomen and Pelvis, Medicine, Policy No. 139
3. Stereotactic Radiosurgery and Stereotactic Body Radiation Therapy, Surgery, Policy No. 16
4. Radiofrequency Ablation of Tumors (RFA), Surgery, Policy No. 92
5. Cryosurgical Ablation of Miscellaneous Solid Tumors, Surgery, Policy No. 132
6. Magnetic Resonance (MR) Guided Focused Ultrasound (MRgFUS) and High Intensity Focused Ultrasound (HIFU) Ablation, Surgery, Policy No. 139
7. Microwave Tumor Ablation, Surgery, Policy No. 189
8. Ablation of Primary and Metastatic Liver Tumors, Surgery, Policy No. 204

BACKGROUND

Hepatic tumors can arise either as primary liver cancer or by metastasis to the liver from other organs. Potentially curative local treatments include surgical resection with tumor-free margins, liver transplantation, ablative techniques, and external-beam radiation therapies. Unfortunately, most hepatic tumors are unresectable at diagnosis, due either to their anatomic location, size and number of lesions, concurrent nonmalignant liver disease, or insufficient hepatic reserve.

The use of external beam radiotherapy, 3-D or more advanced radiotherapy approaches such as intensity-modulated radiotherapy (IMRT) may be of limited use in patients with diffuse, multiple lesions due to the low tolerance of normal liver to radiation compared to the higher doses of radiation needed to kill the tumor.

Various nonsurgical and non-external irradiation based ablative techniques have been investigated that seek to cure or palliate unresectable hepatic tumors by improving locoregional control. These techniques rely on extreme temperature changes, particle and wave physics (microwave or laser ablation), or arterial embolization therapy including chemoembolization, bland embolization, or radioembolization.

Radioembolization, formerly referred to as selective internal radiation therapy or “SIRT”, is the intra-arterial delivery of small beads (microspheres) impregnated with yttrium-90 via the hepatic artery. The microspheres, which become permanently embedded, are delivered to tumor preferentially to normal liver, as the hepatic circulation is uniquely organized, whereby...
tumors greater than 0.5 cm rely on the hepatic artery for blood supply while normal liver is primarily perfused via the portal vein. Yttrium-90 is a pure beta-emitter with a relatively limited effective range and short half-life that helps focus the radiation and minimize its spread. Radioembolization is generally reserved for patients with adequate functional status (ECOG 0-2), adequate liver function and reserve, Child Pugh score A or B, and liver-dominant metastases. Candidates for radioembolization are initially examined by hepatic angiogram to identify and map the hepatic arterial system, and at that time, a mixture of albumin particles is delivered via the hepatic artery to simulate microspheres. After, single-photon emission CT gamma imaging is used to detect possible shunting of the albumin particles into gastrointestinal or pulmonary vasculature.

**UNRESECTABLE PRIMARY LIVER CANCER [HEPATOCELLULAR CARCINOMA (HCC)]**

The majority of patients with HCC present with unresectable disease and treatment options are limited secondary to the chemoresistance of HCC and the intolerance of normal liver parenchyma to tumoricidal radiation doses.

**Other Treatment Options**

- **Radioembolization.** In general, radioembolization is used for unresectable HCC that is greater than 3 cm.
- **Transarterial chemoembolization (TACE) therapy.** Results of two randomized controlled trials have shown a survival benefit using TACE versus supportive care in patients with unresectable HCC.[2,3]
- **Transarterial embolization (TAE).** In one study, patients were randomly assigned to TACE, TAE, or supportive care. One-year survival rates for TACE, TAE, and supportive care were 82%, 75%, and 63%, respectively and 2-year survival rates were 63%, 50%, and 27%, respectively.
- **Targeted therapies.** A 2007 multicenter, randomized, double-blind placebo controlled Phase III trial that enrolled 602 patients with advanced HCC randomly assigned patients to receive sorafenib versus placebo.[4] Overall survival (OS) was significantly longer in the sorafenib group compared with placebo (10.7 versus 7.9 months, respectively; hazard ratio for sorafenib: 0.69; p<0.001).

**UNRESECTABLE METASTATIC COLORECTAL CARCINOMA**

The role of local (liver-directed) therapy (including radioembolization, chemoembolization, and conformal radiation therapy) for complete tumor removal or destruction is widely accepted in clinical practice. Incomplete “debulking” of unresectable metastatic disease in the liver remains controversial.[5]

Fifty to sixty percent of patients with colorectal cancer develop metastases, either synchronously or metachronously. Emphasis on treating patients with potentially curable disease is on complete destruction or removal of all tumor tissue. The majority of patients diagnosed with metastatic colorectal disease are initially classified as having unresectable disease.

**Other Treatment Options**

- In patients with metastatic disease limited to the liver, preoperative chemotherapy is sometimes used in an attempt to downsize the metastases in order to convert the
metastatic lesions to a resectable status (conversion chemotherapy).

- In patients with unresectable disease that cannot be converted to resectable disease, the primary treatment goal is palliative, with survival benefit shown with both second and third-line systemic chemotherapy.[6]
- Advances in chemotherapy have doubled the median survival in this population from less than 1 year to more than 2 years.
- Palliative chemotherapy by combined systemic and hepatic artery infusion therapy (HAI) may increase disease-free intervals for patients with unresectable hepatic metastases from colorectal cancer.
- Ablation techniques (see Cross References)
- Radiation therapy (see Cross References).

UNRESECTABLE METASTATIC NEUROENDOCRINE TUMORS

Neuroendocrine tumors are an uncommon, heterogeneous group of mostly slow-growing, hormone-secreting malignancies, with an average patient age of 60 years. Primary neuroendocrine tumors vary in location, but most are either carcinoids (which most commonly arise in the midgut) or pancreatic islet cells. Carcinoid tumors, particularly if they metastasize to the liver, can result in excessive vasoactive amine secretion including serotonin and are commonly associated with the carcinoid syndrome (diarrhea, flush, bronchoconstriction, and right valvular heart failure).

Although they are considered to be indolent tumors, at the time of diagnosis, up to 75% of patients have liver metastases. The 5-year survival rates with metastases to the liver are less than 20%. Less than 10% of patients are eligible for resection as most patients have diffuse, multiple lesions.

Conventional therapy is largely considered to be palliative supportive care to control, eradicate, or debulk hepatic metastases, often to palliate carcinoid syndrome or local pain from liver capsular stretching.

Other Treatment Options

- Medical treatment includes somatostatin analogs, like octreotide or lanreotide, or systemic chemotherapy. Although patients often achieve symptom relief with octreotide, the disease eventually becomes refractory, with a median duration of symptom relief of approximately 13 months, with no known effect on survival. Systemic chemotherapy for these tumors has shown modest response rates of limited duration, is better for pancreatic neuroendocrine tumors compared to carcinoids, and is frequently associated with significant toxicity.[7]
- Radiofrequency or cryosurgical tumor ablation (see Cross References)
- Transarterial chemoembolization (TACE) therapy. Chemoembolization has shown response rates of nearly 80%, but the effect is of short duration and a survival benefit has not been demonstrated.[7]
- TAE
- Radiation therapy (see Cross References).

UNRESECTABLE INTRAHEPATIC CHOLANGIOCARCINOMA

Cholangiocarcinomas are tumors that arise from the epithelium of the bile duct and are separated into intrahepatic and extrahepatic types. Intrahepatic cholangiocarcinomas appear in the hepatic parenchyma and are also known as peripheral cholangiocarcinomas.[8]
Resection is the only treatment with the potential for cure and 5-year survival rates have been in the range of 20% to 43%.

**Other Treatment Options**

Patients with unresectable disease may select among fluoropyrimidine-based or gemcitabine-based chemotherapy, fluoropyrimidine chemoradiation, or best supportive care.

**MISCELLANEOUS METASTATIC TUMORS**

Small case reports have been published on the use of radioembolization in many other types of cancer with metastases, including breast, head, and neck (including parotid gland), pancreaticobiliary, anal, thymic, thyroid, endometrial, lung, kidney, gastric, small bowel, esophageal, ovarian, cervical, prostatic, bladder, and for melanoma, sarcoma and lymphoma.[9]

**REGULATORY STATUS**

Currently, two commercial forms of yttrium-90 microspheres are available: a glass sphere, TheraSphere® (MDS Nordion, Inc. used under license by BTG International) and a resin sphere, SIR-Spheres® (Sirtex Medical Limited). Noncommercial forms are mostly used outside the U.S. While the commercial products use the same radioisotope (yttrium-90) and have the same target dose (100 Gy), they differ in microsphere size profile, base material (i.e., resin vs. glass), and size of commercially available doses. These physical characteristics of the active and inactive ingredients affect the flow of microspheres during injection, their retention at the tumor site, spread outside the therapeutic target region, and dosimetry calculations.

Note also that the U.S. Food and Drug Administration (FDA) granted premarket approval of SIR-Spheres® for use in combination with 5-flouxuridine (5-FUDR) chemotherapy by HAI to treat unresectable hepatic metastases from colorectal cancer. In contrast, TheraSphere® was approved by humanitarian device exemption (HDE) for use as monotherapy to treat unresectable hepatocellular carcinoma (HCC). In January 2007, this HDE was expanded to include patients with hepatocellular carcinoma who have partial or branch portal vein thrombosis. For these reasons, results obtained with one product do not necessarily apply to other commercial (or noncommercial) products.

**EVIDENCE SUMMARY**

The principal health outcomes associated with treatment of malignancies are typically measured in units of survival past treatment: disease-free survival (DFS), a period of time following treatment where the disease is undetectable; progression-free survival (PFS), the duration of time after treatment before the advancement or progression of disease; and overall survival (OS), the period of time the patient remains alive following treatment.

In order to understand the impact of radioembolization (RE) on these outcomes, well-designed randomized controlled trials (RCTs) are needed that compare this therapy with standard medical and/or surgical treatment of tumors in the liver.

**UNRESECTABLE PRIMARY LIVER CANCER [HEPATOCELLULAR CARCINOMA (HCC)]**

The following literature review on RE for unresectable HCC focused on systematic literature reviews and comparative studies (randomized and nonrandomized).
Systematic Reviews

Tao (2017) reported on a network meta-analysis comparing nine minimally invasive surgeries for treatment of unresectable hepatocellular carcinoma (HCC).[10] The interventions included were transarterial chemoembolization (TACE), TACE plus sorafenib, sorafenib, TACE plus high-intensity focused ultrasound, TACE plus percutaneous ethanol injection, drug-eluting bead (DEB) plus TACE (DEB-TACE), yttrium-90 radioembolization (90Y RE), TACE plus external-beam radiation therapy (EBRT), and ethanol ablation. The network included 17 studies with 2669 patients and 4 studies with 230 patients including 90Y RE. In a pairwise meta-analysis, patients treated with 90Y RE were more likely to achieve complete remission than those who received TACE (odds ratio [OR], 4.5; 95% confidence interval [CI], 1.3 to 15.1). However, in the network meta-analysis, there was no significant difference between the corresponding 8 treatments and TACE with respect to complete remission, partial response, stable disease, and objective response rate. The treatments were ranked for several outcomes using surface under the cumulative ranking curves (SUCRA). TACE plus EBRT had the highest SUCRA ranking in complete remission (77%), partial response (89%), progressive disease (95%), and objective response rate (81%).

Ludwig (2017) conducted a meta-analysis of studies that indirectly compared DEB-TACE with 90Y RE for HCC.[11] Fourteen studies (total N=2065 patients) comparing DEB-TACE or 90Y RE with conventional TACE for primary HCC treatment were included. The pooled estimate of median survival was 23 months for DEB-TACE and 15 months for RE. The estimated 1-year survival was significantly higher for DEB-TACE (79%) than for RE (55%; OR=0.57; 95% CI, 0.36 to 0.92; p=0.02). Survival did not differ statistically significantly at 2 or 3 years but did favor DEB-TACE. At 2 years, survival was 61% for DEB-TACE and 34% or RE (OR=0.65; 95% CI, 0.29 to 1.44; p=0.29) and at 3 years survival was 56% and 21% (OR=0.71; 95% CI, 0.21 to 2.55; p=0.62), respectively.

Two systematic reviews published in 2016 compared RE with TACE for the treatment of unresectable HCC. Lobo (2016) selected 5 retrospective observational studies (total N=533 patients).[12] Survival at 1 year did not differ statistically between RE (42%) and TACE (46%; relative risk [RR], 0.93; 95% CI, 0.81 to 1.08; p=0.33). At 2 years, the survival rate was higher for RE (27% vs 18%; RR=1.36; 95% CI, 1.05 to 1.76; p=0.02), but there was no statistically significant difference in survival rates at 3, 4, or 5 years. Postprocedural complications were also similar in the 2 groups. Facciorusso (2016) included 10 studies (total N=1557 patients), two of which were randomized controlled trials (RCTs).[13] The OR for survival was not statistically significant at 1 year (OR=1.0; 95% CI, 0.8 to 1.3; p=0.93) but favored RE in years 2 (OR=1.4; 95% CI, 1.1 to 1.90; p=0.01) and 3 (OR=1.5; 1.0 to 2.1; p=0.04).

Vente (2009) conducted a meta-analysis evaluating tumor response and survival in patients who received yttrium-90 glass or resin microsphere radioembolization for the treatment HCC or metastases from colorectal cancer (CRC).[14] (See below under unresectable metastatic CRC section for the data from the meta-analysis as pertains to that disease.) Included studies were from 1986 onward and presented tumor response measured by CT scans and data on median survival times. To allow comparability of results with regard to tumor response, the category of “any response” was introduced, and included complete response, partial response, and stable disease. Overall tumor response could only be assessed as any response because response categories were not uniformly defined in the analyzed studies.
In 14 articles, clinical data were presented on tumor response and survival for 425 patients with HCC who had received yttrium-90 radioembolization. Treatment with resin microspheres was associated with a significantly higher proportion of any response than glass microsphere treatment (0.89 vs. 0.78, respectively; p=0.02). Median survival was reported in 7 studies in which survival time was defined as survival from microsphere treatment or from diagnosis or recurrence of HCC. Median survival from microsphere treatment varied between 7.1 and 21.0 months, and median survival from diagnosis or recurrence was 9.4–24.0 months.

The authors of the meta-analysis concluded that yttrium-90 radioembolization is associated with high response rates, both in salvage and first-line settings, but that the true impact on survival will only become known after publication of several ongoing and/or to-be-initiated Phase III studies, as well as the results of trials in which yttrium-90 radioembolization and modern chemotherapy agents are combined with novel biologic agents.

In May 2013 a comparative effectiveness review of local therapies (i.e., ablation, embolization, and radiotherapy) for patients with unresectable HCC was conducted by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center for the Agency for Healthcare Research and Quality (AHRQ).[15] The review sought to report on overall survival and quality of life outcomes and adverse events. Transplant candidates were excluded from this review. Three prospective case series and one retrospective case series with a total of 187 participants met inclusion criteria for review. There were no randomized controlled trials and no comparative trials that met inclusion criteria. Therefore, the strength of evidence was rated as insufficient to evaluate the outcomes of interest. One study reported a 1-year survival rate of 75%; three studies reported a median survival range of 11 to 15 months. Quality of life, local recurrence, and disease progression were not reported in any of the included studies. Adverse events were rare and no liver failure or hepatic abscess was reported. The authors recommended studies that compare various embolization techniques including radioembolization.

