Craig Blackmore: Okay, we have a quorum, so I'm going to call the meeting to order. This is the Washington State Health Technology Clinical Committee. So while Margaret dials us in, there's a couple quick procedural things. The meeting is being recorded. There's a transcript that will be prepared. So, I'll ask all of the committee members, as well as anyone addressing the committee to identify themselves so the transcript knows whose speaking. We also have the ability for people to call in, which is obviously what we're dealing with. So, we call the meeting to order and first item on the agenda is chair remarks, but I'm going to just launch right into the meeting proper and that brings us to HTA program updates, and I'm Craig Blackmore. As I said, I have to introduce myself when I'm speaking to the mic.

Josh Morse: Thank you, Dr. Blackmore. Hello. Thank you Craig. My name is Josh Morse. I'm the HTA program director. Margaret, do we have slides? I'll give a brief presentation on the background of the program and some updates on topics that we're currently looking at.

Margaret Dennis: Josh, I'm not finding the (unintelligible).

Josh Morse: Okay.

Marie Brown: Somebody in here would have one.

Josh Morse: Okay. Thank you, Margaret. So briefly, a brief program overview and some program updates. Today's topics are sleep apnea diagnosis and treatment. Oh, thank you, and bone morphogenic proteins for spinal fusion procedures. The HTA Program started in 2006. It came out of the governor's 5-point plan to improve healthcare in 2005 with an emphasis on evidence-based care to create a more transparent healthcare system and to promote prevention, healthy lifestyles, and healthy choices. The HTA Program is a collaboration among the state agencies that purchase health care, including the Healthcare Authority, Medicaid, Worker's Comp Program, and the Department of

Corrections. Why are we looking at health technologies? Medical technologies are a primary driver of cost in our healthcare system and may account for up to half of all lumped up spending growth and there are, at times, quality gaps in the technologies. The primary purpose of this program is to ensure that the medical treatments and services paid for with state healthcare dollars are safe and proven to work. The program provides resources for the state agencies that purchase healthcare. We developed scientific evidence-based reports on medical devices, procedures, and tests, and we faciliate this committee, the Clinical Committee, to determine which medical devices and procedures are safe, efficacious, and cost-effective. This is a high-level view of our process annually or semiannually the direct of the Healthcare Authority selects technologies for review. We contract with evidence-based practice centers to generate technology assessment reports. We then bring these reports to our clinical committee here for a review and determination in a public meeting and the state agencies then implement the determinations of this committee. Key products of the program include the assessment report and the clinical committee determination. We strive for transparency, and the key focus questions of the program are, "Is it safe? Is it effective? Does it provide value for state purchasing?" The clinical committee decisions must give greatest weight to the most valid and reliable evidence. Objective factors for evidence consideration include the nature and the source of the evidence, the empirical characteristics of the trials for the studies on which the evidence is based, and the consistency of the outcomes within those studies. Additional factors include how recent the data is, how relevant it is to our state population, and the level of bias in the evidence. Currently topics that are under review are for today sleep apnea and bone morphogenetic proteins. In May, the committee will be reviewing upper endoscopy for GERD and GI symptoms, as well as robotic-assisted surgery. In the fall September meeting, we have stereotactic radiosurgery and stereotactic body radiation surgery along with intensity-modulated radiation therapy through radiation treatments. Later, towards the winter, vitamin D testing is currently scheduled. Selected last year but not yet scheduled for the following year include prostate-specific antigen testing, ablation procedures for supraventricular tachycardia, carotid artery stenting, cervical level fusion for degenerative disk disease, cochlear implants, hyperbaric oxygen therapy for wound care and brain injury, and cardiac nuclear imaging. Our program maintains a webpage and

an e-mail distribution list. The primary way we communicate with people who are interested in the program is through the webpage and the e-mail distribution list. If you're not signed up and you're interested in the program, I encourage you to sign up for our e-mail. It's available on our webpage. Thank you.

Craig Blackmore: Thank you, Josh. Craig Blackmore again. Next item on the agenda is previous meeting business. This has 3 parts. First is, we will look to the minutes of the last meeting and then second, we will go individually through the previous two decisions and finalize those decisions or not. So, the first item of official business then on the agenda is regarding the previous minutes from the November meeting. These have been distributed to the committee members and are for the committee members there in your packet, and I will entertain a motion or any, well entertain a motion to approve or any discussion of the minutes.

Seth Schwartz: I move to approve the minutes.

Chris Standaert: Second.

Craig Blackmore: So Dr. Schwartz moved to approve. I have a second from Dr. Standaert. Any discussion of the minutes? All in favor of approval please raise your hands. That's all approved. Next item of business on the agenda is to finalize decisions made at the previous meeting. The first of these is the - which one's first? First is the microprocessor controlled lower limb prosthesis and the decision made at the last meeting has been formalized by committee staff at the request of the committee and that is included in your packets and has been distributed to the committee members. The final decision is detailed on page 8 of that report and I will entertain a motion to approve the final - to approve the decision we made on the prosthesis at the prior meeting or any discussion.

Marie Brown: So move.

Carson Odegard: Second.

Craig Blackmore: I have a motion from Dr. Brown, second from Dr. Odegard, and any discussion? All right. Then we will entertain a vote for a final decision. All in favor of the decision please raise your hands. All approved. Same procedure now on the previous decision about osteochondral allograft/autograft transplantation or the OAT

procedure, as it's sometimes called. Again, the draft decision has been distributed to the committee members and is in the packets, as well as public comments we received, and again I will entertain a motion to approve.

Kevin Walsh: Make a motion to approve.

Craig Blackmore: Thank you.

Marie Brown: Second.

Craig Blackmore: So I have a motion to approve from Dr. Walsh, second from Dr. Brown and any discussion on the topic before we go. All right. So final approval of our draft decision on osteochondral transplantation. All in favor of the decision, please raise your hands. And again, all approved. Okay, that completes our previous business meeting. We move on to the new topic, which is sleep apnea and first item on the agenda is scheduled open public comments. So we have a signup sheet, which has been outside at the entrance. If there's anybody who has not notified us in advance but wishes to address the committee on this topic, there's an opportunity for you to sign in and do that. We've also received some people who have contacted us in advance that wish to address the committee on the topic and have time for that. We're just going to take a second and see if anybody has signed up who didn't tell us in advance. Okay, and we had some people who told us in advance on this topic.

Josh Morse: No.

Craig Blackmore: Okay. So we don't have any scheduled public comments. Nobody has notified us in advance. We have two people who have told us, or who have expressed a desire to address the committee. Okay, so first on the list is Edward Weaver and for all the speakers who address the committee we ask that you identify yourself, tell us who you represent if you're representing more than just yourself, if there's a group or an organization. Tell us if you receive any financial support from an organization, commercial entity related to this topic, and then feel free to go ahead. We limit time for the public speakers to 5 minutes, and do we have somebody whose going to - so Margaret will let you know when you get to 1 minute, is that right, Margaret?

Margaret Dennis: 1 minute, 30 seconds.

Craig Blackmore: 1 minute and 30 seconds, and very good. Dr. Weaver.

Ed Weaver: Thank you, Dr. Blackmore. My name is Ed Weaver. I have no commercial conflicts of interest, and I receive no financial support from anybody related to this commercial industry or otherwise. I am a practicing sleep medicine and sleep surgeon at the University of Washington, Seattle Medical Center. While this experience might pose a potential conflict of interest, I believe it also (unintelligible). My comments are focused on key question #5 - what is the comparative fact of different treatments for obstructive sleep apnea in adults? I will specifically focus on surgical treatment, a niche that I have that is relatively unusual. I was anticipating more public comments from the rest of the questions, so I'd be happy to add insight there, given the opportunity. My qualifications: I am here as an official representative of the American Academy of Otolaryngology Head and Neck Surgery, which has a major interest in sleep apnea. I receive no financial support, but they asked me to represent them as I've been in leadership positions in that organization. Other qualifications include board certification in sleep medicine and otolaryngology, continuous funding for sleep apnea outcomes research by NIH for almost 10 years, co-authorship on 5 surgery studies and 2 guidelines cited in the review, and co-authorship of the Medicare of local coverage determination for surgical treatment of obstructive sleep apnea. I wish to congratulate the Healthcare Authority for taking on this topic. It's of great importance, I believe, to public health in our community and abroad. I congratulate the reviewers for doing really an outstanding review of a very complex topic that expands wide literature, and I congratulate the committee for taking this on and to translate that review into practical real-world practice. So thank you for all your efforts. I know it's a lot. I wish to make 3 points regarding the context of the surgery review, specially, but first I will start with my recommendation. Sorry, one (unintelligible) than I expected so I'm going to limit it to 2 comments. My recommendation is that we follow the Medicare local coverage determination for surgical treatment of sleep apnea. Unfortunately, that is not included in the packet, but it is available online. I do have 1 recommended change - that the definition of sleep apnea worthy of treatment with surgery matches that specified for C-PAP. Specifically, that includes mild sleep apnea with symptoms and not just moderate or severe sleep apnea. In terms of context for surgical treatment, I have it

that there are an array of surgical treatments and in the review, those are all lumped together as a single entity of surgery. In the local coverage determination, those reviews were done specifically for each specific procedure, and that granularity is necessary in order to make a determination, because when lumped together some of the treatments work and some don't work and when lumped together leaves confusion such that the only conclusion is insufficient evidence. When looked at in a more granular fashion, that can be specified more specifically. Thank you for giving me this 5 minutes today to present to the conference. Thank you.

Craig Blackmore: Thank you, Dr. Weaver. Next on the list is Robert Michelson.

Robert Michaelson: My name is Robert Michaelson. I am the vice president of the Washington State Chapter of the American Society for Metabolic and Bariatric Surgery. I'm a bariatric surgeon, and I've been practicing exclusively bariatric surgery for the past 6 years. The position of the Washington State Chapter of the American Society of Metabolic and Bariatric Surgery is that obstructive sleep apnea is known by a large body of scientific and medical evidence to increasing body mass index. We also know that it's a comorbid condition of morbid obesity and treatment with morbid obesity improves sleep apnea. We feel that it is an imperative to diagnose and treat sleep apnea prior to bariatric surgery because of the high risk of this patient population and feel that perioperative mortality associated with untreated sleep apnea exceeds the risks and benefits of the surgery. As an organization, we feel that as an obesity-related comorbidity, bariatric surgery is an effective treatment for obstructive sleep apnea and should be considered one of the surgical options for treatment of this life-threatening condition.

Craig Blackmore: Thank you. So that is all that we've received in terms of requesting the opportunity to address the committee. Is there anyone else who had wished to address the committee who has not had an opportunity to sign up? Okay. We're a little bit ahead of schedule, so we always like to make sure that the time window that we post in advance corresponds to the time when we offer public comments so that nobody misses the opportunity. What I'm going to do is have us continue and then I'm going to continue with the agency utilization and outcomes and then we'll come back and check one more time to see if anybody had expected to talk in that window. One thing I haven't done is to see - is there

anyone on the phone who had wished to address the committee who has not yet had the opportunity?

Marie Brown: I think they're on mute.

Craig Blackmore: Okay. We're going to un-mute you if you're there on the phone. Hang on just a second.

Marie Brown: Would it be possible to hear more from the first presenter, if no one else chooses to speak?

Craig Blackmore: I think given that we are a little bit ahead of schedule that we could entertain that. Let's see what's going on. So, is there anyone on the phone who had wished to address the committee on the topic of sleep apnea? Okay, if you're on the phone, we're going to put you back on mute. If you don't tell us that you wish to address us on this topic. Okay, we will re-mute the phones. Dr. Weaver, we cut you off and we heard a lot about your introduction and we didn't get to hear all your comments, and we have a little extra time. So, if you'd like to give us another 3 minutes, you'd be welcome.

Ed Weaver: Ed Weaver again speaking for the sake of the transcript. The second point that I'd like to make beyond the fact that surgery is more than one single entity that spans all the way from minimally invasive techniques to maximally invasive techniques like craniofacial reconstruction and bariatric surgery. In addition to that, I would like to highlight a key concept reviewed in the literature. When one reviews the literature of surgical treatments, there are some level 1 studies. Paradoxically, the level 1 studies are usually in the treatments that are least effective. The reason that is, is because it's only in those less effective, minimally-invasive treatments where we're even capable of doing level 1 studies. It's very difficult to randomize invasive therapy in those patients who are not willing to accept a flip of a coin to determine whether you do a very invasive procedure, and certainly it's difficult to blind those, either in the placebo control or in the double-blinded fashion. So, inherently, it's almost impossible to do a level 1 study in the more invasive treatments that actually have a greater effect. I congratulate the reviewers for relaxing their criteria, including level 2 studies in the review. However, the emphasis is still on the level 1 studies on the treatments that have less effect and have already been determined not to be used for sleep apnea in many cases, and yet

were included in the lumping of surgery so key in interpreting the literature. Then the third point I was going to make is that surgery is reserved for cases where C-PAP has failed and the patients are otherwise untreated. In this context, surgery benefits far outweigh the risks. Surgery usually does not eliminate the sleep apnea, but it usually improves it. Most of the studies reviewed of the relevant surgeries show consistent clinical benefits superior to no treatment and even comparable to C-PAP or oral appliance therapy. In published higher-level studies, level 2 mostly, surgery consistent improved symptoms, quality of life, car crash risk, instant cardiovascular disease, and mortality - consistently. These are the clinical outcomes that are important to patients. Compared to no treatment, surgery benefits outweigh risks in most studies where both (unintelligible) directly including mortality risk and patient satisfaction. These subtle nuances were not covered in the review. It's too granular for the scope of that review. Lastly, the instructions offer an opportunity to present emerging data, which I would like to do. Compared to no treatment, the most common surgery, uvulopalatopharyngoplasty, is highly cost effective with an incremental cost-effective ratio $8487 per quality adjusted life year with an upper-bound 95% cost effective ratio of $22,000. This analysis used published costs and outcomes stated in applied markup modeling techniques. This pending study has been presented at scientific meetings but has not yet been published and is not included in the evidence review. So, in conclusion, the surgery literature shows that the Medicare covered procedures have important and consistent clinical benefit above and beyond the adverse effects, and emerging data indicates it is highly cost effective. I recommend following the Medicare policy for surgical treatment but also including symptomatic mild sleep apnea, like what's in the policy in 2005 to 2011, and is still in the policy for C-PAP. Thank you.