Randomized Controlled Trials

In 2014, Kolligs reported results of a small pilot randomized controlled trial (RCT) comparing RE with TACE for the treatment of unresectable HCC, the SIR-TACE study.[16] The study included 28 subjects with unresectable HCC, preserved liver function, and an Eastern Cooperative Oncology Group [ECOG] Performance Status of 2 or less, with no vascular invasion or extrahepatic spread, who had 5 or fewer liver lesions or a single lesion of 10 cm or less. Patients were randomized to RE (n=13) or TACE (n=15). Over posttreatment follow up, PR rates were 13.3% for TACE and 30.8% for RE, with rates of disease control (CR, SD, PR) of 73.3% for TACE and 76.9% for RE. Median progression-free survival (PFS) was 3.6 months for TACE and 3.7 months for RE.

In 2014, Pitton reported results from a small RCT comparing RE with TACE with drug eluting beads TACE (DEB-TACE) for the treatment of unresectable HCC.[17] The study included 24 patients, 12 randomized to each group. No deaths occurred within 30 days of the procedure for either group. There were no statistically significant differences between the groups in terms of in PFS (180 days for RE vs 216 for TACE; p=0.619) and overall survival (OS; 592 days for RE vs 788 for TACE; p=0.927).

Nonrandomized Comparison Studies
Padia (2017) reported on a single-center, retrospective study (2010-2015) comparing segmental RE with segmental chemoembolization in 101 patients with localized, unresectable HCC not amenable to ablation. Patients receiving chemoembolization had poorer ECOG Performance Status ratings and Child-Pugh class while those receiving RE had larger and more infiltrative tumors. Overall complete remission was 84\% with RE and 58\% with chemoembolization (p=0.001). Median PFS was 564 days and 271 days (p=0.002) and median OS was 1198 days and 1043 days (p=0.35), respectively, for the RE group and the chemotherapy group.

In 2016, Soydal reported a retrospective study comparing outcomes of patients receiving RE and TACE for HCC. Each group included 40 patients. RE patients had a mean survival of 39 months versus 31 months for TACE (p=0.014). There was no significant difference in chronic complications and recurrence of disease.

In 2016, Oladeru reported a retrospective study based on SEER registry data comparing survival outcomes of patients receiving RE and external beam radiation of HCC. A total of 189 patients with unresectable HCC (77 receiving RE, 112 external beam radiotherapy) receiving treatment between 2004 and 2011 were evaluated. Median OS for RE was 12 months versus 14 months for external beam radiotherapy. Median disease-specific survival was identical for both groups at 14 months. After adjustment for differences between patients, multivariable survival analysis showed no association of treatment and OS or disease-specific survival.

In 2015, El Fouly reported results of a nonrandomized study comparing yttrium-90 RE with TACE among 86 patients with intermediate stage, nonresectable HCC. Sixty-three patients at one institution were treated with TACE, while 53 patients at a second institution were treated with RE. Median OS in for TACE and RE was not significantly different between groups (18 months for TACE vs 16.4 months for RE); similarly median time to progression (TTP) was not significantly different between groups (6.8 months for TACE vs 13.3 months for RE). TACE patients had higher numbers of treatment sessions, hospital times, and rates of adverse events. Also in 2015, Gramenzi conducted a retrospective cohort study to compare RE with yttrium-90 with sorafenib for intermediate- or advanced-stage HCC. 15 Patients with HCC refractory to other therapies and no metastases or systemic chemotherapy were included, 74 of whom were treated with sorafenib and 63 treated with RE. Median OS between groups was similar (14.4 months for sorafenib-treated patients vs. 13.2 months for RE-treated patients). After propensity-score matching of 32 subjects in each group, there were no significant differences in median OS or 1-, 2-, and 3-year survival rates between groups.

**RE AS A BRIDGE TO LIVER TRANSPLANTATION FOR PRIMARY HCC**

Salem (2016) reported on results of a phase 2 RCT comparing conventional TACE and TheraSphere radioembolization (Y90) for treatment of unresectable, unablatable HCC. Twenty-four patients were assigned to Y90 and 21 patients to conventional TACE; the ultimate goal of treatment for these patients was liver transplantation. The primary outcome was time to progression using intention-to-treat analysis. Median follow-up was 17 months. In the conventional TACE group, there were 7 transplants at a median of 9 months (range, 3-17 months). In the Y90 group, there were 13 transplants at a median of 9 months (range, 4-15 months). Median time to progression exceeded 26 months in the Y90 group and 6.8 months in the conventional TACE group (hazard ratio, 0.12; 95\% CI, 0.03 to 0.56; p=0.007). Median survival was 19 months in Y90 and 18 months in conventional TACE (p=0.99). Adverse events...
were similar between groups, with the exception of more diarrhea (21% vs 0%) and hypoalbuminemia (58% vs 4%) in the conventional TACE group. A limitation of the OS analysis was the censoring of the survival outcome at liver transplantation given that transplantation is related to the treatment effect.

In 2014, Kulik reported results of a pilot RCT of yttrium-90 RE with or without sorafenib for patients with HCC awaiting liver transplantation.[23] The study randomized 23 subjects; after accounting for losses due to self-withdrawal from the study, failure to confirm HCC, and death, the modified intention-to-treat (ITT) population included 10 subjects randomized to RE alone and 10 randomized to RE with sorafenib. Overall, 17 of 20 patients underwent liver transplantation, with no difference in median time-to-transplant between groups. However, the addition of sorafenib was associated with increased peritransplant biliary complications, and acute rejection.

In a 2013 retrospective review, Tohme reported on 20 consecutive HCC patients on liver transplant waiting lists who received radioembolization as bridge therapy.[24] When radioembolization began, Milan criteria (extent of disease) for liver transplantation were met by 14 patients and sustained until transplantation. Of the 6 patients who did not meet Milan criteria initially, radioembolization was able to downstage 2 patients to meet Milan criteria. Complete or partial radiologic response to radioembolization on modified Response Evaluation Criteria In Solid Tumors (RECIST) occurred in 9 patients. Additionally, on pathologic examination, 5 patients who met Milan criteria had complete tumor necrosis with no evidence of viable tumor.

In 2014, Ramanathan reported on multimodality therapy, including radioembolization, for 715 HCC patients of which 231 were intended for transplant.[25] In the intention-to-treat with transplantation arm, 60.2% were able to receive a transplant. Survival rates posttransplant were 97.1% and 72.5% at 1 and 5 years, respectively. Tumor recurrence rates were 2.4%, 6.2%, and 11.6% at 1, 3, and 5 years, respectively. Since this study included multimodality therapy, it is not possible to isolate the effect of radioembolization.

Lewandowski (2009) compared RE with chemoembolization in the efficacy of downstaging 86 patients with HCC from stage T3 to T2 (potentially making patients liver transplant candidates).[26] Patients were treated with either radioembolization using yttrium-90 microspheres (n=43) or TACE (n=43). Median tumor size was similar between the two treatment groups (5.7 and 5.6 cm, for TACE vs. radioembolization, respectively.) Partial response rates were 61% versus 37% for radioembolization vs. TACE, respectively, with downstaging from T3 to T2 in 58% of patients treated with RE versus 31% with TACE (p<0.05).

**UNRESECTABLE METASTATIC COLORECTAL CARCINOMA (CRC)**

**Systematic Reviews**

A 2010 technology assessment,[6] a 2009 Cochrane review,[27] and a 2009 systematic review with meta-analysis[14] all concluded that data from large Phase III trials were needed in order to fully understand the impact of radioembolization on survival in patient with CRC metastases in the liver.

Two additional systematic reviews were published in 2013:
Rosenbaum considered radioembolization, either as monotherapy or concomitant with chemotherapy, to be an emerging treatment for CRC liver metastases, with a limited amount of data from heterogeneic studies.[28] This review evaluated 13 articles on radioembolization as monotherapy and 13 studies on radioembolization combined with chemotherapy for chemoresistant, unresectable CRC liver metastasis. Heterogeneity between studies prohibited pooling of data. This heterogeneity included varying patient inclusion criteria such as the amount of intrahepatic and extrahepatic tumor burden, patient performance status, previous systemic treatments, and protocols for assessing tumor response. CR, PR, and stable disease (SD) rates ranged from 29% to 90% with radioembolization alone and from 59% to 100% for radioembolization with chemotherapy. At 12 months, survival ranged from 37% to 59% with radioembolization alone and from 43% to 74% for radioembolization combined with chemotherapy. As with prior reviews, the authors concluded that additional data is needed from high-quality randomized trials.

In contrast to the prior systematic reviews, Saxena considered the evidence sufficient to recommend increased utilization of radioembolization as salvage treatment for CRC liver metastases.[29] The review evaluated a total of 979 patients in 20 studies including two RCTs[30,31]. The majority of patients had previously undergone at least 3 lines of chemotherapy (range 2-5). After radioembolization, the average reported CRs and PRs from 16 studies was 0% (range, 0%-6%) and 31% (range, 0%-73%), respectively. The median time to intrahepatic progress was 9 months (range 6-16 months) and the median survival time was 12 months (range 8.3-36 months). The mean rate of acute toxicity was 40.5% (range 11% to 100%); most cases were mild and did not require intervention. Despite concluding that radioembolization was safe and effective, the authors noted the need for continued evaluation of clinical outcomes.

Randomized Controlled Trial

A phase 3 RCT by van Hazel of 530 patients compared patients receiving modified FOLFOX chemotherapy and FOLFOX chemotherapy plus SIRT in patients with previously untreated liver-dominant metastatic disease.[32] Bevacizumab was allowed as additional treatment at the discretion of the treating physician. About 40% of patients had extrahepatic metastases at randomization. About 28% of patients had more than 25% liver involvement of metastases. The primary end point was overall (any site) progression-free survival (PFS). Secondary end points included liver-specific outcomes such as PFS in the liver, tumor response rate, and liver resection rate. The primary end point of PFS at any site showed no difference between groups (10.2 months vs 10.6 months control vs RE; hazard ratio, 0.93; p=0.43). Secondary liver-specific end points of median PFS in the liver and objective response rate in the liver were improved in the RE group (liver PFS, 12.6 months vs 20.5 months control versus RE; liver response rate, 68.8% vs 78.7% control vs RE). Overall survival outcomes have not yet been published. The investigators plan to analyze overall survival of this study in combination with 2 other studies of chemotherapy with and without RE that have not yet been completed. This combined preplanned analysis should be able to determine the efficacy of RE (in combination with current chemotherapy regimens) in first-line treatment of unresectable metastatic CRC.

Nonrandomized Studies

Since the systematic reviews were published, a number of additional nonrandomized studies have reported outcomes of RE for patients with CRC liver metastases who failed or were not candidates for chemotherapy.[33-36] The majority of these were noncomparative studies which
precluded conclusions on the survival benefit of RE compared to other treatments. There was a wide range of clinical response to RE; although the rate of complete response was low, partial response averaged 35% and stable disease was reported in 32-71% of patients. The few studies that compared RE to best supportive care reported a statistically significant survival benefit with RE. The rates of Grade 3-4 toxicities ranged from 0% to 39% and included absolute lymphocyte, alkaline phosphatase, bilirubin, and albumin. Factors associated with poorer prognosis included large tumor volume, poor radiological response to treatment, and the number of prior chemotherapy treatments.

**MELANOMA METASTASES IN THE LIVER**

The evidence related to the use of RE for melanoma consists of relatively small observational studies, many of which focus on patients with uveal melanoma in whom the liver is the most common site of metastatic disease.

**Randomized Controlled Trials**

No randomized controlled trials were identified for radioembolization of melanoma metastases in the liver.

**Nonrandomized Comparative Studies**

In 2014, Xing conducted a retrospective observational study to compare outcomes for patients with unresectable melanoma (both uveal and cutaneous) liver metastases refractory to standard chemotherapy treated with either yttrium-90 RE (n=28) or best supportive care (n=30). The groups were similar at baseline in terms of Child-Pugh class, ECOG performance status scores, age, sex, and race. However, patients treated with RE had significantly larger tumor size at baseline than those treated with best supportive care (mean, 7.28 cm vs 4.19 cm; p=0.02). Median OS from diagnosis of melanoma liver metastases was longer in RE-treated subjects (19.9 mo. vs 4.8 mo.; p<0.000), as was the median OS from diagnosis of the primary melanoma (119.9 months vs 26.1 months; p<0.001). Pre- and post-treatment imaging studies were available for 24/28 (85.7%) of those treated with RE. Of those, no patients had a CR; 5 patients (17.9%) had PR, 9 patients (32.1%) had SD, and 10 patients (35.7%) had PD. Two patients receiving RE had major (grade 5) clinical toxicities (ascites and hepatic encephalopathy and eventual mortality). Significant factors for longer OS were ≤10 metastatic liver lesions, absence of extrahepatic metastases, and Child-Pugh class A. Although this study was retrospective and included small sample sizes, it included relatively long-term follow-up and provided comparison between RE and best supportive care.

**Nonrandomized Non-comparative Studies**

In 2014, Eldredge-Hindy retrospectively evaluated outcomes for the use of yttrium-90 RE in 71 patients with biopsy-confirmed uveal melanoma liver metastases. The median time from the diagnosis of liver metastases to RE was 9.8 months (95% CI, 7.4 to 12.2 months), and 82% of patients had received prior liver-directed therapies. Sixty-one patients (86%) had CT or magnetic resonance imaging (MRI) evaluation of treatment response at 3 months post-RE. Of those, 5 patients (8%) had a PR, 32 patients (52%) had SD, and 24 patients (39%) had DP. Median OS RE was 12.3 months (range, 1.9-49.3 months).

Six small studies (n=8–32) reported on use of RE in patients with hepatic metastases from melanoma. Four of the studies included only patients with ocular melanoma, and two included patients with ocular, cutaneous, or other site melanoma. Three studies excluded...
those patients with poor performance status. Median age was in the 50s for four studies and 61 in one study. One article did not describe any previous treatment and one described it incompletely. Four studies reported tumor response data, by RECIST criteria.

- **Treatment response.** Among 32 patients in the study by Gonsalves, one patient had a CR (3%), one had a PR, 18 patients had SD (56%) and 12 patients had PD (38%). In the study of 13 patients published by Klingenstein, none had a CR, 8 had a PR (62%), 2 had SD (15%) and 3 had PD (23%). Nine of 11 patients in the article by Kennedy provided response data: one had CR, 6 had PR, 1 had SD and 1 had PD. Of the 8 patients in the Schelhorn study, four (50%) had SD and 4 (50%) had PD. Memon reported PD and SD in 13 (81%) patients and PD in 3 (19%) patients.

- **Survival.** Median survival in Gonsalves, Klingenstein, Schelhorn, and Kennedy were 10.0 months, 19 months, 20 months, and not yet reached, respectively.