Craig Blackmore: Thank you. The next item on the agenda is the Agency Utilization and Outcomes.

Steve Hammond: This is Dr. Steve Hammond, Department of Corrections, Chief Medical Officer. Do you have a preference which screen I highlight?

Marie Brown: Would you mind being a little closer to the microphone?

Steve Hammond: Okay, is that better?

Marie Brown: Yes.

Steve Hammond: Okay. How's that?

Marie Brown: Much better.

Steve Hammond: Do you have preference which screen I highlight? Okay. I'm Dr. Steve Hammond, Chief Medical Officer of the Department of Corrections. Sleep apnea, in particular obstructive sleep apnea, or OSA, has been increasingly recognized, diagnosed and treated over the past 20 to 30 years. Prevalence estimates range from 2 to 7% in the adult population, and in adults over 60, prevalence may be as high as 20%. Diagnosis is usually made on the basis of polysomnography. However, diagnostic procedures and criteria are not well standardized, but we do know from the evidence that severe obstructive sleep apnea, as defined by an apnea-hypopnea index of greater than 30, is independently predictive of increased risk of nonfatal cardiovascular morbidity and all cause mortality. The most common treatment for obstructive sleep apnea is continuous positive airway pressure, or CPAP. CPAP has been demonstrated to improve some measures of disordered breathing during sleep. It has not been demonstrated with high quality evidence to decrease major morbidity or mortality and it is not tolerated by a significant proportion of patients. This is the perspective of the agency medical directors group. This topic was selected for review by the HTA Program over 4 years ago. The review by the HTA was delayed somewhat because we knew that a review was underway by the AHRQ and actually our current evidence report is largely based on that review, although with some additional review. In any case, when the topic was initially selected the concern for safety was rated as medium, concern for efficacy was rated as high, and the concern for cost was rated as medium. In light of subsequent review of literature and agency utilization data, the level of concern has been adjusted somewhat. At present, the concern for safety is more in the low to medium range. Concern for efficacy remains high, and concern for cost is medium to high. We do know that 12 to 15 million dollars per year is being spent by the state agencies on diagnosis and treatment of sleep apnea, and we think that given the high prevalence the potential for continued increase in utilization is quite substantial. These are the current coverage policies for the different state agencies. The Public Employees Benefit Package covers diagnosis and treatment of sleep apnea without limitation.

The Medicaid program does cover diagnosis and CPAP treatment of sleep apnea under certain conditions. The most significant ones include that the polysomnography must be performed in an approved sleep center and the sleep apnea must be considered chronic. Labor and Industries covers diagnosis and treatment of sleep apnea related to occupational illness or injury without limitations. The Department of Corrections covers diagnosis and treatment of sleep apnea when deemed medically necessary in individual cases. Diagnosis and treatment of sleep apnea is infrequently authorized in the Department of Corrections. So these are some questions raised by the state agencies. As regards to safety, sleep studies and treatment with CPAP and oral appliances appears safe. Surgical treatment does carry risk related to surgery and safety must be judged in relation to health benefits, which we believe are uncertain given the currently available evidence. As regards effectiveness, improvement in sleep measures and daytime sleepiness has been demonstrated with CPAP treatment. However, the relation of these improvements to major health outcomes, such as reduction in mortality or cardiovascular morbidity, is unproven. Cost is significant and cost effectiveness is indeterminate, because effectiveness for major health outcomes is unproven, and I think that's an important point to keep in mind when you look at the cost effectiveness studies that had been performed. Again, the potential for increased utilization is high given the high prevalence of sleep apnea in the population. So, these are state agency utilization and cost data. We have at the top the Public Employees Benefits Data, in the middle Medicaid, and in the bottom L&I. We see first that there is a rising number of cases that receive diagnosis and treatment related to sleep apnea over the 4 year period of study. I should point out that this, to some degree, reflects new diagnoses. Someone who received services for sleep apnea in 2006 and remained in the plan would be counted again in following years if services were rendered, but we do see rising numbers of patients receiving services for sleep apnea, and we see rising costs. We see similar trends for the Medicaid population, the exception being 2010. We think our utilization data for the year 2010 are artificially low related to the implementation of the new billing system that year, and you'll see the same effect in the data that are presented this afternoon on bone morphogenetic protein. Finally for Labor and Industries, the utilization rates are quite low, which perhaps is not surprising. Sleep apnea would not commonly be considered an occupational illness or injury. We had the opportunity with the Medicaid data

to look at utilization contrasting between adults and children. Actually, this is presented as overall utilization in the top rows and then utilization in adults in the bottom row. So, you need to look at the difference between the 2 to see the utilization for children. But what is of note here is that the difference here is about 4 million dollars out of a total of about 24 million. So, it is of note that approximately 1/6 of the utilization for sleep apnea in Medicaid was for non-adults. This chart shows the trends of the major diagnoses that are associated with sleep apnea diagnosis and treatment services, and there are a couple of interesting points to note here. The diagnoses, although we studied a larger list, they really come down to 3 major diagnoses, one being a very nonspecific diagnosis, the tan color being sleep apnea not otherwise specified, the light blue color being another somewhat nonspecific diagnosis, hypersomnia with sleep apnea not otherwise specified and then the maroon area represents a more specific diagnosis, obstructive sleep apnea, which presumably would be made on the basis of diagnostic testing. So, again, because we are following patients with the sleep apnea diagnosis throughout the 4-year period of study, it's perhaps not surprising that the proportion with more specific diagnoses tends to grow. That was for the Public Employees Benefits data, and we see really a similar pattern for the Medicaid data. This chart breaks down the total cost for sleep apnea related services. We tried to break out costs related to diagnostic studies, as opposed to treatment and then thirdly just for office visits for these patients, and we see that 4 patients with a diagnosis of sleep apnea there is roughly equal costs divided between the costs of the diagnostic studies and treatment, and these are really the lion's share of the costs related to a sleep apnea diagnosis. Also in the lower part of the chart, we see that among sleep apnea treatments, CPAP is far and away the greatest proportion of treatment account for about 95% of the cost of treatment with surgery and appliances being relatively minor. This chart shows the comorbidities that are listed in our database for those patients who received services for sleep apnea, and I think this is noteworthy because sleep apnea is frequently thought to be associated with multiple comorbidities. I checked with Margaret right before the meeting and this includes a diagnosis that was included at anyplace in the database for these patients who did receive sleep apnea studies. So, this should be a fairly accurate representation of the prevalence of comorbidities with the sleep apnea diagnosis and we see, not surprisingly, that type 2 diabetes and hypertension there is a fair degree of comorbidity in the 20 to 30% range, but I think it's also

noteworthy that we're not seeing very high levels of comorbidities along with these sleep apnea diagnoses. This one over here is pulmonary hypertension, a fairly small percentage. Over here, we see congestive heart failure a little under 10%, and coronary artery disease about 10%. Depression is a significant comorbidity in the range of 10-20% prevalent in this group with sleep apnea diagnoses. So, this is the view from the state agencies. After over 20 years of widespread clinical use and voluminous research, the diagnostic evaluation lacked standardization and instrumentation, measurement methodology and diagnostic criteria. We know that severe sleep apnea, as represented by an apnea-hypopnea index greater than 30 is predictive of both higher mortality and higher nonfatal cardiovascular morbidity. We know that CPAP improves certain sleep measures, such as apnea-hypopnea index, arousal indicators, oxygen saturation, and also improves sleepiness scores during the day. We do not have strong evidence or even adequate evidence to demonstrate that treatment of sleep apnea improves quality of life, improves neurocognitive function, improves mortality or cardiovascular morbidity, or other major objective health outcomes. So, the position that the agencies find themselves in is that they are charged with making evidence-based purchasing decisions and also to be good stewards of state resources and the state of the evidence supporting sleep apnea diagnosis and treatment is quite weak. Let me give just one example. We know that practitioners are requesting preoperative screening for sleep apnea and yet the evidence in favor of that, evidence supporting that is of very poor quality and with inconclusive results. We know this is being requested because Labor and Industries needs to preauthorize such requests. We don't have a good idea how much this is being done in the Medicaid and the PEBB programs. So that's one example of the quandary that we find ourselves in, in trying to manage utilization. So, our recommendations are that diagnosis and treatment of sleep apnea be covered with conditions. This was the best we could come up with to address our concerns. We think it is appropriate and justified for the agency to be able to approve the manner of diagnostic evaluation, although there is no direct evidence that treatment of severe sleep apnea with an apnea-hypopnea index greater than 30 does improve health outcomes given the higher morbidity in that group. We think there may be some justification for coverage of treatment in severe sleep apnea and we think that the agencies need to maintain discretion on determining medical necessity in covering

sleep apnea diagnosis and treatment, and I believe that's it. Any questions? Yes, Dr. Odegard?

Carson Odegard: Yes, I have a question about, I mean obviously there's quite a constant trend of utilization and obviously costs over the last 4 years. Did you find any correlation between portable testing, increased utilization of portable testing, portable monitors and the increased utilization?

Steve Hammond: Yeah. Good question and I don't have the answer. Margaret, in your data pull, did you break out, or were you able to distinguish between diagnostic testing done in the sleep lab setting versus other methods?

Margaret Dennis: I didn't really hear the question - the full question.

Steve Hammond: Oh. Do you want me to repeat it?

Margaret Dennis: Yeah.

Steve Hammond: I think the question was, did we distinguish along with the rising trends and utilization, a rising trend in use of sort of, I guess, nonconventional diagnostic approaches, such as home monitoring or other less-intensive approaches than polysomnography.

Margaret Dennis: Right. It looks like the home testing does not appear to be covered, and so we didn't see it in the databases at all. Okay. All we saw was the lab tests.

Carson Odegard: Okay, thank you.

Craig Blackmore: Any other questions for Dr. Hammond?

Marie Brown: I had just one question. I know that the frequency is increasing and the prevalence is increasing, but you talk about prevalence estimates of adults over 60 at 20%? And it looks like from our evidence vendor, they were saying about, less than 20% for older adults or? In the regular population about 2-7%? So would, I'm just questioning your prevalence estimation of 20%.

Steve Hammond: Oh, okay. Well 2 comments. I don't know - those figures were taken straight from the evidence report. I don't know that we have evidence of increasing prevalence, although one might suspect that was the case with the increasing prevalence of

obesity, but I don't know that those were shown in the evidence report, and I just quoted that figure of an estimated prevalence of greater than 20% in adults older than 60 out of the report, and are you our sleep apnea expert?

Akran Khan: I am the sleep apnea expert and I can.

Craig Blackmore: So, if I might interrupt. We're getting a little ahead of ourselves.

Marie Brown: Right, sorry.

Craig Blackmore: I'm going to come back to you if I may, Dr. Kahn, and at this phase we sort of do things arbitrarily, but at this phase, it's really targeted questions about the information that Dr. Hammond can share with us and then we'll get into the more open discussion later. So, if I could defer this issue until a little later. Are there any other questions specifically around the information that the agencies can provide for us regarding the state. Dr. Hammond and his colleagues are here throughout the discussion, so there will be other opportunities.

Chris Standaert: Yeah, I just want to ask, I agree we need to not divert into this topic yet, because I think it's an important question, but when we get into our discussion, there is going to be a question, I know from me, as to the prevalence of the diagnosis of sleep apnea in the population of the people served by the state agencies. So I'm going to ask you about the relevant prevalence in the general population versus what we have in the state so we get some idea as to whether we're matching or not. So I don't know if you have that data. If you can get it, that would be great.

Steve Hammond: I think we can get those data from looking at the absolute numbers that we have and then looking at the number of enrollees. I don't have that at my fingertips, but it looks like Margaret might.

Chris Standaert: So if Margaret can do that. I don't need it now, but when we get that discussion, I'm going to ask that of you. So if it may take some time to do it, that would be great.

Steve Hammond: Okay.

Craig Blackmore: Any other questions at this point? Okay, I'm going to take a step backwards now. We're in the published window for open and

public comments. So, I'm going to ask again if there's anybody who had intended to address the committee and didn't arrive until more recently and wishes to do so? If you could please make yourself known. And then the second thing, Margaret, if I could have you one more time check with the phone and see if there's anyone who has called in.

Ethan Balk: Hi, this is Ethan Balk. I just called in a few seconds ago.

Gerilynn Novajar: Gerilynn Novajar from the HCA. I've been on the line.

Craig Blackmore: And do either of you wish to address the committee, provide public input at this time?

Gerilynn Novajar: I do not.

Ethan Balk: No, I do not.

Craig Blackmore: Okay. Thank you. Okay. So now they're back to muted?

Marie Brown: No.