- **Toxicity.** Gonsalves reported 4 patients (12.5%) with grade 3-4 liver toxicity. Klingenstein observed one patient with marked hepatomegaly. Kennedy described one grade 3 gastric ulcer. Memon reported Grade 3 toxicity in two (12%) (absolute lymphocyte toxicity) and 1 (7%) (aspartate aminotransferase toxicity) patients; and grade 4 bilirubin toxicity in 1 patient. One study[^42] (n=12) did not include any toxicity data.

### Unresectable Metastatic Neuroendocrine Tumors

#### Systematic Reviews

A 2012 systematic review evaluated the safety and efficacy of chemoembolization, bland embolization, and radioembolization in patients with unresectable metastatic neuroendocrine tumors (mNET) in the liver.^[45] A total of 37 studies with 1575 total patients were reviewed for response to treatment, survival outcome, and toxicity. The authors reported that each of these therapies were found to be safe and effective, and recommended additional prospective trials to compare relative efficacy and toxicity.

In 2014, a meta-analysis of 12 studies that met inclusion criteria reported complete and partial responses of 50% for radioembolization of metastatic neuroendocrine tumors (mNET) in the liver.^[46] Weighted average disease control was 86%. It was noted that patients with pancreatic mNET was marginally associated with poorer response (p=0.03). The authors concluded that the meta-analysis confirmed the effectiveness of radioembolization of hepatic mNET.

#### Randomized Controlled Trials

No randomized controlled trials were found for radioembolization of metastatic neuroendocrine tumors in the liver.

#### Nonrandomized Comparative Studies

Engelman retrospectively compared locoregional therapies including transarterial, liver-directed therapies including RE, hepatic artery embolization (HAE), and hepatic artery chemoembolization (HACE) in 42 patients treated for metastatic neuroendocrine tumors.^[47] Treatment decisions were at the discretion of the referring physician and interventional radiologist, but the decision to proceed with therapy was typically based on progression of symptoms nonresponsive to octreotide therapy or rapid progression of liver tumor burden on imaging. Seventeen patients had HACE, 13 had HAE, and 12 had RE. Among the 27 patients...
with symptoms from their liver metastases, there were no statistically significant differences in symptom improvement at 3 months after first liver-directed therapy across treatment modalities (6/13 for HACE; 4/8 for HAE; 5/6 for RE; p=0.265). There were no differences between treatment modalities in radiographic response at 6 months postprocedure (p=0.134), TTP (p=0.968), or OS (p=0.30).

**Nonrandomized Non-Comparative Studies**

In 2015 Peker reported on 30 patients with unresectable hepatic mNET who received resin-based RE.[48] Post-treatment response was assessed by imaging using the Response Evaluation Criteria in Solid Tumors (RECIST) guidelines. Mean follow-up was 23 months. Median OS was 39 months (range 12.6-65.4 months) with 1- and 2-year survival rates of 71% and 45%, respectively. PR was 43%, CR 3%, SD 37%, and PD 17%. The following were not significant prognostic factors: extrahepatic disease, radiographic response, age, and primary NET site.

In 2010, Cao reported the outcomes of 58 patients with unresectable neuroendocrine liver metastases from 2 different hospitals treated with yttrium-90 microspheres (SIR-Spheres) from 2003 to 2008. Data were examined retrospectively from a database.[49] Response was assessed with radiographic evidence before and after radioembolization and measured by RECIST guidelines. Patients typically had a CT scan within 3 months of treatment and every 3 to 6 months until disease progression or death. Systemic chemotherapy was routinely given at 1 institution but not the other. Mean patient age at the time of radioembolization was 61 (range: 29-84 years), and 67% of patients were men. Primary tumor site was variable and included small bowel, pancreas, colon, thyroid, lung, and unknown. Thirty-one patients underwent surgical resection of their primary tumor, which was classified as low-grade in 15, intermediate-grade in 7, and high-grade in 7. Forty-three percent of patients had extrahepatic metastatic disease at study entry. Prior therapies before radioembolization included liver resection in 19 patients, TAE or TACE in 6, ablation or percutaneous ethanol injection in 10, previous chemotherapy in 20, concurrent chemotherapy in 34, and post-radioembolization chemotherapy in 5 patients. Median follow-up was 21 months (range 1-61 months). Fifty-one patients were evaluable, and 6 achieved a complete response, 14 a partial response, 14 had stable disease, and 17 had disease progression. Overall survival (OS) rates at 1, 2, and 3 years were 86, 58, and 47%, respectively. Median survival was 36 months (range: 1-61 months). Prognostic factors for survival included extent of tumor involvement of the liver, radiographic response to treatment, presence of extrahepatic disease at the time of radioembolization, histological grade of tumor, and whether patients were responders (versus nonresponders) to radioembolization. Factors that were not significant prognostic features included age, sex, ECOG status, and previous therapy.

King reported outcomes in patients treated in a single-institution prospective study.[7] Thirty-four patients with unresectable neuroendocrine liver metastases were given radioactive microspheres [SIR-Spheres] and concomitant 7-day systemic infusion of 5-FU, between 2003 and 2005. Mean patient age was 61 years (range: 32-79 years), and 65% were men. Mean follow-up was 35.2 +/- 3.2 months. The mean interval from diagnosis of hepatic metastases and treatment with SIR therapy was 36.6 +/- 6.7 months. Primary tumor sites were variable and included bronchus (n=1), thyroid (n=2), gastrointestinal (n=15), pancreas (n=8), and unknown (n=8). Subjective changes from baseline hormone symptoms were reported every 3 months. Twenty-four patients (71%) had, at baseline assessment, symptoms of carcinoid syndrome, including diarrhea, flushing, or rash. At 3 months, 18 of 33 patients (55%) reported...
improvement of symptoms, as did 16 of 32 (50%) at 6 months. Radiologic tumor response was observed in 50% of patients and included 6 CR (18%), and 11 PR (32%). Mean OS was 29.4 +/- 3.4 months.

INTRAHEPATIC CHOLANGIOCARCINOMA

Systematic Reviews

In 2015, Al-Adra reported results from a systematic review of studies reporting outcomes for RE for ICC.[50] The review included 12 publications, 7 of which were published in abstract form only. Of the peer reviewed manuscripts, three were described as prospective cohort studies.[51-53] The overall weighted median survival was 15.5 months (range 7-22.2 months), based on 11 included studies. A weighted mean PR was seen in 28% of patients and stable disease was seen in 54% at 3 months posttreatment.

In 2015, Boehm conducted a meta-analysis to compare hepatic artery-based therapies including hepatic arterial infusion (HAI), TACE, DEB-TACE, and yttrium-90 RE for unresectable ICC.[54] Twenty studies met inclusion criteria, five of which evaluated yttrium-90 RE. Median OS across studies was 22.8 months for HAI, 13.9 months for RE, 12.4 months for TACE, and 12.3 months for DEB-TACE. CR or PR occurred in 56.9% of patients treated with HAI, compared with 27.4% of those treated with RE and 17.3% of those treated with TACE. While HAI showed the highest median OS, it also had the highest rate of grade III and IV toxicity.

Randomized Controlled Trials

No randomized controlled trials were found for radioembolization of ICC.

Nonrandomized Studies

Additional nonrandomized studies published after the above systematic reviews are included here.

Jia (2017) retrospectively reviewed all 24 patients who underwent Y90 RE for unresectable and failed first-line chemotherapy for ICC at a single institution.[55] Mean follow-up was 11 months (range, 3-36 months). Median OS from time of diagnosis was 24 months (range, 18-30 months) and from the RE procedure was 9 months (range, 6-12 months). Survival rates at 6, 12, and 30 months was 70%, 33%, and 20%, respectively.

Mosconi (2016) retrospectively analyzed 23 consecutive patients with unresectable or recurrent ICC at a single institution.[56] Overall median survival was 18 months (95% CI, 14 to 21 months). Survival was significantly longer in treatment-naive patients (52 months) than in those who received other treatments before RE (16 months; p=0.009).

Rayar (2015) reported successful downstaging after RE in eight patients with unresectable ICCs. Initial unresectability was due to involvement of hepatic veins or portal veins of the future liver remnant.[57] After RE there was significant decrease in tumor volume and all patients were subsequently able to undergo successful resection. At median follow-up of 15.6 months (range 4-40.7 months) after medical treatment and 7.2 months (range 0.13-36.4 months) after surgery, five patients were still alive, one of which was alive at 40 months after medical treatment. Two patients had tumor recurrence.
METASTATIC BREAST TUMORS

Systematic Reviews

One systematic review included six studies with a total of 198 patients with breast cancer metastases in the liver.^[58^] Five studies reported tumor response. Overall disease control (complete response, partial response, and stable disease) at 2-4 months post-treatment ranged from 78% to 96%. Median survival was reported in four studies and ranged from 10.8 to 20.9 months. Adverse effects included gastric ulceration in 10 patients (5%) and treatment-related mortality in 3 patients (2%). The authors concluded that these studies showed safety and effectiveness of treatment and strongly encouraged comparative studies, in particular, combining radioembolization with systemic therapy.

Nonrandomized Studies

Table 1. Retrospective Case Series of Radioembolization for Liver Metastases in Breast Cancer

<table>
<thead>
<tr>
<th>Study (Year)</th>
<th>Populations</th>
<th>Outcomes</th>
</tr>
</thead>
</table>
| Pieper et al (2016)^[59^] | 44 women with unresectable liver-dominant breast metastases who had failed 2+ lines of chemotherapy who underwent yttrium-90 RE at a single center from 2006-2015 | ORR: 29%  
Disease control rate: 71%  
Median TTP: 101 d  
Median survival: 184 d  
Grade 2 toxicity: 1 (cholecystitis)  
Grade 3 toxicity: 1 (duodenal ulceration) |
| Gordon et al (2014)^[60^] | 75 women with stable extrahepatic disease who had hepatic tumor progression after systemic chemotherapy treated with yttrium-90 RE at a single center | 30-day mortality: 4%  
Median OS: 6.6 mo (95% CI, 5.0 to 9.2 mo)  
Median hepatic TTP: 3.2 mo (95% CI, 1.2 to 8.5 mo)  
Median distant TTP: 4.1 mo (95% CI, 2.7 to 7.0 mo) |
| Saxena et al (2014)^[61^] | 40 women with unresectable, chemoresistant breast cancer–related liver metastases treated from 2006-2012 at a single institution who had received at least 1 line of systemic chemotherapy | Grade 1 or 2 clinical toxicity: 40%  
Of 38 women with ≥1 mo follow-up:  
CR: 5%  
PR: 26%  
SD: 39%  
PD: 29%  
Median survival: 13.6 mo |
| Cianni et al (2013)^[62^] | 52 women with chemotherapy-refractory breast cancer and inoperable liver metastases; chemotherapy administered previously to all patients, surgery in 17.3%, TACE in 3.8%, and RFA in 3.8% | CR: 0%  
PR: 56%  
SD: 35%  
PD: 10%  
Median OS: 11.5 mo |
| Haug et al (2012)^[52^] | 58 women with chemotherapy-refractory breast cancer and unresectable hepatic metastases | Mean follow-up: 27.5 wk  
CR: 0%  
PR: 25.6%  
SD: 62.8%  
PD: 11.6%  
Median OS: 47 wk |
| Jakobs et al (2008)^[63^] | 30 (29 women, 1 man) patients who underwent RE with resin microspheres in a single-session, whole-liver treatment for | For 23 patients with follow-up data, after median follow-up of 4 mo:  
PR: 61%  
SD: 35% |
breast cancer metastases and had failed prior polychemotherapy regimens

<table>
<thead>
<tr>
<th>Study</th>
<th>Patients</th>
<th>Treatment</th>
<th>Response Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangash et al (2007)(^\text{[64]})</td>
<td>27 women</td>
<td>progressive liver metastases from breast cancer while on polychemotherapy</td>
<td>CR: 39%, PR: 39%, SD: 52%, PD: 9%</td>
</tr>
<tr>
<td>Coldwell et al (2007)(^\text{[65]})</td>
<td>44 patients</td>
<td>hepatic metastases at 3 hospitals who failed 1st-, 2nd-, or 3rd-line treatment for primary breast tumor and were not candidates for RFA, TACE, resection, IMRT, or SRT</td>
<td>PR: 47%</td>
</tr>
</tbody>
</table>

CI: confidence interval; CR: complete response; ECOG: Eastern Cooperative Oncology Group; IMRT: intensity-modulated radiotherapy; ORR: response rate; OS: overall survival; PD: progressive disease; PR: partial response; RE: radioembolization; RFA: radiofrequency ablation; SD: stable disease; SRT: stereotactic radiotherapy; TACE: transarterial chemoembolization; TTP: time to progression.

OTHER METASTATIC TUMORS IN THE LIVER

Data on the use of radioembolization in other tumors metastatic to the liver are limited and included numerous methodologic limitations such as patient heterogeneity, lack of a control group, and patient numbers too small to draw meaningful conclusions. For example, a retrospective data analysis was reported in 2014 by Michl on RE for liver metastases from pancreatic cancer. Nineteen patients were included, 16 of whom had received previous palliative chemotherapy.\(^\text{[66]}\) Median local PFS in the liver was 3.4 months (range 0.9-45.0). Median OS was 9 months (range 0.9-53.0) and 1-year survival was 24%. Adverse effects were grade <3 (e.g., nausea, vomiting, fatigue, fever, abdominal pain) in the short term and long-term effects included liver abscess, gastroduodenal ulceration, cholestasis and cholangitis, ascites, and spleen infarction. The lack of a control group precludes conclusions about any survival benefits and complication rates of RE.

RADIOEMBOLIZATION AS A BRIDGE TO HEPATIC RESECTION

In 2013, Vouche reported on 83 patients treated with radioembolization as a technique to control or limit tumor progression in unresectable, unilobar hepatic disease and to hypotrophy a small future liver remnant.\(^\text{[67]}\) Patients included in the study had right unilobar disease with HCC (n=67), cholangiocarcinoma (n=8), or metastatic CRC (n=8). One month after radioembolization, significant right lobe atrophy (p=0.003), left lobe hypertrophy (p<0.001), and future liver remnant hypertrophy (p<0.001) were observed and remained during follow-up. Successful right lobectomy was later performed in 5 patients, and 6 patients received liver transplants. However, further studies are needed to assess radioembolization as a bridge to hepatic resection.

PRACTICE GUIDELINE SUMMARY

NATIONAL COMPREHENSIVE CANCER NETWORK GUIDELINES

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.

Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Primary Hepatocellular Carcinoma

National Comprehensive Cancer Network (NCCN) guidelines recommend that patients with unresectable/inoperable disease who are eligible to undergo embolization therapy and have tumor lesions larger than 5 cm should be treated using arterial embolic approaches (chemoembolization, bland embolization, or radioembolization) or systemic therapy.[8] Patients with lesions 3-5 cm can be considered for combination therapy with ablation and arterial embolization, and tumors 3 cm or smaller should be treated with ablation (all category 2A recommendations). The guidelines note that randomized, controlled studies on the use of radioembolization therapy in the treatment of patients with HCC are needed and participation in prospective clinical trials is preferred for all stages of disease.