Craig Blackmore: Okay, is there anybody else on the line? We've now un-muted you if you're there. So, if there's anybody whose there who wishes to address the committee, if you could please identify yourselves at this point. Okay, then we will close the open and public comment portion of the meeting for this topic. Thank you, Margaret. So, the next thing I'd like to do is I would like to introduce our clinical expert. It's the practice of the committee to invite a clinical expert to provide context and clinical and technical details around the technology under study. So, Dr. Kahn has been kind enough to participate with us. If you could please introduce yourself and just tell us your background.

Akram Khan: My name is Akram Khan. I am a board certified sleep physician. I'm a practicing pulmonologist, intensivist, and sleep physician at Oregon Health and Science University, and I'm faculty at the medical school at Oregon Health and Science University. I practice sleep medicine. That's my main area of practice and focus.

Craig Blackmore: Thank you. Do you have any financial conflicts to disclose.

Akram Khan: None that I know of.

Craig Blackmore: Thank you, and what we will ask is that you - we will have questions periodically, and we'll rely on your expertise around the clinical aspects of the treatments and around sort of the context. It's really the role of the evidence vendor to summarize the literature, but there's always a lot of issues that aren't directly addressed in the literature that we need to better understand. So, we thank you for being here. Next item on the agenda is the vendor report.

Ken Gleitsmann: My name is Ken Gleitsmann, and I represent the Center for Evidence Based Policy today talking about sleep apnea, both diagnosis and treatment in adults. Moving through our standard methodology, we'll start with background, go through the key questions, talk about methods, findings, a few statements on policy, and then summarize the report followed by questions that you have, which we'd like to save for that point. As was said earlier, obstructive sleep apnea refers to sleep disordered breathing and is associated with significant morbidity and mortality, as expressed by hypertension, stroke, cardiovascular disease, diabetes, and neuropsych impairment and is prevalent in 2 to 7% of the general adult population. The figure that was discussed earlier is straight out of the ARC Report and prevalence has been shown to be increased with increasing age and is 20% over the age of 60. So, the diagnosis for background, as was said previously, is by a sleep lab study, which is considered the standard, and during this process, this recording of apneic events, key events per hour are made into an index, the apnea-hypopnea index, and the diagnosis of OSA is made when that index is greater than or equal to 15 or greater than 5 with noticeable daytime symptoms. In addition to that diagnosis, there are other ways that have been used to diagnose sleep apnea, and those are portable monitors, various questionnaires, such as the Epworth Sleepiness Scale, STOP questionnaire and others that we'll discuss. There are predictive models that are developed to screen candidates for further formal sleep studies. The monitors that can be used either in a home or sleep lab setting record numbers of channels of information and they are categorized by 4 categories starting with the sleep lab study, which is a type 1 monitor and down to the type 4 monitor, which have fewer, as you go down the types, type 4 has fewer information channels than type 1. For treatment, the standard treatment for sleep apnea is CPAP, continuous positive airway pressure, and there are mandibular and oral devices used, various types of surgery to

include airway and bariatric surgery, and many other interventions from weight loss regimens all the way down to various drug therapies. So, this talk will be organized around key questions, and there are 8. Notice that the first 4 have to do with diagnosis and then 5, 6, and 7 have to do with treatment, and the last question is cost effectiveness. So, our search strategy, which has been discussed earlier, was to look as a sole source of evidence at the ARC review, which was authored by Balk in 2011 that looked at both diagnosis and treatment of OSA in adults and I might say that Dr. Balk is going to be on the line with us as also an expert on this source of information. Our Center for Evidence Based Policy updated the ARC report and searched through the termination date of our search of November 2011, and Washington HTA added a key question, #8, on cost effectiveness, and those studies were looked at for the past 10 years. Our methods in the quality assessment area were that the Balk report methodology for grading the quality of studies uses a scale from A to C, which you'll see in the report having to do with the degree of biased that is found in the reports. Key question 8 was methodology quality assessment used by an instrument developed at our center, and that scale goes from good to poor. So, turning to key question #1 on diagnosis, this slide on the left shows 3 types of portable or home monitors and then presents the evidence, the numbers of studies and types of studies that were used to study these, and then the strength of evidence is presented on the right. In talking through these, the type 2 monitors had a low strength of evidence that they were useful in predicting OSA as compared to the sleep lab studies and the type 3 and 4 monitors had a moderate strength of evidence with a lot more evidence being presented and subjected to meta-analysis that are presented here and the performance measures for those tests all were ranked as a moderate strength of evidence for their ability to predict OSA with the same accuracy as the sleep lab. Across all monitors, the last row shows that there is insufficient evidence to compare one specific type of monitor to the other due to the heterogeneity of the monitor types. So, the next comparison after the monitors is to compare the questionnaire types to the sleep lab studies, and there are various types of questionnaires from the Berlin down to the Epworth Sleepiness Scale. The next comparison is the clinical prediction rules, and there were a number of studies that looked at 10 different kinds of clinical prediction rules, and again the comparator here was the sleep lab study. The last point on the slide was that there was no prediction rule that had been independently validated externally,

and their variable accuracy was noted for predicting sleep apnea. They weren't felt to be accurate for diagnosing severe sleep apnea digging into the comparators, the comparison between questionnaires and the sleep lab studies. This is a complex slide. We won't discuss it in detail, but on the left, you see the questionnaire types and the number of studies that reviewed this particular comparison, and then their ability to predict either the high or low degree of sleep apnea, and the question categories that the questionnaires actually go into are listed on the slides and again with the performance standards on the right. You see that the first 4 of the slides had the binary risk prediction scale used and the last 2 have their own scales that are defined by the test themselves. So, as to question #1, the diagnostic ability of the various comparators, there was moderate strength of evidence that the type 3 and 4 portable monitors are diagnostic of OSA predictability and the low strength of evidence that type 2 portable monitors and the Berlin questionnaire and the clinical prediction rules had that ability to predict an insufficient evidence that comparing the portable monitors to one another or the questionnaires themselves individually had that prediction ability. Moving to question #2, which is about phase testing, there was a single prospective study of commercial drivers that was thought to be low quality and overall found that there was insufficient evidence that phase testing is useful in testing to aid the diagnosis of OSA. Key question 3 is regarding preoperative testing and there was insufficient evidence that there is utility of preoperative testing for the usefulness of preop screening for sleep apnea in that it would be able to improve the postoperative outcomes. Now we're moving to key question 4. These are the important clinical longterm outcomes and there are 2 slides that will show the outcome on the left column followed by the number of studies, the followup period, and then the predictability of AHI, the apnea-hypopnea index, as a predictor of those longterm outcomes and as was previously said, they all cause mortality. AHI is a predictor of that. The higher the AHI level of greater than 30 had an all cause mortality risk increased. The next two rows regarding cardiovascular mortality have mixed results, and the last row in the nonfatal cardiovascular events also have AHI as an outcome predictor in that clinical outcome. The next slide again organized in the same fashion shows stroke as the first outcome and that is not associated with AHI as a predictor but stroke has a low event rate wide confidence intervals suggesting that this study lacks power even though it was for 12 years for followup. Hypertension is showing mixed results in the evidence, and the