NCCN indicated that there is limited evidence available on the utility of radioembolization as a bridge to liver transplant for patients on a liver transplant waiting list. However, some NCCN centers use radioembolization as a bridge to transplant.

Primary Intrahepatic Cholangiocarcinoma

Recommendations for unresectable intrahepatic cholangiocarcinoma (ICC) include chemotherapy, clinical trial, and supportive care.[8] The guidelines note that, due to the rarity of this disease, there have been no RCTs on locoregional therapies such as radioembolization for cholangiocarcinoma. However, retrospective series have reported RE to be safe and effective for unresectable ICC. Based on the current evidence, a 2B recommendation was made in favor or RE and other locoregional therapies for unresectable ICC. A 2B recommendation is defined as based upon lower-level evidence and NCCN consensus that the intervention is appropriate. This is a lower level recommendation than the standard 2A NCCN recommendations which are also based upon lower-level evidence but with uniform consensus.

Metastatic Colorectal Cancer

Use of intra-arterial embolization including RE is a category 3 recommendation for highly selected patients with chemotherapy-resistant/-refractory disease without obvious systemic disease, with predominant hepatic metastases.[5,68] Category 3 is the lowest level recommendation, defined by NCCN as a recommendation based on any level of evidence but reflects major disagreement.

Metastatic Neuroendocrine Tumors

For unresectable liver metastases (carcinoid or neuroendocrine tumors of the pancreas, e.g., islet cell), recommendations include hepatic regional therapy which includes radioembolization (category 2B lower-level evidence with NCCN consensus).[1]

Metastatic Breast Cancer

Current recommendations do not address the use of radioembolization in the treatment of metastatic breast cancer.[69]

Metastatic Melanoma

Current recommendations do not address the use of radioembolization in the treatment of metastatic melanoma.[70]
Primary Hepatocellular Carcinoma

ACR Appropriateness Criteria consider radioembolization with beta-emitting Y90 beads to be an emerging treatment option for HCC, with outcomes similar to those with transarterial chemoembolization (TACE) and transarterial embolization (TAE), but with the possibility of less patient discomfort and toxicity. The guideline also reports that radioembolization has “shown the ability to effectively downstage patients for potential transplant or resection. Therefore, ACR recommendations are that radioembolization may be appropriate for solitary HCC tumor <3cm, and usually appropriate, particularly in the presence of portal vein thrombosis or extensive bilobar disease, for solitary HCC tumor of 5 cm and for multiple tumors, at least one of which is >5cm.

Metastatic Colorectal Cancer

The ACR reports that published evidence suggests that TACE and radioembolization provide similar survival benefit and may be appropriate for patients with metastatic liver-dominant colorectal tumors ≥5cm, or for solitary colorectal liver metastasis.

Metastatic Neuroendocrine Tumors

The ACR reports increasing research into the use of radioembolization in this patient population, with early small studies suggesting therapeutic equivalency with more traditional arterial embolization techniques. Radioembolization is recommended as usually appropriate for symptomatic neuroendocrine metastases in the liver when medication fails to control symptoms.

AMERICAN COLLEGE OF RADIOLOGY (ACR)/AMERICAN SOCIETY FOR RADIATION ONCOLOGY (ASTRO)/SOCIETY OF INTERVENTIONAL RADIOLOGY (SIR)

The list of indications in the ACR/ASTRO/SIR guidelines “include, but are not limited to:”

- Unresectable and/or inoperable primary or secondary liver malignancies that are liver dominant but not necessarily exclusive to the liver; and
- Performance status that will allow them to benefit from the therapy (e.g., ECOG performance status of 0 or 1 or KPS of 70 or more); and
- Life expectancy of at least 3 months

RADIOEMBOLIZATION BRACHYTHERAPY ONCOLOGY CONSORTIUM

Members met as an independent group of experts in interventional radiology, radiation oncology, nuclear medicine, medical oncology, and surgical oncology. Using level 2A evidence (panel consensus with low-level evidence), 14 recommendations were made. They concluded that there was sufficient evidence to support the safety and efficacy of yttrium-90 microsphere therapy and that its use requires multidisciplinary management, adequate patient selection, and meticulous angiographic technique. They also stated that the initiation of clinical trials was necessary to further define the role of yttrium-90 microsphere therapy in relation to other currently available therapies.
SUMMARY

PRIMARY HEPATOCELLULAR CARCINOMA (HCC)

Studies have demonstrated that radioembolization is comparable to transarterial chemoembolization (TACE), which is considered to be the therapy of choice for patients with unresectable primary hepatocellular carcinoma (HCC) in terms of tumor response and overall survival. However, disadvantages of TACE include the necessity of multiple treatment sessions and hospitalization, its contraindication in patients with portal vein thrombosis, and its poorer tolerance by patients. Therefore, radioembolization may be considered medically necessary for the treatment of unresectable primary HCC or as a bridge to transplantation in primary HCC.

METASTATIC COLORECTAL CANCER IN THE LIVER

A major cause of morbidity and mortality in patients with colorectal disease metastatic to the liver is liver failure, as this disease tends to progress to diffuse, liver-dominant involvement. Therefore, the use of radioembolization to decrease tumor bulk and/or halt the time to tumor progression and liver failure may lead to prolonged progression free and overall survival in patients with no other treatment options (i.e., those with chemotherapy refractory liver-dominant disease). Other uses include palliation of symptoms from tumor bulk. Radioembolization for the treatment of unresectable hepatic metastases from colorectal cancer may be considered medically necessary in carefully selected patients, when criteria are met.

There is insufficient evidence on the use of radioembolization for the treatment of unresectable hepatic metastases from colorectal cancer when the patient does not meet criteria. Therefore, radioembolization for the treatment of unresectable hepatic metastases from colorectal cancer is considered investigational when criteria are not met.

METASTATIC NEUROENDOCRINE TUMORS IN THE LIVER

Studies of radioembolization for treatment of metastatic neuroendocrine tumors in the liver have included heterogeneous patient populations, making interpretation of survival data difficult. However, relief of symptoms from carcinoid syndrome has been reported in a proportion of patients. Surgical debulking of liver metastases has shown palliation of hormonal symptoms; similarly, debulking by radioembolization may lead to symptom relief in some patients. Therefore, radioembolization for the treatment of unresectable hepatic metastases from neuroendocrine tumors may be medically necessary in carefully selected patients when criteria are met.

There is insufficient evidence on the use of radioembolization for the treatment of hepatic metastases from neuroendocrine tumors when the patient does not meet criteria. Therefore, radioembolization for the treatment of hepatic metastases from neuroendocrine tumors is considered investigational when criteria are not met.

METASTATIC MELANOMA IN THE LIVER

In patients with uveal melanoma, the liver is the most common site of metastatic disease. Studies of radioembolization for treatment of metastatic melanoma (uveal or cutaneous) in...
the liver consists of one comparative study and several relatively small observational studies. In general, these studies predict good tumor response to radioembolization and report significant increases in overall survival compared to those treated with best supportive care. Therefore, radioembolization may be considered medically necessary for the treatment of diffuse, symptomatic hepatic metastases from melanoma when criteria are met.

There is insufficient evidence on the use of radioembolization for the treatment of hepatic metastases from melanoma when the patient does not meet criteria. Therefore, radioembolization for the treatment of hepatic metastases from melanoma is considered investigational when criteria are not met.

PRIMARY INTRAHEPATIC CHOLANGIOCARCINOMA (ICC)

The current evidence on the use of radioembolization (RE) in patients with primary intrahepatic cholangiocarcinoma (ICC) is limited to data from small studies that do not compare the health outcomes of RE with other treatments. These study designs make interpretation of the data on tumor response and survival difficult to interpret. However, ICC is a rare tumor, so large comparative studies may never become available. The available studies have consistently reported beneficial effects in patients who are not candidates for surgical tumor resection. Because there are currently limited treatment options for these patients, radioembolization may be medically necessary for the treatment of unresectable primary ICC. Since surgical resection is currently the preferred treatment for these tumors, radioembolization is considered investigational for resectable primary ICC.

MISCELLANEOUS METASTATIC TUMORS IN THE LIVER

The current evidence on the use of radioembolization in intrahepatic cholangiocarcinoma and metastatic tumors in the liver other than those from colorectal carcinoma, melanoma or neuroendocrine tumors is too limited to draw meaningful conclusions due to methodologic limitations such as small numbers of heterogeneous patients. Therefore, radioembolization for these other tumors, including metastatic tumors from breast and pancreatic cancer, is considered investigational.

REFERENCES


15. Belinson, S, Yang, Y, Chopra, R, Shankaran, V, Samson, D, Aronson, N. Local Therapies for Unresectable Primary Hepatocellular Carcinoma [Internet]. AHRQ Comparative Effectiveness Reviews. 2013 May. PMID: 23844445


19. Soydal, C, Arslan, MF, Kucuk, ON, Idilman, R, Bilgic, S. Comparison of survival, safety, and efficacy after transarterial chemoembolization and radioembolization of Barcelona

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


45. Yang, TX, Chua, TC, Morris, DL. Radioembolization and chemoembolization for unresectable neuroendocrine liver metastases - a systematic review. *Surgical oncology*. 2012 Dec;21(4):299-308. PMID: 22846894

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.


72. ACR–SIR Practice parameter for radioembolization with microsphere brachytherapy device (RMBD) for treatment of liver malignancies. Revised 2014. [cited 07/24/2018]; Available from: [https://www.acr.org/-/media/ACR/Files/Practice-Parameters/RMBD.pdf](https://www.acr.org/-/media/ACR/Files/Practice-Parameters/RMBD.pdf)


74. BlueCross BlueShield Association Medical Policy Reference Manual "Radioembolization for Primary and Metastatic Tumors of the Liver." Policy No. 8.01.43
NOTE: CPT code 37243 can be used for both radioactive and non-radioactive embolization procedures performed for numerous conditions/locations. Only radioactive embolization for the liver is addressed in this policy.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>37243</td>
<td>Vascular embolization or occlusion, inclusive of all radiological supervision and interpretation, intraprocedural roadmapping, and imaging guidance necessary to complete the intervention; for tumors, organ ischemia, or infarction</td>
</tr>
<tr>
<td></td>
<td>75894</td>
<td>Transcatheter therapy, embolization, any method, radiological supervision and interpretation</td>
</tr>
<tr>
<td></td>
<td>77399</td>
<td>Unlisted procedure, medical radiation physics, dosimetry and treatment devices, and special services</td>
</tr>
<tr>
<td></td>
<td>77778</td>
<td>Interstitial radiation source application; complex</td>
</tr>
<tr>
<td>HCPCS</td>
<td>C2616</td>
<td>Radiopharmaceutical therapy, by intra-arterial particulate administration</td>
</tr>
<tr>
<td></td>
<td>S2095</td>
<td>Brachytherapy source, non-stranded, yttrium-90, per source</td>
</tr>
</tbody>
</table>

**Date of Origin:** December 2010
**Ovarian, Internal Iliac, and Gonadal Vein Embolization, Ablation, and Sclerotherapy**

**Effective:** June 1, 2018

**Next Review:** April 2019

**Last Review:** April 2018

**IMPORTANT REMINDER**

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

**DESCRIPTION**

Embolization involves occlusion of blood flow through the ovarian, internal iliac, and gonadal veins with coils, foam, or a chemical sclerosant as a treatment of pelvic congestion syndrome or varicoceles.

**MEDICAL POLICY CRITERIA**

**Note:** This policy does not address surgical ligation of the spermatic vein(s) or uterine artery embolization.

I. Embolization, ablation, and sclerotherapy of ovarian veins, internal iliac veins, or gonadal veins is considered **investigational** for the treatment of the following conditions:
   
   A. Pelvic congestion syndrome
   B. Varicoceles.

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.
BACKGROUND

Enlarged ovarian and internal iliac veins can lead to pelvic congestion syndrome in women, and enlarged gonadal and internal iliac veins can lead to a varicoceles in men. Each are discussed separately below.

PELVIC CONGESTION SYNDROME

Pelvic congestion syndrome (PCS), also called pelvic venous incompetence, is a rare condition characterized by chronic pelvic pain. Although this condition is primarily found in women it can also be found in men. PCS is often aggravated by standing for long periods of time, and often manifests during or after pregnancy. The syndrome is thought to be associated with dilated and refluxing incompetent pelvic veins, similar to what happens in varicose veins of the legs. However, the cause of PCS is unclear. Furthermore, there are no definitive diagnostic criteria for PCS. Instead the diagnosis is generally based on a combination of symptoms, tenderness on physical exam, and documentation of pelvic vein dilation or incompetence after excluding all other causes for the nonspecific findings. Although imaging may show vein dilation or incompetence, these findings are common nonspecific findings and therefore no diagnostic.

There is no standard treatment approach for PCS, and the optimum treatment is unknown. Instead, therapy is individualized and based on symptoms. Medical therapy is generally the first line of treatment, as it is low risk and non-invasive. Other methods, such as embolization has been proposed as an alternative to surgical treatment for patients who fail medical therapy with analgesics. Embolization therapy involves the occlusion of blood flow through the ovarian and internal iliac veins with coils, glue, or chemical sclerosants. The internal iliac veins may be treated at the same time or a later date to prevent recurrence.

VARICOCELES

A varicocele is the dilation of the pampiniform plexus of the gonadal veins. Varicocele’s are present in 15 to 20% of post-pubertal males, and generally get larger over time. Most varicoceles occur in the left hemiscrotum because the left gonadal vein is one of the longest veins in the body and it enters the left renal vein at a perpendicular angle increasing pressure which can dilate the veins and cause incompetence of the valves, similar to what happens in varicose veins of the legs. Although varicoceles on the left are more common, bilateral varicoceles can occur; however, this could be caused by a possible underlying pathology warranting more investigation. Symptoms of a varicocele include dull, aching, left scrotal pain, which is often aggravated by standing for long periods of time, testicular atrophy, and decreased fertility. Although there are no clear guidelines regarding the established treatment for varicoceles, surgical ligation is the preferred first-line treatment.

EVIDENCE SUMMARY

The primary beneficial outcomes of interest for treatments of pelvic pain in both men and woman are symptom reduction and improvement in the ability to function. These are subjective outcomes that are typically associated with a placebo effect. Therefore, data from adequately...
powered, randomized controlled trials (RCTs) with sufficient long-term follow-up are required to control for the placebo effect, determine its magnitude, and to determine whether any treatment effect from provides a significant advantage over placebo or other treatment options.

**TREATMENT FOR PELVIC CONGESTION SYNDROME**

**Health Technology Assessments**

In 2016, Champaneria published a health technology assessment from the National Institute for Health Research that examined the diagnosis and treatment of pelvic vein incompetence and chronic pelvic pain in women.[1] Forty studies were included in the review; six association studies, ten studies involving ultrasound, two studies involving magnetic resonance venography, 21 case series, and one poor-quality randomized trial of embolization. The authors found that there were no consistent diagnostic criteria for pelvic congestion syndrome (PCS). Although the studies have showed associations between chronic pelvic pain (CPP) and pelvic vein incompetence (PVI), the prevalence of PVI ranged widely. The authors identified that transvaginal ultrasound with doppler and magnetic resonance venography are both useful screening methods; however, there is limited data on the accuracy of these methods for PCS. Finally, although the research showed embolization provides symptomatic relief in the majority of women, these studies were small case series. The authors concluded that more research is needed to determine what the diagnostic criteria for PCS are, and the efficacy of embolization as a treatment for PCS.