type 2 diabetes, that's instant diabetes, is - let me back up for just a minute. The instant diabetes is positively associated to AHI as an outcome predictor and no quality of life outcomes were statistically associated with that predictor. So, in looking at the evidence overall for key question #4, there was a high strength of evidence that greater AHI greater than 30 events an hour is an independent predictor of all cause mortality but a low strength of evidence that this higher AHI was associated with instant diabetes. This is an association that certainly can be confounded by obesity and overall there was insufficient evidence to determine the association of AHI with other longterm clinical outcomes, such as cardiovascular mortality, nonfatal cardiovascular disease, and hypertension. Now, moving to treatment, which starts with key question #5, there are a number of studies that were eligible to address this question; 132 of them were randomized clinical trials, and we're going to organize this by comparators comparison groups. First, we'll talk about a comparison between CPAP and SHAM CPAP, and you can see the variety of trials that were taken as this evidence base in the report. The mean AHI baseline varied from 5 to 30 in this group, and followup varied from a week to 3 months. Most studies had small sample sizes. Very important point, no objective clinical outcomes were reported, and there was significant heterogeneity in the meta-analysis results. These studies did not have compliance rate data reported. So, in looking at a table of comparing CPAP to SHAM CPAP with the outcomes again on the left, the number of studies shown there, and these first 3 outcomes had a significant favor in favor of CPAP in the meta-analysis results. The results of the objective sleepiness and wakefulness test had mixed results and the quality of life differences showed no difference between these 2 comparators. Again, comparing CPAP to SHAM CPAP continuing neurocognitive and psych testing, the results were mixed, as they were for blood pressure. So, in general, there were no significant differences in what were essentially the miscellaneous groups of further outcomes to include clinical outcomes, other sleep measures, and other wakefulness sleepiness tests and a single instrument that's the functional outcomes of sleep questionnaire. Next, comparator will be CPAP versus control. The controls were various and this slide should say controls no treatment or placebo with lactose tablets, or conservative measures, such as weight loss counseling, and that's all to be compared with CPAP. In this group of studies, the baseline AHI ranged from 10 to 65 events per hour. There were some methodological concerns that are

noted there. A significant one was in the crossover trials, there was a lack of washout. So, again, a table to look at this comparator, the CPAP versus control, in the objective clinical outcomes. No significant improvement was noted, just in a single study of congestive heart failure over a 3-month period with CPAP. However, CPAP was favored in the meta-analysis of the next outcomes, AHI, ESS, and other sleep measures. Again, the same table, same comparators looking at different outcomes and with basically mixed results in the outcomes that are listed to include objective sleepiness, quality of life, neuropsych tests, blood pressure, and hemoglobin A1c. No differences were noted between the reports. CPAP was then compared with the various types of CPAP devices, and they range from oral to nasal to auto-titrating to C-flex flexible, bilevel, and humidified or not, and there were no studies that evaluated CPAP of the clinical outcomes of these various devices, and no studies reported consistent improvements with the different or modified CPAP devices in this evidence. Turning to devices of mandibular or oral devices, the trials are summarized here with a variety of qualities. Assessments of the mean baseline in this group is 19-34 so slightly less severe AHI. In the first comparator and also in the MAD to SHAM MAD comparators, there were again no studies that reported objective clinical outcomes and the particulars of just comparison showed that in the outcome of AHI in the 5 studies there was a significant favoring of the MAD group, as was there in the next outcome of the Epworth Sleepiness Scale, and in the various oral devices, those would be devices, such as tongue stabilizing devices. There were no significant differences between those comparators. Now turning to comparators with MAD versus CPAP, there were 10 trials of a variety of qualities, and in this group the baseline AHI was 18-40. Sample sizes, again, were small with a lot of clinical heterogeneity because of so many different devices and again, no studies reported objective clinical outcomes. In looking specifically at these outcomes, CPAP was favored in the AHI outcome. Again, we're trying to look at comparing MAD to CPAP. In comparing the 2, CPAP was also favored in the other sleep measures. No sleep significant difference noted with the Epworth Sleepiness Scale or neurocognitive functioning. We're now going to turn to surgery, comparing surgery to control treatments, there were 6 trials, which compared various surgical interventions through SHAM surgery, conservative therapy, and no treatment. The surgical interventions that have been previously mentioned are soft palate surgeries and a variety of other types of surgery. The mean

baseline in this group of studies was 5-40 events per hour and all of these patients had less than severe OSA. Comparing surgery with CPAP, there were 12 studies with a very large and with comparing the surgical mortalities to CPAP, 2 of which were RCTs, and the mean baseline AHI in this group varied from 5-80. In the randomized control trials, no difference in outcomes were noted when radiofrequency application to soft palate versus CPAP was followed at 2 months and with mandibular advancement surgery against CPAP at 12 months. There were no differences in the outcomes. In comparing surgery with MADs, there was 1 RCT that compared soft palate surgery to the MAD treatment, and it was a B quality. Baseline AHI mean in this study was 19 events per hour and 4 years of followup. Again, clinical outcomes were not evaluated, but in this study there was a 50% reduction in the sleep index at 12 months with the MAD then in those who had soft palate surgery and in that group the difference was 80 versus 60% of that 50% reduction, and it was statistically significant at 12 months and also at 4 years of followup. The last part of the treatment, question #5, is adverse events. The randomized clinical trials were not powered to compare these, and the event data was not collected for control or SHAM. The ARC report used a different methodology to examine larger cohort studies that looked at these for these adverse events, and in CPAP generally 5-15% of patients reported problems with using that, such as claustrophobia, discomfort, and nosebleeds, which all resolved after discontinuing CPAP. For adverse events for MADs, 2-4% of patients complained of pain with jaw pain or TMJ and some dental crown damage was noted occurring in 6% of patients in one study. Across the surgical interventions, there were adverse events that occurred in soft palate and bariatric surgery, primarily reporting perioperative mortalities of 0.2% in each of the categories in the largest cohorts that were reporting, and most of the complications of surgery and adverse events were reported qualitatively so that the overall strength of evidence for this very long key question #5 for treatment is that there's a moderate strength of evidence to say that treating OSA with CPAP is effective versus SHAM or control, that it's superior to MAD regarding improving the sleep measures, but clinical outcomes were not studied, and that MAD is an effective treatment for sleep apnea versus its comparator of SHAM. Low strength of evidence that intensive weight loss programs are an effective way of treating patients who are obese who have sleep apnea. Further, insufficient evidence with patients as to whether CPAP or MAD, which patients will benefit the most from those treatments.