**Systematic Reviews**

A 2016 systematic review by Mahmoud identified 20 case series (total N=1081 patients) who underwent vein embolization for pelvic congestion syndrome.[2] The authors did not require any particular diagnostic criteria for pelvic congestion syndrome. The length of follow-up in the studies ranged from one month to six years. Seventeen studies (n=648 patients) reported the proportion of patients who reported symptom relief. Overall, 571 (88.1%) patients reported short-term symptom relief and 77 (11.9%) reported little or no relief. Seventeen studies (n=721 patients) reported symptom relief at 12 months. A total of 88.6% had symptom improvement and 13.4% reported little or no relief. Only one study used a comparison group, but patients in it received conservative treatment because they were ineligible for vein embolization therapy, so outcomes after the two interventions cannot be compared.

A systematic review by Daniels (2016) assessed the effectiveness of sclerotherapy or embolization for the treatment of chronic pelvic pain.[3] The review included 21 case series and one poor-quality randomized trial. Due to the overall low quality and heterogeneity of the studies, a meta-analysis was not performed. However, the authors reported that approximately 75% of women who underwent embolization experienced early pain relief. Adverse events noted included, transient pain following foam embolization and a small (<2%) risk of coil migration.

In 2015 Hansrani published a systematic review that evaluated the effectiveness of transvenous occlusion as a treatment of chronic pelvic pain.[4] Thirteen studies were included comprising 866 women. The authors noted that all 13 studies were of poor methodological quality, and most studies did not use objective outcome measures or have consistent follow-up of outcomes. Studies on embolization for treatment of PCS were rated as poor due to lack of randomization and control groups, unclear patient selection criteria, and heterogeneous outcome measures that did not permit between-study comparison or estimates of overall...
treatment effects. There was one RCT included in the review, in which embolization resulted in significantly better pain reduction than hysterectomy, but the study also had significant limitations, including but not limited to, the randomization protocol was not described, and the hysterectomy patients (bilateral compared to unilateral salpingo-oophorectomy) were not blinded to their treatment allocation, small sample size limits the ability to rule out the role of chance as an explanation of study findings, and a discrepancy between reported outcomes in text and data tables. The authors recommended that more high quality studies are needed that compare embolization, with other treatments, including surgical treatments, hormonal therapy, and other noninvasive treatments.

**Randomized Controlled Trials**

No randomized controlled trials have been published comparing embolization therapy for pelvic congestion syndrome to an alternative or sham/placebo treatment. Randomized controlled trials are especially needed in situations such as this where the primary symptom is pain, a subjective outcome for which a placebo response to treatment is likely.

**Nonrandomized Studies**

The remainder of the published literature regarding the clinical outcomes of embolization therapy consists of case series and retrospective reviews.[5-26] Collectively, conclusions concerning safety and effectiveness cannot be reached from these studies due to significant limitations in the data, including but not limited to:

- Lack of established diagnostic criteria for pelvic congestion syndrome. Without consistent criteria for patient selection it is unknown which patients are most likely to benefit, or not benefit, from treatment. Furthermore, it is unknown how results from the various case series can be applied to the overall population of patients with this condition.
- Lack of randomization and comparison groups. Failure to randomize patients to different treatment groups may introduce bias on the part of both the study participant and researchers in favor of the new technology. As noted above, for pain treatments, a comparator (preferably sham treatment) is necessary, in order to guard against this bias and to distinguish treatment from placebo effects.
- Retrospective design and failure to control for other treatments. Retrospective study designs do not allow for control of co-treatments or confounding factors that may influence results. This design may also introduce bias to interpretation of results. Control for additional factors, such as other medical therapies, is necessary to isolate treatment response to embolization therapy.
- Failure to define relevant study endpoints. Bias may also be introduced by failure to define study endpoints and treatment success prior to commencement of the study.

**Adverse Effects**

The following adverse effects associated with embolization of the uterine and internal iliac veins, though uncommon, have been reported in the literature.[5,13]

- Embolization of coils to the pulmonary circulation
- Embolization of coils to the renal circulation
- Accidental embolization of glue fragments
- Perforations of the ovarian vein with extravasation of contrast
- Transient cardiac arrhythmia
Treatment of Varicoceles

Systematic Reviews

In 2012 Kroese published results from a systematic review and meta-analysis that examined the effect of treatment, surgery or embolization, for varicoceles in subfertile men. Ten studies were included in the review, which comprised 894 men. The authors concluded that there is evidence to suggest treatment improves a couple’s chance of pregnancy; however, findings are inconclusive. Furthermore, the available evidence is of low quality and limited to men from couples with subfertility problems. Therefore further research is needed to determine the efficacy of treatment, surgery or embolization, for the treatment of varicoceles.

Randomized-Controlled Trials

No randomized controlled trials have been published comparing embolization therapy for the treatment of varicoceles to an alternative or sham/placebo treatment. Randomized controlled trials are especially needed in situations such as this where the primary symptom is pain, a subjective outcome for which a placebo response to treatment is likely.

Nonrandomized studies

The remainder of the published literature regarding the clinical outcomes of embolization therapy consists of case series and retrospective reviews. Collectively, conclusions concerning safety and effectiveness cannot be reached from these studies due to significant limitations in the data.

PRACTICE GUIDELINE SUMMARY

PELVIC CONGESTION SYNDROME

American Congress of Obstetricians and Gynecologists

No relevant policy positions on embolization for treating pelvic congestion syndrome were identified on the American Congress of Obstetricians and Gynecologists (ACOG) website.

Society for Vascular Surgery (SVS) and the American Venous Forum

The 2011 Society for Vascular Surgery (SVS) and the American Venous Forum (AVF) guidelines for the care of patients with varicose veins and associated chronic venous diseases provided a Grade 2B recommendation in favor of coil embolization, plugs, or transcatheter sclerotherapy for treatment of PCS. A Grade 2B recommendation is defined as a weak recommendation based on medium quality evidence.

SUMMARY

There is not enough research to show that embolization, ablation, or sclerotherapy improves long term health outcomes for people with pelvic congestion syndrome or varicoceles, compared to other forms of therapy. Therefore, embolization, ablation, or sclerotherapy of ovarian veins, internal iliac veins, or gonadal veins are considered investigational for the treatment of pelvic congestion syndrome or varicoceles.


## CODES

**NOTE:** There are no specific codes for ovarian and internal iliac vein embolization; however, the following codes may be used:

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>36012</td>
<td>Selective catheter placement, venous system: second order or more selective, branch (eg, left adrenal vein, petrosal sinus)</td>
</tr>
<tr>
<td></td>
<td>37241</td>
<td>Vascular embolization or occlusion, inclusive of all radiological supervision and interpretation, intraprocedural roadmapping, and imaging guidance necessary to complete the intervention; venous, other than hemorrhage (eg, congenital or acquired venous malformations, venous and capillary hemangiomas, varices, varicoceles)</td>
</tr>
<tr>
<td></td>
<td>75894</td>
<td>Transcatheter therapy, embolization, any method, radiological supervision and interpretation</td>
</tr>
</tbody>
</table>

**HCPCS** None

*Date of Origin:* October 2005

---

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage.

Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
**Transgender Services**

**Effective:** April 1, 2018

**Next Review:** September 2018
**Last Review:** November 2017

---

**IMPORTANT REMINDER**

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

---

**DESCRIPTION**

Medical and surgical treatments of gender dysphoria in transgender individuals involves psychotherapy, hormonal therapy and, in some cases, gender affirmation surgery.

---

**MEDICAL POLICY CRITERIA**

*Note:* Member contracts for covered services vary. Member contract language takes precedence over medical policy.

**I. Medical Treatments of Gender Dysphoria**

A. Psychotherapy may be considered **medically necessary** as a treatment of gender dysphoria

B. Continuous hormone therapy may be considered **medically necessary** as a treatment of gender dysphoria when all of the following criteria are met:

1. Clinical records document that the patient has the capacity to make fully informed decisions and consent for treatment; and

2. A licensed mental health professional has diagnosed gender dysphoria as defined by the DSM-5 criteria (see Appendix 1); and
3. At least one of the following criteria must be met for a period of 3 or more months prior to the initiation of hormone therapy:
   a. Documentation of living as the desired gender; and/or
   b. Psychotherapy with a licensed mental health professional.

II. Surgical Treatments of Gender Dysphoria may be considered *medically necessary* when either A. or B. are met:
   A. Gender affirmation surgery (see Policy Guidelines) may be considered *medically necessary* in the treatment of gender dysphoria when all of the following criteria are met:
      1. Age at least 18 years (Note: *age requirement will not be applied to mastectomy in Female-to-Male patients with documented provider determination of medical necessity of earlier intervention*); and
      2. Clinical records document that the patient has the capacity to make fully informed decisions and consent for treatment, and that any other mental health condition, if present, is adequately controlled; and
      3. At least 2 licensed mental health professionals have diagnosed gender dysphoria as defined by the DSM-5 criteria (see Appendix 1), and recommend surgical treatment (Note: *only 1 mental health professional referral is required for mastectomy in Female-To-Male patients*); and
      4. Documentation of continuous hormonal therapy for at least 12 months, unless there is a documented contraindication to hormonal therapy (Note: *hormonal therapy is not required prior to mastectomy in Female-To-Male patients*); and
      5. Twelve months of living in a gender role that is congruent with the patient’s gender identity.

   B. When the criteria in II.A. above are met or have been met, the following procedures may be considered *medically necessary* when clinical information is submitted expressly documenting that the particular requested procedure would improve otherwise documented significant gender dysphoria:
      1. Breast augmentation
      2. Hair removal
      3. Hair transplantation
      4. Nipple/areola reconstruction in the absence of concurrent or prior subcutaneous or simple/total mastectomy
      5. Mastopexy

III. Other than gender affirmation surgeries listed in the Policy Guidelines, and/or surgeries in criteria II above, additional treatments to change specific appearance characteristics are considered *not medically necessary* as treatments of gender dysphoria including, but not limited to the following:
   A. Abdominoplasty
   B. Blepharoplasty
C. Brow lift  
D. Calf implants  
E. Cheek/malar implants  
F. Chin/nose implants  
G. Collagen injections  
H. Face-lift  
I. Facial bone reduction  
J. Forehead lift  
K. Lip reduction  
L. Liposuction  
M. Neck tightening  
N. Pectoral implants  
O. Reduction thyroid chondroplasty  
P. Rhinoplasty  
Q. Suction-assisted lipoplasty of the waist  
R. Voice modification surgery  
S. Voice therapy/lessons  

IV. Reversal of gender affirmation surgery is considered not medically necessary as a treatment of gender dysphoria.

**NOTE:** A summary of the supporting rationale for the policy criteria is at the end of the policy.

**POLICY GUIDELINES**

**GENDER AFFIRMATION SURGERY**

Surgical treatment for gender dysphoria differs depending upon the birth gender of the individual. The World Professional Association for Transgender Health (WPATH) indicated that, “(p)hysicians who perform surgical treatments for gender dysphoria should be urologists, gynecologists, plastic surgeons, or general surgeons, and board-certified as such by the relevant national and/or regional association. Surgeons should have specialized competence in genital reconstructive techniques as indicated by documented supervised training with a more experienced surgeon.”[1]

**Female-To-Male (FTM)**

For females transitioning to males, the following procedures may be included as part of gender affirmation surgery:[1,2]

- Hysterectomy  
- Mastectomy (subcutaneous mastectomy [CPT code 19304] or simple/total mastectomy [CPT code 19303], which may include related nipple/areola reconstruction [CPT code
Note: The use of CPT code 19318 (reduction mammaplasty) is incorrect coding.

- Metoidioplasty
- Nipple/areola reconstruction related to subcutaneous or simple/total mastectomy with nipple/areola excision or repositioning
- Penile prostheses implantation
- Phallic reconstruction/Phalloplasty
- Salpingo-oophorectomy
- Scrotoplasty
- Testicular prostheses implantation
- Urethroplasty
- Vaginectomy

**Definitions:**

Subcutaneous mastectomy: skin-sparing mastectomy which removes tissue through an incision under the breast, leaving the skin, areola, and nipple intact.

Simple/total mastectomy: removal of the entire breast and commonly any excess skin, including the areola and nipple.

**Male-To-Female (MTF)**

For males transitioning to females, the following procedures may be included as part of gender affirmation surgery:[1]

- Clitoroplasty
- Labiaplasty
- Orchiectomy
- Penectomy
- Vaginoplasty

**CROSS REFERENCES**

1. Endometrial Ablation, Surgery, Policy No. 01
2. Cosmetic and Reconstructive Surgery, Surgery, Policy No. 12
3. Reconstructive Breast Surgery/Mastopexy, and Management of Breast Implants, Surgery, Policy No. 40
4. Reduction Mammaplasty, Surgery, Policy No. 60
5. Autologous Fat Grafting to the Breast and Adipose-derived Stem Cells, Surgery, Policy No. 182
6. Medication Policy Manual, Do a find (Ctrl+F) and enter drug name in the find bar to locate the appropriate policy.

**BACKGROUND**

This policy supports applicable professional association statements,[1,3-6] and is also intended to support the Affordable Care Act (ACA) Section 1557 final implementing regulations published on May 18, 2016, and applicable state requirements[7].
A clinical diagnosis of gender dysphoria is required prior to treatment of the disorder. Treatments typically include psychotherapy, hormone therapy and in some cases surgical gender affirmation procedures. Psychotherapy followed by hormone therapy is often the first medical treatment sought, although not all transgender individuals on hormone therapy choose to undergo gender-affirming surgery.[4]

**Psychotherapy**

Psychotherapy provided by a mental health professional typically includes an initial assessment of gender identity and dysphoria, the historical development of gender dysphoric feelings, and severity of resulting stress caused by the condition.[1] The goal of therapy is to assess, diagnose, and discuss treatment options, if needed, and is typically required prior to hormone therapy and/or surgical treatment.

**Hormone Therapy**

Hormone therapy is undertaken in order to feminize or masculinize individuals' bodies to conform to their desired gender identities. For transgender individuals, hormone replacement therapy (HRT) causes the development of many of the secondary sexual characteristics of their gender identity. Prescribed hormones differ depending upon the natal gender of the individual. For MTF individuals, hormone treatment may include estradiol, finasteride, and spironolactone. For FTM individuals, hormone treatment may include androgenic hormones such as testosterone.

**Surgical Treatment**

Surgical treatment for gender dysphoria differs depending upon the natal gender of the individual. For MTF individuals, surgery may involve removal of the testicles and penis and the creation of a pseudo vagina, clitoris, and labia. Complications of MTF genital surgery may include necrosis of the vagina and labia, neovaginal prolapse, fistulas from the bladder or bowel into the vagina, stenosis of the urethra, and small or short vaginas.[1,8]

For FTM individuals, surgery may involve removal of the uterus, ovaries, and vagina, and creation of a neophallus and scrotum with scrotal and/or penile prostheses. The creation of a neophallus for FTM patients is a multistage reconstructive procedure. Currently, techniques for penile reconstruction procedures vary and complications may include frequent urinary tract stenoses and fistulas, donor site scarring and necrosis of the neophallus.[1,2] In addition, breast size does not significantly decrease with hormonal therapy and as a result, FTM patients may choose to undergo mastectomy to remove breast tissue. For many patients this may be the only surgery undertaken.[1] Mastectomy may involve a complete resection of all breast tissue; however, the nipple/areola sparing technique is typically performed to preserve the nipple/areola.

There are various additional aesthetic surgical procedures which may be sought in order to complete the physical gender transformation and align an individual to their gender identity. However, conflicting opinions exist regarding whether these procedures are essential in treating gender dysphoria.

The WPATH recommends that patients, “engage in 12 continuous months of living in a gender role that is congruent with their gender identity…” prior to gender reassignment surgery so that patients may socially adjust to their desired gender role.[1] WPATH notes that changing a
gender role may have personal and social consequences which should be adequately explored prior to undergoing an irreversible surgery.

**EVIDENCE SUMMARY**

Evidence regarding the treatment of gender dysphoria in transgender individuals primarily consists of two systematic reviews consisting of small cohort studies. Randomized clinical trials (RCTs) comparing gender dysphoria treatments with the non-treatment are ideal, however, there are challenges in conducting RCTs to evaluate treatments of gender dysphoria due to several factors, such as small patient populations and ethical concerns regarding the high morbidity and mortality rates associated with non-treatment. Therefore, large RCTs are not anticipated. This policy relies on the following systematic reviews and non-randomized studies, as well as professional association recommendations to support applicable federal and state requirements.

**SYSTEMATIC REVIEWS**

Only one of two systematic reviews is considered good quality (Murad, 2007) and reported on the resolution of gender dysphoria psychiatric comorbidities, quality of life, and sexual satisfaction outcomes for individuals treated with both hormonal and surgical treatments for gender identity disorder (GID).[9]

In 2009, Murad assessed quality of life and other psychosocial outcomes of transgendered individuals with GID, receiving hormonal therapy as part of gender affirmation surgery.[9] Twenty-eight cohort studies were included in the review which included pooled data from 1,833 patients with GID (1,093 MTF and 801 FTM). Significant improvements were reported after gender affirmation compared to pre-treatment status: 80% of patients reported improvement in gender dysphoria (95% CI = 68-89%; 8 studies) 78% reported significant improvement in psychological symptoms (95% CI = 56-94%; 7 studies) 80% reported significant improvement in quality of life (95% CI = 72-88%; 16 studies); and 72% reported significant improvement in sexual function (95% CI = 60-81%; 15 studies). Significant study heterogeneity was reported for all outcomes. Although the authors acknowledge the low quality of evidence used in the analysis, gender affirmation that included hormonal interventions in patient with GID was thought to likely improve symptoms of gender dysphoria and overall quality of life.

In 2009, Elamin evaluated the use of sex steroids on cardiovascular risk in transgender individuals.[10] A total of 16 studies were included in the review with a total of 1,471 male-to-female (MTF) patients and 651 female-to-male (FTM) patients. Steroid use was associated with increased serum triglycerides in both MTF and FTM patients and a nonsignificant effect on HDL-cholesterol and systolic blood pressure in FTM patients. Authors noted that the quality of evidence was low due to methodological limitations of included studies, including but not limited to, heterogeneity of patient population and variable follow-up periods and uncontrolled study design.

**NONRANDOMIZED STUDIES**

Primary evidence is limited to cohort studies with a variety of methodological limitations, including but not limited to small sample size, short-term follow-up, lack of comparison group, and varied treatment methods. Despite these limitations, significant improvements in quality of life, psychological comorbidities, and sexual functioning were consistently reported in patients who received gender-confirming medical treatments.[11]
Imbimbo evaluated the clinical and psychosocial profile of male-to-female transgendered individuals who had undergone reconstructive surgery.[12] The average age of patients was 31 years old, 72% had high educational levels, half of patients' contemplated suicide at some point prior to surgery and 4% had attempted suicide. Improved sex life satisfaction was reported in 75% of patients, with almost all patients' reporting satisfaction with their new sexual status. Additional studies sought to evaluate the sociodemographic profile of transgender individuals with GID in an effort to better characterize and provide treatment for this population.[13]

Heylens assessed comorbidities and psychosocial factors at various phases of the gender affirmation process in 57 patients with GID.[14] The Symptom Checklist-90 (SCL-90) was administered at three time points: baseline, after the start of hormone therapy, and after sex reassignment surgery (SRS) (also known as [aka] gender affirmation surgery). Psychopathological parameters include overall psychoneurotic distress, anxiety, agoraphobia, depression, somatization, paranoid ideation/psychoticism, interpersonal sensitivity, hostility, and sleeping problems and the psychosocial parameters consist of relationship, living situation, employment, sexual contacts, social contacts, substance abuse, and suicide attempt. The greatest improvement in psychoneurotic distress was observed after the initiation of hormone therapy (p<0.001). In addition, significant decreases in anxiety, depression, interpersonal sensitivity and hostility were reported after hormone therapy. No significant differences were observed in pre- and postoperative assessments.

Fisher described clinical and sociodemographic features of 140 transmen (n=48) and transwomen (n=92) with GID and without affirmation surgery.[15] The following assessment tests were administered: the Body Uneasiness Test (a self-rating scale exploring different areas of body-related psychopathology), Symptom Checklist-90 Revised (a self-rating scale to measure psychological state), and the Bem Sex Role Inventory (a self-rating scale to evaluate gender role). Authors reported that transmen displayed significantly better social functioning than transwoman.

Gorin-Lazard reported a case series which assessed a variety of gender dysphoria symptoms with hormonal treatment preceding gender affirmation surgery. Pre- and post- hormone treatment self-esteem (Social Self-Esteem Inventory), mood (Beck Depression Inventory), QoL (Subjective Quality of Life Analysis), and global functioning (Global Assessment of Functioning) scores were compared in 49 patients.[16] Hormone therapy was reported to be an independent factor in greater self-esteem, a reduction in depression, and improved QoL scores.

Gomez-Gil evaluated symptoms of social distress, anxiety and depression in 187 transgendered individuals.[17] Of those included in the study, 120 had undergone hormonal sex-reassignment (SR) (aka gender affirmation) treatment and 67 had not. Social anxiety was assessed with the Social Anxiety and Distress Scale (SADS) and depression and anxiety were assessed with the Hospital Anxiety and Depression Scale (HADS). The non-hormone group was reported to be significantly younger than the treatment group (mean age 25.9 vs. 33.6 years, p=0.001) and was less likely to have undergone surgical interventions (p<0.001). After adjusting for confounding factors, the authors reported that patients who were receiving hormone treatment had significantly lower prevalence of depression, anxiety, and social anxiety than those not receiving hormones.
Johansson reported long-term (five-year) outcomes of transgendered individuals (n=42) with GID who had completely transitioned (n=32), were in progress (n=5) or who were on hormone therapy (n=5).[18] Authors reported that no patient regretted affirmation and clinicians rated the global outcome as favorable in 62% of the cases, compared to 95% according to the patients themselves, with no differences between the subgroups. At follow-up, more than 90% of patients reported stable or improved work situations, partner relations and sex-life. However 5-15% of patients reported dissatisfaction with hormonal treatment, results of surgery, total gender affirmation procedure, or their present general health.

Asscheman evaluated the long-term (one-year) effects of cross-sex hormones in 966 male-to-female (MTF) and 365 female-to-male (FTM) transgendered individuals.[19] MTF patients received different doses of estrogen and cyproterone acetate and FTM patients received parenteral/oral testosterone esters or testosterone gel. Hormone treatment levels varied at pre- and post-surgical affirmation time points. High mortality rates were reported in the MTF group when compared to the general population (51%); however, this increased rate was due to non-hormone-related causes such as suicide, acquired immunodeficiency syndrome (AIDS), cardiovascular disease, drug abuse and other unknown causes. No significant increase in mortality was observed in FTM patients compared to the general population.

**PRACTICE GUIDELINE SUMMARY**

**WORLD PROFESSIONAL ASSOCIATION FOR TRANSGENDER HEALTH**

The World Professional Association for Transgender Health (WPATH) is a multidisciplinary professional society representing the specialties of medicine, psychology, social sciences and law that has published clinical guidelines regarding health services for patients with gender disorders. In 2012, WPATH updated their evidence and consensus-based guideline regarding, the *Standards of Care (SOC) for the Health of Transsexual, Transgender, and Gender Nonconforming Peoples.*[1] WPATH listed the following options for individuals seeking treatment for gender dysphoria:

- Changes in gender expression and role (which may involve living part time or full time in another gender role, consistent with one’s gender identity);
- Hormone therapy to feminize or masculinize the body;
- Surgery to change primary and/or secondary sex characteristics;
- Psychotherapy (individual, couple, family, or group) for purposes such as exploring gender identity, role, and expression; addressing the negative impact of gender dysphoria and stigma on mental health; alleviating internalized transphobia; enhancing social and peer support; improving body image; or promoting resilience.

WPATH guidelines describe surgical procedures as “irreversible changes to the body.” Therefore, WPATH guidelines recommend the appropriate care should be taken to ensure patients have sufficient time (at least 24 hours) to consider all the information and can provide informed consent. WPATH notes, “(t)hese surgeries may be performed once there is written documentation that this assessment has occurred and that the person has met the criteria for a specific surgical treatment. By following this procedure, mental health professionals, surgeons, and patients share responsibility for the decision to make irreversible changes to the body.”

**Physical Interventions for Adolescents**
WPATPATH guidelines state that physical interventions for adolescents fall into three categories or stages:

1. Fully reversible interventions. These involve the use of GnRH analogues to suppress estrogen or testosterone production and consequently delay the physical changes of puberty. Alternative treatment options include progestins (most commonly medroxyprogesterone) or other medications (such as spironolactone) that decrease the effects of androgens secreted by the testicles of adolescents who are not receiving GnRH analogues. Continuous oral contraceptives (or depot medroxyprogesterone) may be used to suppress menses.

2. Partially reversible interventions. These include hormone therapy to masculinize or feminize the body. Some hormone-induced changes may need reconstructive surgery to reverse the effect (e.g., gynaecomastia caused by estrogens), while other changes are not reversible (e.g., deepening of the voice caused by testosterone).

3. Irreversible interventions. Reversible and irreversible interventions are outlined in the standards of care, specifying intervention sequencing in adolescents. It is also stated that “[t]wo goals justify intervention with puberty suppressing hormones: (i) their use gives adolescents more time to explore their gender nonconformity and other developmental issues; and (ii) their use may facilitate transition by preventing the development of sex characteristics that are difficult or impossible to reverse if adolescents continue on to pursue sex reassignment.”

**Referral for Surgery**

WPATPATH guidelines indicate that surgical treatments can be initiated by a referral from a qualified mental health professional. One or two referrals may be required depending upon the type of surgery requested. “The mental health professional provides documentation—in the chart and/or referral letter—of the patient’s personal and treatment history, progress, and eligibility.” WPATPATH guidelines specifically recommend the following:

- One referral from a qualified mental health professional is needed for breast/chest surgery (e.g., mastectomy, chest reconstruction, or augmentation mammoplasty).
- Two referrals—from qualified mental health professionals who have independently assessed the patient—are needed for genital surgery (i.e., hysterectomy/salpingo-oophorectomy, orchiectomy, genital reconstructive surgeries).

**Criteria for Breast/Chest Surgery (One Referral)**

WPATPATH lists the following criteria for mastectomy and creation of a male chest in FTM patients:

1. Persistent, well-documented gender dysphoria;
2. Capacity to make a fully informed decision and to consent for treatment;
3. Age of majority in a given country;
4. If significant medical or mental health concerns are present, they must be reasonably well controlled.

Hormone therapy is not a prerequisite.
Criteria for Genital Surgery (Two Referrals)

WPATH lists the following criteria for genital surgery:

1. Persistent, well-documented gender dysphoria;
2. Capacity to make a fully informed decision and to consent for treatment;
3. Age of majority in a given country;
4. If significant medical or mental health concerns are present, they must be reasonably well controlled.
5. 12 continuous months of hormone therapy as appropriate to the patient’s gender goals (unless hormones are not clinically indicated for the individual).

In addition, WPATH made specific recommendations regarding breast augmentation procedures:

Breast Augmentation

The WPATH guideline recommends MTF patients undergo feminizing hormone therapy for a minimum of 12 months prior to augmentation surgery and lists specific criteria for breast augmentation (implants/lipofilling).

THE ENDOCRINE SOCIETY

In 2017, the Endocrine Society in conjunction with American Association of Clinical Endocrinologists, American Society of Andrology, European Society for Pediatric Endocrinology, European Society of Endocrinology, Pediatric Endocrine Society, and World Professional Association for Transgender Health published updated guidelines for the treatment of gender-dysphoric/gender-incongruent persons.[20] The guideline employed transparent methods for evidence review and for rating the quality of evidence. Guidelines were referenced as recommendations or suggestions, by the numbers 1 and 2, respectively. Evidence was ranked as very low-quality | † † † †; low quality | † † † †; moderate quality | † † † †; and high quality | † † † † †. The consortium made the following statements:

1.0 Evaluation of Youth and Adults

1.1 We advise that only trained mental health professionals (MHPs) who meet the following criteria should diagnose gender dysphoria (GD)/gender incongruence in adults: (1) competence in using the Diagnostic and Statistical Manual of Mental Disorders (DSM) and/or the International Statistical Classification of Diseases and Related Health Problems (ICD) for diagnostic purposes, (2) the ability to diagnose GD/gender incongruence and make a distinction between GD/gender incongruence and conditions that have similar features (e.g., body dysmorphic disorder), (3) training in diagnosing psychiatric conditions, (4) the ability to undertake or refer for appropriate treatment, (5) the ability to psychosocially assess the person’s understanding, mental health, and social conditions that can impact gender-affirming hormone therapy, and (6) a practice of regularly attending relevant professional meetings. (Ungraded Good Practice Statement)

1.2. We advise that only MHPs who meet the following criteria should diagnose GD/gender incongruence in children and adolescents: (1) training in child and adolescent developmental psychology and psychopathology, (2) competence in using the DSM and/or the ICD for diagnostic purposes, (3) the ability to make a distinction between GD/gender incongruence and conditions that have similar features (e.g., body
dysmorphic disorder), (4) training in diagnosing psychiatric conditions, (5) the ability to undertake or refer for appropriate treatment, (6) the ability to psychosocially assess the person’s understanding and social conditions that can impact gender-affirming hormone therapy, (7) a practice of regularly attending relevant professional meetings, and (8) knowledge of the criteria for puberty blocking and gender-affirming hormone treatment in adolescents. (Ungraded Good Practice Statement)

1.3. We advise that decisions regarding the social transition of prepubertal youths with GD/gender incongruence are made with the assistance of an MHP or another experienced professional. (Ungraded Good Practice Statement). 