Insufficient evidence that various CPAP devices can be recommended and comparison of various oral devices, other than MAD, shows insufficient evidence. The efficacy of the surgical interventions for sleep apnea is also without sufficient evidence versus the controls of CPAP or MAD, and the merits of the surgery and MAD compared to CPAP also has insufficient evidence to show that. The management modalities, except for the point about the intensive weight loss programs, there is insufficient evidence that they are effective for treating sleep apnea. Key question #6 deals with treatment compliance, and in that the studies that are noted there with various qualities. The important point is the baseline AHI is in the moderate range with the CPAP compliance. Followup varying from 3 months to 4 years and then 1 study that looked at compliance with MAD that was followed for 31 months with poor quality results that were not clearly reported in that study. In compliance with CPAP, most studies found that the higher sleep index AHI baseline predicted greater compliance. Said another way, patients with higher AHI did better using CPAP and also that was noted with higher baseline Epworth Sleep Study scores, and the only quality A study that was in the largest trial, the treatment compliance had been followed for 8 years and that study had an AHI baseline of 50. So in the largest study, interestingly enough, 16% of the patients discontinued CPAP in a year and by 2 years, 32% had discontinued CPAP. So the overall question of treatment compliance has a moderate strength of evidence that compliance is greater in patients with more severe, higher sleep apnea-hypopnea indices in patients using CPAP, or at least prescribed CPAP, and higher ESS scores also associated with higher compliance with CPAP. There was insufficient evidence that MAD compliance has predictors of compliance and all other treatments, as well. Key question #7 discusses interventions, which can perhaps improve compliance with these devices and interventions. There were 18 RCTs and overall, the evidence was of low strength that these interventions do not improve CPAP compliance among overweight patients with more severe CPAP. There was no particular intervention of all of the interventions looked at to include education, telemonitoring, behavioral interventions that was more promising than others. Key question #8 talks about cost and cost effectiveness. There was a systematic review that was of good quality that was eligible for this in 1 individual study. The systematic review developed a de novo economic model that evaluated cost effectiveness. We're talking about treating with CPAP sleep apnea patients. There were 3 clinical endpoints that were evaluated related to qualities

and those were the Epworth Sleep study Scale, blood pressure, and road traffic accidents. In the base case analysis, 50-year-olds with specified cardiovascular risks were analyzed comparing the cost of qualities with CPAP versus MAD versus conservative management. So what were the qualities associated with those? The model was characterizing 4 health states for a lifetime horizon prognosis. The first was sleep apnea, the second sleep apnea with post coronary heart disease, and the third was sleep apnea post stroke, and the fourth health state was death. In this model of this hypothetical correlative men age 50 with cardiovascular factors, CPAP was associated with higher cost and higher qualities compared to the dental devices or conservative management. CPAP showed an incremental cost effectiveness compared with the dental devices of 4000 pounds per quality and the thought was that CPAP would be considered cost effective at a quality threshold of 20,000 per quality. The effective CPAP on the Epworth Sleep Scale has an ICER that is below the threshold of 20,000 pounds for moderate and severe sleep apnea. It was impossible to estimate the differential effect of the baseline severity of sleep apnea as measured by AHI on cardiovascular disease or on road traffic accidents. The single study that looked at cost effectiveness in oral devices was a cost utility analysis of these appliances in treating sleep apnea using a Markov Model, which simulated the costs and benefits of treatment with MADs or CPAP based on their effect of the outcomes we discussed, the 3, and there were many assumptions taking the AHI index as a surrogate for the effectiveness of reducing all adverse events related to sleep apnea, and in this study the ICERs per additional quality were just under $3000 for MADs and just over $13,000 for CPAP. So, our overall key question cost and cost effectiveness summary of evidence is that first, these certainly are not true cost effectiveness studies. They are model studies only. The quality of the models was moderate. The longterm outcomes were unknown, and the surrogate outcomes were extrapolated based on a number of assumptions to get to these longterm outcomes. So, the overall strength of evidence is insufficient or low that CPAP is a cost effective treatment. A brief policy statement here, this is the CMS National Coverage decision that was covering diagnostic procedures for sleep apnea in 2009 and the evidence at the bottom of the slide is what was used for this NCD that basically covers all 4 of the monitor devices to include the lab sleep studies and then there was a 2008 NCD that covers the other side of the coin and that was CPAP therapy for OSA, and that first line should say 12-week trial period. AHI was part of this

coverage decision or with lower AHI with additional symptoms. Also, patient education and a positive diagnosis through either a sleep lab study or a home study and the evidence that supports this coverage decision is shown there. So, overall, to sum the entire report, there was a high strength of evidence again using AHI greater than 30 events per hour that this is an independent predictor of all cause mortality. There is moderate strength of evidence that types 3 and 4 portable monitors are effective in diagnosis of sleep apnea, that treating sleep apnea with CPAP versus Sham or control has a moderate strength of evidence for that, as does the use of CPAP in that it's superior to MAD with regard to the sleep measures, but the clinical outcomes were not studied. MAD is an effective treatment for OSA versus its controls. Compliance is greater in patients with more sever sleep apnea and of those patients that are treated with, or at least prescribed CPAP. Then again moderate strength of evidence for higher sleep scales were also associated with higher compliance with CPAP. Low strength of evidence for the type 2 monitor, the Berlin questionnaire and the clinical prediction roles. Low strength of evidence for a higher AHI being associated with instant diabetes, again, that association being confounded by obesity. Low evidence showing intensive weight loss programs are effective in treating obese patients with sleep apnea, low evidence that particular interventions do not improve CPAP compliance among these overweight patients with more severe sleep apnea and insufficient or low evidence for cost effectiveness for CPAP therapy. Insufficient evidence, and without reading each one of these lines, there was insufficient evidence around comparing the various portable monitors all but the Berlin questionnaire. Phase testing had insufficient evidence, as did preoperative screening of all preop patients with sleep apnea. Insufficient evidence determining an association of AHI with longterm clinical outcomes, and for which patients may benefit most from either CPAP or MAD and for the various types of CPAP trying to choose between the CPAP devices with their methods of delivery and different regimens. Lastly, insufficient evidence to compare various oral devices other than MAD to one another. Insufficient evidence for the surgical interventions, OSA versus CPAP or MAD. Insufficient evidence for the merits of surgery or MAD when compared to CPAP, and the management modalities, excepting intensive weight loss programs, are effective with OSA treatment. The evidence is insufficient for that and lastly, insufficient evidence for the predictors of compliance in all other treatments with MAD. Thank you.

Craig Blackmore: Thank you. So, a lot of information, which is great. Just sort of on a procedural note, we try to divide things up for the committee members on questions directed specifically around the vendor report and then questions to more bring all the information in and guide our deliberation. It's about a quarter of 10. What I'd like to do is just take about 15 minutes and focus specifically on what we've heard in the report, any questions around clarity for the presentation, the information that's here, and then we'll take a little break, and then we'll come back and have the more open deliberation among the panel members. So, specific questions related to the vendor report.