1.4. We recommend against puberty blocking and gender-affirming hormone treatment in prepubertal children with GD/gender incongruence. (1 | ⬤ ⬤ ⬤ ⬤)

1.5. We recommend that clinicians inform and counsel all individuals seeking gender-affirming medical treatment regarding options for fertility preservation prior to initiating puberty suppression in adolescents and prior to treating with hormonal therapy of the affirmed gender in both adolescents and adults. (1 | ⬤ ⬤ ⬤ ⬤)

2.0 Treatment of Adolescents

2.1. We suggest that adolescents who meet diagnostic criteria for GD/gender incongruence, fulfill criteria for treatment, and are requesting treatment should initially undergo treatment to suppress pubertal development. (2 | ⬤ ⬤ ⬤ ⬤)

2.2. We suggest that clinicians begin pubertal hormone suppression after girls and boys first exhibit physical changes of puberty. (2 | ⬤ ⬤ ⬤ ⬤)

2.3. We recommend that, where indicated, GnRH analogues are used to suppress pubertal hormones. (1 | ⬤ ⬤ ⬤ ⬤)

2.4. In adolescents who request sex hormone treatment (given this is a partly irreversible treatment), we recommend initiating treatment using a gradually increasing dose schedule after a multidisciplinary team of medical and MHPs has confirmed the persistence of GD/gender incongruence and sufficient mental capacity to give informed consent, which most adolescents have by age 16 years. (1 | ⬤ ⬤ ⬤ ⬤).

2.5. We recognize that there may be compelling reasons to initiate sex hormone treatment prior to the age of 16 years in some adolescents with GD/ gender incongruence, even though there are minimal published studies of gender-affirming hormone treatments administered before age 13.5 to 14 years. As with the care of adolescents 16 years of age, we recommend that an expert multidisciplinary team of medical and MHPs manage this treatment. (1 | ⬤ ⬤ ⬤ ⬤)

2.6. We suggest monitoring clinical pubertal development every 3 to 6 months and laboratory parameters every 6 to 12 months during sex hormone treatment. (2 | ⬤ ⬤ ⬤ ⬤)

3.0 Hormonal Therapy for Transgender Adults

3.1. We recommend that clinicians confirm the diagnostic criteria of GD/gender incongruence and the criteria for the endocrine phase of gender transition before beginning treatment. (1 | ⬤ ⬤ ⬤)

3.2. We recommend that clinicians evaluate and address medical conditions that can be exacerbated by hormone depletion and treatment with sex hormones of the affirmed gender before beginning treatment. (1 | ⬤ ⬤ ⬤)

3.3. We suggest that clinicians measure hormone levels during treatment to ensure that endogenous sex steroids are suppressed and administered sex steroids are maintained in the normal physiologic range for the affirmed gender. (2 | ⬤ ⬤ ⬤ ⬤)
3.4. We suggest that endocrinologists provide education to transgender individuals undergoing treatment about the onset and time course of physical changes induced by sex hormone treatment. (2 | ★★★★)

4.0 Adverse Outcome Prevention and Long-term Care

4.1. We suggest regular clinical evaluation for physical changes and potential adverse changes in response to sex steroid hormones and laboratory monitoring of sex steroid hormone levels every 3 months during the first year of hormone therapy for transgender males and females and then once or twice yearly. (2 | ★★★)

4.2. We suggest periodically monitoring prolactin levels in transgender females treated with estrogens. (2 | ★★★)

4.3. We suggest that clinicians evaluate transgender persons treated with hormones for cardiovascular risk factors using fasting lipid profiles, diabetes screening, and/or other diagnostic tools. (2 | ★★★)

4.4. We recommend that clinicians obtain bone mineral density (BMD) measurements when risk factors for osteoporosis exist, specifically in those who stop sex hormone therapy after gonadectomy. (1 | ★★★)

4.5. We suggest that transgender females with no known increased risk of breast cancer follow breast-screening guidelines recommended for non-transgender females. (2 | ★★★)

4.6. We suggest that transgender females treated with estrogens follow individualized screening according to personal risk for prostatic disease and prostate cancer. (2 | ★★★)

4.7. We advise that clinicians determine the medical necessity of including a total hysterectomy and oophorectomy as part of gender-affirming surgery. (Ungraded Good Practice Statement)

5.0 Surgery for Sex Reassignment and Gender Confirmation

5.1. We recommend that a patient pursue genital gender-affirming surgery only after the MHP and the clinician responsible for endocrine transition therapy both agree that surgery is medically necessary and would benefit the patient’s overall health and/or well-being. (1 | ★★★)

5.2. We advise that clinicians approve genital gender affirming surgery only after completion of at least 1 year of consistent and compliant hormone treatment, unless hormone therapy is not desired or medically contraindicated. (Ungraded Good Practice Statement)

5.3. We advise that the clinician responsible for endocrine treatment and the primary care provider ensure appropriate medical clearance of transgender individuals for genital gender-affirming surgery and collaborate with the surgeon regarding hormone use during and after surgery. (Ungraded Good Practice Statement)

5.4. We recommend that clinicians refer hormone treated transgender individuals for genital surgery when: (1) the individual has had a satisfactory social role change, (2) the individual is satisfied about the hormonal effects, and (3) the individual desires definitive surgical changes. (1 | ★★★)

5.5. We suggest that clinicians delay gender-affirming genital surgery involving gonadectomy and/or hysterectomy until the patient is at least 18 years old or legal age of majority in his or her country. (2 | ★★★)

5.6. We suggest that clinicians determine the timing of breast surgery for transgender males based upon the physical and mental health status of the individual. There is insufficient evidence to recommend a specific age requirement. (2 | ★★★)
AMERICAN COLLEGE OF OBSTETRICIANS AND GYNECOLOGY

In 2011, American College of Obstetricians and Gynecology (ACOG) published a committee opinion regarding health care services for transgendered individuals. Although this guideline is not based in evidence, ACOG does make the following recommendations, “Obstetrician–gynecologists should be prepared to assist or refer transgender individuals for routine treatment and screening as well as hormonal and surgical therapies. Hormonal and surgical therapies for transgender patients may be requested, but should be managed in consultation with health care providers with expertise in specialized care and treatment of transgender patients.”

In addition, ACOG guidelines made specific recommendations regarding hormone therapy, surgery and screening for both female-to-male and male-to-female patients:

Female-to-Male Transgender Individuals

Hormones

Methyltestosterone injections every 2 weeks are usually sufficient to suppress menses and induce masculine secondary sex characteristics. Before receiving androgen therapy, patients should be screened for medical contraindications and have periodic laboratory testing, including hemoglobin and hematocrit to evaluate for polycythemia, liver function tests, and serum testosterone level assessments (goal is a mid-normal male range of 500 microgram/dL), while receiving the treatment.

Surgery

Hysterectomy, with or without salpingo-oophorectomy, is commonly part of the surgical process. An obstetrician–gynecologist who has no specialized expertise in transgender care may be asked to perform this surgery, and also may be consulted for routine reasons such as dysfunctional bleeding or pelvic pain. Reconstructive surgery should be performed by a urologist, gynecologist, plastic surgeon, or general surgeon who has specialized competence and training in this field.

Screening

Age-appropriate screening for breast cancer and cervical cancer should be continued unless mastectomy or removal of the cervix has occurred. For patients using androgen therapy who have not had a complete hysterectomy, there may be an increased risk of endometrial cancer and ovarian cancer.

Male-to-Female Transgender Individuals

Hormones

Estrogen therapy results in gynecomastia, reduced hair growth, redistribution of fat, and reduced testicular volume. All patients considering therapy should be screened for medical contraindications. After surgery, doses of estradiol, 2–4 mg/d, or conjugated equine estrogen, 2.5 mg/d, are often sufficient to keep total testosterone levels to normal female levels of less than 25 ng/dL. Nonoral therapy also can be offered. It is recommended that male-to-female transgender patients receiving estrogen therapy have an annual prolactin level assessment and visual field examination to screen for prolactinoma.
Surgery

Surgery usually involves penile and testicular excision and the creation of a neovagina. Reported complications of surgery include vaginal and urethral stenosis, fistula formation, problems with remnants of erectile tissue, and pain. Vaginal dilation of the neovagina is required to maintain patency. Other surgical procedures that may be performed include breast implants and nongenital surgery, such as facial feminization surgery.

Screening

Age-appropriate screening for breast and prostate cancer is appropriate for male-to-female transgender patients. Opinion varies regarding the need for Pap testing in this population. In patients who have a neocervix created from the glans penis, routine cytologic examination of the neocervix may be indicated. The glands are more prone to cancerous changes than the skin of the penile shaft, and intraepithelial neoplasia of the glans is more likely to progress to invasive carcinoma than is intraepithelial neoplasia of other penile skin.

SUMMARY

The research lacks well-designed studies comparing the safety and effectiveness of non-treatment for gender dysphoria with treatments such as hormone therapy and gender affirmation surgery. However, there are challenges in conducting large studies to evaluate existing treatments, and such studies are not expected in the near future. Although additional research is needed, the research has consistently suggested significant improvement in symptoms and overall quality of life in those who have received treatment for gender dysphoria. Therefore, treatment of gender dysphoria in transgender individuals may be considered medically necessary when specified policy criteria are met.

The World Professional Association for Transgender Health (WPATH) Standards of Care (SOC) for the Health of Transsexual, Transgender, and Gender Nonconforming Peoples describe reversible and irreversible interventions, and the ideal order and timing of these approaches. Surgery as an intervention is considered irreversible by WPATH. Therefore, reversal of gender affirmation surgery is considered not medically necessary as a treatment of gender dysphoria.

REFERENCES


6. Resolution #114. American Medical Association House of Delegates. “Removing Barriers to Care for Transgender Patients” [cited; Available from:


### CODES

**NOTES:**

- Codes 31552, 31554, 31580, 31584, 31587, and 31591 are not appropriate to use to represent voice modification. Unlisted code 31599 should be reported instead.
- Codes 55970 and 55980 are non-specific. The specific procedure code(s) must be requested in place of these non-specific codes.

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>11970</td>
<td>Replacement of tissue expander with permanent prosthesis</td>
</tr>
<tr>
<td></td>
<td>11971</td>
<td>Removal of tissue expander(s) without insertion of prosthesis</td>
</tr>
<tr>
<td></td>
<td>15775</td>
<td>Punch graft for hair transplant; 1 to 15 punch grafts</td>
</tr>
<tr>
<td></td>
<td>15776</td>
<td>Punch graft for hair transplant; more than 15 punch grafts</td>
</tr>
<tr>
<td></td>
<td>15820</td>
<td>Blepharoplasty, lower eyelid</td>
</tr>
<tr>
<td></td>
<td>15821</td>
<td>; with extensive herniated fat pad</td>
</tr>
<tr>
<td></td>
<td>15822</td>
<td>Blepharoplasty, upper eyelid</td>
</tr>
<tr>
<td></td>
<td>15823</td>
<td>; with excessive skin weighting down lid</td>
</tr>
<tr>
<td></td>
<td>17380</td>
<td>Electrolysis epilation, each 30 minutes</td>
</tr>
<tr>
<td></td>
<td>19303</td>
<td>Mastectomy, simple, complete</td>
</tr>
<tr>
<td></td>
<td>19304</td>
<td>Mastectomy, subcutaneous</td>
</tr>
<tr>
<td></td>
<td>19316</td>
<td>Mastopexy</td>
</tr>
<tr>
<td></td>
<td>19318</td>
<td>Reduction mammoplasty</td>
</tr>
<tr>
<td></td>
<td>19324</td>
<td>Mammoplasty, augmentation; without prosthetic implant</td>
</tr>
<tr>
<td></td>
<td>19325</td>
<td>Mammoplasty, augmentation; with prosthetic implant</td>
</tr>
<tr>
<td></td>
<td>19350</td>
<td>Nipple/areola reconstruction</td>
</tr>
<tr>
<td></td>
<td>30400</td>
<td>Rhinoplasty, primary; lateral and alar cartilages and/or elevation of nasal tip</td>
</tr>
<tr>
<td></td>
<td>30410</td>
<td>; complete, external parts including bony pyramid, lateral and alar cartilages, and/or elevation of nasal tip</td>
</tr>
<tr>
<td></td>
<td>30420</td>
<td>; including major septal repair</td>
</tr>
<tr>
<td></td>
<td>30430</td>
<td>Rhinoplasty, secondary; minor revision (small amount of nasal tip work)</td>
</tr>
<tr>
<td></td>
<td>30435</td>
<td>; intermediate revision (bony work with osteotomies)</td>
</tr>
<tr>
<td></td>
<td>30450</td>
<td>; major revision (nasal tip work and osteotomies)</td>
</tr>
<tr>
<td></td>
<td>31599</td>
<td>Unlisted procedure, larynx</td>
</tr>
<tr>
<td></td>
<td>53400</td>
<td>Urethroplasty; first stage, for fistula, diverticulum, or stricture (eg, Johannsen type)</td>
</tr>
<tr>
<td></td>
<td>53405</td>
<td>Urethroplasty; second stage (formation of urethra), including urinary diversion</td>
</tr>
<tr>
<td></td>
<td>53410</td>
<td>Urethroplasty, 1-stage reconstruction of male anterior urethra</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>53415</td>
<td>Urethroplasty, transpubic or perineal, 1-stage, for reconstruction or repair of prostatic or membranous urethra</td>
<td></td>
</tr>
<tr>
<td>53420</td>
<td>Urethroplasty, 2-stage reconstruction or repair of prostatic or membranous urethra; first stage</td>
<td></td>
</tr>
<tr>
<td>53425</td>
<td>Urethroplasty, 2-stage reconstruction or repair of prostatic or membranous urethra; second stage</td>
<td></td>
</tr>
<tr>
<td>53430</td>
<td>Urethroplasty, reconstruction of female urethra</td>
<td></td>
</tr>
<tr>
<td>54125</td>
<td>Insertion of testicular prosthesis</td>
<td></td>
</tr>
<tr>
<td>54200</td>
<td>Laparoscopy, surgical; orchietomy</td>
<td></td>
</tr>
<tr>
<td>55175</td>
<td>Scrotoplasty; simple</td>
<td></td>
</tr>
<tr>
<td>55180</td>
<td>Scrotoplasty; complicated</td>
<td></td>
</tr>
<tr>
<td>55899</td>
<td>Phallic reconstruction/Phalloplasty (Unlisted procedure, male genital system)</td>
<td></td>
</tr>
<tr>
<td>55970</td>
<td>intersex surgery; male to female</td>
<td></td>
</tr>
<tr>
<td>55980</td>
<td>intersex surgery; female to male</td>
<td></td>
</tr>
<tr>
<td>56625</td>
<td>Vulvectomy simple; complete</td>
<td></td>
</tr>
<tr>
<td>56800</td>
<td>Plastic repair of introitus</td>
<td></td>
</tr>
<tr>
<td>56805</td>
<td>Clitroplasty for intersex state</td>
<td></td>
</tr>
<tr>
<td>57106</td>
<td>Vaginectomy, partial removal of vaginal wall</td>
<td></td>
</tr>
<tr>
<td>57110</td>
<td>Vaginectomy, complete removal of vaginal wall;</td>
<td></td>
</tr>
<tr>
<td>57291</td>
<td>Construction of artificial vagina; without graft</td>
<td></td>
</tr>
<tr>
<td>57292</td>
<td>Construction of artificial vagina; with graft</td>
<td></td>
</tr>
<tr>
<td>57295</td>
<td>Revision (including removal) of prosthetic vaginal graft; vaginal approach</td>
<td></td>
</tr>
<tr>
<td>57296</td>
<td>Revision (including removal) of prosthetic vaginal graft; open abdominal approach</td>
<td></td>
</tr>
<tr>
<td>57335</td>
<td>Vaginoplasty for intersex state - the physician uses various plastic surgery techniques to correct a small, underdeveloped vagina due to the overproduction of male hormones</td>
<td></td>
</tr>
<tr>
<td>57426</td>
<td>Revision (including removal) of prosthetic vaginal graft, laparoscopic approach</td>
<td></td>
</tr>
<tr>
<td>58150</td>
<td>Total abdominal hysterectomy (corpus and cervix), with or without removal of tube(s), with or without removal of ovary(s)</td>
<td></td>
</tr>
<tr>
<td>58180</td>
<td>Supracervical abdominal hysterectomy (subtotal hysterectomy), with or without removal of tube(s), with or without removal of ovary(s)</td>
<td></td>
</tr>
<tr>
<td>58260</td>
<td>Vaginal hysterectomy, for uterus 250 g or less</td>
<td></td>
</tr>
<tr>
<td>58262</td>
<td>Vaginal hysterectomy, for uterus 250 g or less; with removal of tube(s), and/or ovary(s)</td>
<td></td>
</tr>
<tr>
<td>58270</td>
<td>Vaginal hysterectomy, for uterus 250 g or less; with repair of enterocele</td>
<td></td>
</tr>
<tr>
<td>58275</td>
<td>Vaginal hysterectomy, with total or partial vaginectomy;</td>
<td></td>
</tr>
<tr>
<td>58290</td>
<td>Vaginal hysterectomy, for uterus greater than 250 g</td>
<td></td>
</tr>
<tr>
<td>58291</td>
<td>Vaginal hysterectomy, for uterus greater than 250 g; with removal of tube(s) and/or ovary(s)</td>
<td></td>
</tr>
<tr>
<td>58541</td>
<td>Laparoscopy, surgical, supracervical hysterectomy, for uterus 250 g or less</td>
<td></td>
</tr>
<tr>
<td>58542</td>
<td>Laparoscopy, surgical, supracervical hysterectomy, for uterus 250 g or less; with removal of tube(s) and/or ovary(s)</td>
<td></td>
</tr>
<tr>
<td>58543</td>
<td>Laparoscopy, surgical, supracervical hysterectomy, for uterus greater than 250 g</td>
<td></td>
</tr>
<tr>
<td>58544</td>
<td>Laparoscopy, surgical, supracervical hysterectomy, for uterus greater than 250 g; with removal of tube(s) and/or ovary(s)</td>
<td></td>
</tr>
<tr>
<td>58550</td>
<td>Laparoscopy, surgical, with vaginal hysterectomy, for uterus 250 g or less</td>
<td></td>
</tr>
<tr>
<td>58552</td>
<td>Laparoscopy, surgical, with vaginal hysterectomy, for uterus 250 g or less; with removal of tube(s) and/or ovary(s)</td>
<td></td>
</tr>
</tbody>
</table>
Gender dysphoria is defined by the Diagnostic and Statistical Manual of Mental Disorders DSM-5V as:

**Gender Dysphoria in Children:**

A. A marked incongruence between one's experienced/expressed gender and assigned gender, of at least 6 months duration, as manifested by at least six of the following (one of which must be Criterion A1):

1. A strong desire to be of the other gender or an insistence that one is the other gender (or some alternative gender, different from one's assigned gender)
2. In boys (assigned gender), a strong preference for cross dressing or simulating female attire; or in girls (assigned gender), a strong preference for wearing only typical masculine clothing and a strong resistance to wearing of typical feminine clothing.
3. A strong preference for cross-gender roles in make-believe play of fantasy play.
4. A strong preference for toys, games, or activities stereotypically used or engaged in by the other gender.
5. A strong preference for playmates of the other gender.
6. In boys (assigned gender), a strong rejection of typically masculine toys, games and activities and a strong avoidance of rough and tumble play; or in girls (assigned gender), a strong rejection of typically feminine toys, games and activities.
7. A strong dislike of one's sexual anatomy.
8. A strong desire for the primary and/or secondary sex characteristics that match one's experienced gender.

B. The condition is associated with clinically significant distress or impairment in social, school, or other important areas of functioning.

*Specify if:*
APPENDIX 1

With a disorder of sex development (e.g., a congenital adrenogenital disorder such as 2.55.2 [E25.0] congenital adrenal hyperplasia or 259.0 [E34.50] androgen insensitivity syndrome)

Coding note: Code the disorder of sex development as well as gender dysphoria.

Gender Dysphoria in Adolescents and Adults:

A. A marked incongruence between one's experienced/expressed gender and assigned gender, of at least 6 months duration, as manifested by at least two of the following:

1. A marked incongruence between one's experienced/expressed gender and primary and/or secondary sex characteristics (on in young adolescents, the anticipated secondary sex characteristics).
2. A strong desire to be rid of one's primary and/or secondary sex characteristics because of a marked incongruence with one's experienced/expressed gender (on in young adolescents, a desire to prevent the development of the anticipated secondary sex characteristics.)
3. A strong desire for the primary and/or secondary sex characteristics of the other gender.
4. A strong desire to be of the other gender) or some alternative gender different from one's assigned gender).
5. A strong desire to be treated as the other gender (or some alternative gender different from one's assigned gender).
6. A strong conviction that one has the typical feelings and reactions of the other gender (or some alternative gender different from one's assigned gender).

B. The condition is associated with clinically significant distress or impairment in social, occupational, or other important areas of functioning.

Specify if:

With a disorder of sex development (e.g., a congenital adrenogenital disorder such as 2.55.2 [E25.0] congenital adrenal hyperplasia or 259.0 [E34.50] androgen insensitivity syndrome)

Coding note: Code the disorder of sex development as well as gender dysphoria.

Specify if:

Post transition: The individual has transitioned to full-time living in the desired gender (with or without legalization of gender change) and has undergone (or is preparing to have) at least one cross-sex medical procedure or treatment regimen—namely regular cross-sex treatment or gender reassignment surgery confirming the desired gender (e.g., appendectomy, vaginoplasty in the natal male; mastectomy or phalloplasty in the natal female)."

Date of Origin: September 2014
UMP Transgender Policy Supplement for Male to Female Breast Reconstruction

**Note:** This policy is to be used in conjunction with the UMP specific Transgender Services Clinical Criteria and Policy. Both policies’ criteria must be met for preauthorization to be considered for male to female breast reconstruction.

Breast reconstruction (M to F) will require preauthorization with following criteria:

- Documentation of continuous hormonal therapy for at least 12 months, unless there is a documented contraindication to hormonal therapy; **AND**
- Documentation from surgeon of current cup size and proposed changes as well as photo documentation; **AND**
- No measureable cup size growth, defined as less than an A cup, in one or both breasts; **OR**
- Asymmetry where one breast did not have a measureable cup size growth, defined as less than an A cup.

Example: Client presents with response with one breast B cup and one breast A cup = **NON-COVERED**
Example: Client presents with response with one breast B cup and one breast with no measurable cup size = **COVERED**
Uniform Medical Plan (UMP) Transgender Services:
Clinical Criteria and Policy

Policy effective 01-01-17; Revised 06/01/18

UMP members should refer to Regence medical policy 153 for information about UMP’s coverage of transgender services, with the exception of information in the “Medical Policy Criteria” box in policy 153. Instead of the criteria listed in that box, the UMP-specific clinical criteria outlined below must be met to receive transgender surgical services.

I. Medical Treatments for Gender Dysphoria

A. Psychotherapy may be considered medically necessary as a treatment of gender dysphoria.

B. Continuous hormone therapy may be considered medically necessary as a treatment of gender dysphoria when all of the following criteria are met:

1. Clinical records document that the patient has the capacity to make fully informed decisions and consent for treatment; and hormone therapy is part of a comprehensive, patient-centered treatment plan; and

2. A licensed behavioral health practitioner or a licensed physician, advanced registered nurse practitioner (ARNP), physician’s assistant (PA) or psychologist is treating the patient for primary care or transgender services and:

   a. Assesses the patient and makes or confirms the diagnosis of gender dysphoria as defined by the DSM-V criteria, and

   b. Determines or confirms that the gender dysphoria is not due to another mental or physical health condition.

3. Providers diagnosing and treating patients for medical treatment for gender dysphoria must document these minimum credentials and competencies as part of a comprehensive, patient-centered treatment plan:

   a. Meet the requirements of professional licensure and practice according to the scope of practice for their licensure.

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
b. Attest to specialized competencies in managing hormone therapy for persons diagnosed with gender dysphoria which may include documented, supervised training or mentoring with a more experienced physician.

c. Be knowledgeable of and practice the standards of care for the health of transsexual, transgender, and gender-nonconforming people as developed by the World Professional Association for Transgender Health (WPATH).

II. Surgical Treatments of Gender Dysphoria

A. Gender reassignment surgery (see UMP clinical criteria policy and Regence medical policy 153 guidelines) may be considered medically necessary in the treatment of gender dysphoria when all of the following criteria are met:

1. Age at least 18 years. For patients younger than 18 years of age, mastectomy may be considered medically necessary in female to male surgical procedures. Other requirements outlined in this section must be met to proceed with mastectomy in those younger than 18 years of age.

2. Clinical records document that the patient has the capacity to make fully informed decisions and consent for treatment as part of a comprehensive, patient-centered treatment plan; and that any other mental health condition, if present, is adequately controlled.

3. The multidisciplinary treatment team must have documented the diagnosis of gender dysphoria and recommend surgical treatment as part of a comprehensive, patient-centered plan of care. The plan of care and recommendation for surgical treatment must meet the criteria in sections a. through d. below.

a. The multidisciplinary treatment team consists of the following: two licensed mental health professionals,* the medical provider who has managed the hormone therapy and primary medical care and/or transgender services prior to surgical evaluation, and the surgeon(s) recommending the surgical procedures; and

November 1, 2018

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
Only one mental health professional referral is required for mastectomy in female-to-male patients.

b. A surgical evaluation by a surgeon(s) who will perform the gender reassignment surgery as part of a comprehensive, patient-centered plan of care. Upon completion, the surgeon must forward the results of the surgical evaluation and recommendations for surgical treatment to other treatment team members; and

c. Plan of care documentation must include the patient’s signature to document understanding of the treatment plan, surgical treatment, risks and benefits of the surgery; and

d. A comprehensive referral letter for surgery, written and signed by a member of the treatment team, with a prior authorization request for surgery must be submitted to the plan.

4. Documentation of continuous hormonal therapy for at least 12 months, unless there is a documented contraindication to hormonal therapy. Hormonal therapy is not required prior to mastectomy in female-to-male patients.

5. Twelve months of living in a gender role that is congruent with the patient’s gender identity.

6. If the referring medical provider or mental health provider requests surgical intervention prior to the patient’s completion of 12 months of hormone therapy and living in desired gender, the multidisciplinary treatment team must submit evidence of medical necessity and clear rationale for the proposed surgical intervention. The multidisciplinary treatment team must submit written documentation to the plan that includes:

   a. A comprehensive, coordinated treatment plan with evidence that all treatment plan criteria for surgery and treatment goals have been met; and
b. Clear rationale for the variation from the 12-month period for either/or hormone therapy and living in desired gender; and

c. Documentation that the proposed surgical provider accepts the treatment plan and surgical intervention proposed by the coordinated clinical team’s treatment plan with less than 12 months living in desired gender and on hormone therapy; and

d. Patient understands the treatment plan, risks and benefits of surgery prior to completing the 12-month period; and

e. The plan will determine authorization and consent to care based on medical necessity from the documentation outlined in II.A.

B. Prior authorization is required for all proposed surgical interventions. Section II.A of this policy lists the requirements and documentation that must be submitted for prior authorization review. Surgeries are not required to be completed at the same time and, instead, may be performed and receive prior authorization in progressive stages. UMP covers the following procedures with prior authorization that meet medical necessity criteria:

1. Abdominoplasty, which also meets restorative function medical criteria;
2. Blepharoplasty, which also meets restorative function medical criteria;
3. Breast reconstruction (male-to-female patients);
4. Bilateral mastectomy with or without chest reconstruction;
5. Cliteroplasty;
6. Colovaginoplasty;
7. Colpectomy;
8. Genital surgery;
9. Genital electrolysis as required as part of the genital surgery, including electrolysis of the graft site, as required for genital surgery;
10. Hysterectomy;
11. Labiaplasty;
12. Laryngoplasty;

These criteria do not imply or guarantee approval. Please check with your plan to ensure coverage. Preauthorization requirements are only valid for the month published. They may have changed from previous months and may change in future months.
13. Metoidioplasty;
14. Orchiectomy;
15. Penectomy;
16. Phalloplasty;
17. Placement of testicular prosthesis;
18. Rhinoplasty, which also meets restorative function medical criteria;
19. Salpingo-oophorectomy;
20. Scrotoplasty;
21. Urethroplasty;
22. Vaginectomy; and
23. Vaginoplasty.

C. Other than gender reassignment surgeries listed in this policy, surgery and/or additional treatments to change specific appearance characteristics are considered not medically necessary as treatments of gender dysphoria, including, but not limited to the following:

1. Brow lifts;
2. Calf implants;
3. Cheek/malar implants;
4. Chin/nose implants;
5. Collagen injections;
6. Drugs for hair loss or growth;
7. Facial or trunk electrolysis;
8. Facial feminization;
9. Face lift;
10. Forehead lift;
11. Hair transplantation;
12. Jaw shortening;
13. Lip reduction;
14. Liposuction;
15. Mastopexy;
16. Neck tightening;
17. Pectoral implants;
18. Reduction thyroid chondroplasty;
19. Removal of redundant skin;
20. Suction-assisted lipoplasty of the waist;
21. Trachea shave;
22. Voice modification surgery; and

UMP is administered by a third-party vendor under contract with the Washington State Health Care Authority.
UMP Transgender Policy Supplement for Genital Electrolysis

Note: This policy is to be used in conjunction with the UMP specific Transgender Services Clinical Criteria and Policy. Both policies’ criteria must be met for preauthorization to be considered for electrolysis.

Genital electrolysis as required as part of the genital surgery is covered with prior authorization and is limited to the genitals and, if applicable, the graft site, as required for genital surgery. Electrolysis not meeting these guidelines and the guidelines for Surgical Treatments of Gender Dysphoria outlined in the Transgender Services Clinical Criteria and Policy is not covered.