Man: Can you go back to slide 16, please? So for nonfatal cardiovascular disease, there is 1 study with a quality of A, and then on slide 18, there's insufficient evidence to determine an association of AHI with other longterm clinical outcomes, including what's 'nonfatal cardiovascular mortality?' What is nonfatal mortality. I don't understand what nonfatal mortality is.

Steve Hammond: That should be morbidity.

Man: Okay.

Steve Hammond: Sorry.

Man: So, if on slide 16, if the quality of the study that looked at nonfatal cardiovascular disease showed an increased risk with increased AHI, why is it deemed insufficient evidence.

Steve Hammond: The evidence is related to, I think the answer to your question is related to the number of studies and we can pull that study for you and tell you what the size of the N was and so on.

Man: So, it's not necessarily the quality that would determine - it's not quality alone that would give you a grade of insufficient evidence. It would be if the N is powered - if the N is not enough to power that statement or that association?

Steve Hammond: Or if the number of studies to support that. If there were 18 studies versus 1 study, yes.

Craig Blackmore: If I might interject. I mean, if we're showing a difference that's statistically significant then the study had sufficient power. So we

can't say it's an underpowered study if it's shown a difference, which I think this has indicated.

Steve Hammond: Yes, that is true.

Man: So, I guess I'd like to maybe expand on this a little and understand what you're criteria are for high strength of evidence, low strength of evidence, and insufficient evidence.

Steve Hammond: Yeah, thanks for the question. Yes, I'm just going to read this right out of the report in terms of what that definition of insufficient evidence is. It's either unavailable or does not permit a conclusion. There are sparse or no data. In general, when only study has been published, the evidence was considered insufficient unless the study was particularly large, robust, and of good quality. There is a longer description of that, but essentially it says if there were too few studies in a particular evidence base, then the evidence overall is insufficient. I'd be happy to read the whole.

Man: You don't have to do it right now, but could you go back and let us know later what was the N in that study?

Steve Hammond: Yes.

Man: Thank you.

Craig Blackmore: So, I'm hearing then that one study, even if it's a high-quality RCT if it's felt to be too small, you would still say that the evidence was insufficient. Is that correct?

Steve Hammond: That's the way I read this.

Craig Blackmore: And then, can you also tell us how you define low and high strength of evidence in this particular scale.

Steve Hammond: This is grading a body of evidence out of the ARC report. High - there is a high confidence that the evidence reflects the true effect. Further research is very unlikely to change confidence in the estimate effect. I'm just summarizing.

Craig Blackmore: Thank you.

Steve Hammond: Moderate - There is moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate effect and may change the estimate. Low - There is low confidence that the evidence reflects the true effect. Further research is likely to change the confidence in the estimate effect and likely to change the estimate, and again, just to reinforce, insufficient is evidence is either unavailable or does not permit a conclusion. There is sparse or no data. In general, specifically they say when only 1 study has been published, the evidence is considered insufficient unless the study was particularly large, robust, and of good quality and we're pulling that in for you.

Man: So, Craig, I don't know if this is the time to do it. Tell me if it's not, but I want to direct a question to Dr. Balk about their grading system in comparison to this.

Craig Blackmore: Sure.

Man: So, Dr. Balk?

Ethan Balk: Yes, I'm on.

Man: Do you agree that if a study is given a quality level of A and shows an association that it might be deemed insufficient if the end is not adequate?

Ethan Balk: Yes. So the approach that AHRQ and the ETC Program is currently taking is that if there's only 1 study, 1 trial, and it's not essentially a mega trial, and that really hasn't been defined yet. We're in still in discussions about that, but I think for this purpose, say over 1000 people, then it's reasonable to consider the evidence to be insufficient. The idea being that there is some evidence to suggest that the first trial tends to be positive and then subsequent trials that come out often disagree with that and then in the long run it's found not to be. The significant finding goes away. So, it's just the risk that the conclusion is not valid with only 1 trial is the reason for looking at it as insufficient. I'm looking at the report right now. I believe the study you're talking about is (unintelligible) 2005, which is a prospective noncomparative study, and it's looking at predictors of cardiovascular disease. It is a large study. It had 1650 people, but it's a very different analysis. It's not a trial, so it's not comparing

one CPAP to something else. So in that setting with there just being the 1 study, we deemed it insufficient.

Man: Isn't that a situation where you can't randomize? I mean, you're talking about their AHI, so that's an objective finding relative to the patient. You can't randomize them to having a high or a low AHI. So that type of study seems to be the best you're going to be able to do. How do you handle that situation?

Ethan Balk: Well, this question was looking at predictors of outcomes. So, actually randomized trials are not necessarily the best study design for that question, and none of the trials that we looked at address this question. I'm not sure if that's getting at your question, though.

Man: Well the point is, if you have 1600 patients and a prospective observational trial, then that probably is the best type of study you can design for this that would seem to meet your criteria for a mega trial. So, I'm just…

Ethan Balk: Well, it's not a trial. Everybody had - actually I'm not even sure everybody was treated. This was just following 1600 people who were diagnosed with sleep apnea looking at their baseline AHI and seeing how they did over 10 years.

Akran Khan: It was a prospective observational study of 4 groups of 400 people each. I'd be happy to show you the slides from the study and what they did was, they followed a bunch of people who were normal, people who were snorers, people who had mild sleep apnea and severe sleep apnea who were not willing to try any treatment and they followed them for a period of 10 years, and they compared fatal and nonfatal cardiovascular events over a period of time. So, each arm had approximately 400 people.

Man: So, this is listed as a quality A study, and it's nonrandomized, and we have no way to control for volunteer effects, then why is this an A quality study if it's a nonrandomized cohort observational trial?

Ethan Balk: The way that we grade the study is in the context of the study design. So this was a good quality or A-quality observational study that did what we believe to be a high quality analyses for the question at hand, which was not a comparative question. It

was simply whether different features or characteristics of predictors of outcomes.

Man: That's a good question. I guess what I'm trying to clarify along the same lines is the idea that it looks to me like having a very high AHI, I'm calling 30 very high. I don't know how to say very, but having an AHI of greater than 30 is associated with mortality and nonfatal cardiovascular events potentially, what I'm having trouble seeing is - is there evidence that intervening on this and giving people CPAP and lowering their AHI actually results in a relative reduction in those risks consequently. So, does the CPAP treat the findings of the test - meaning it lowers the AHI - which is what's found on the test - without actually affecting any longterm health outcomes? Or do we have any data that by lowering the AHI you actually impact these longterm outcomes? Does that make sense?

Ethan Balk: The short answer, at least from our review, is no. There's no evidence about that. The predictor studies for key question #4 do not address that issue and essentially none of the trials looked at clinical outcomes. So they also do not address that issue. So we found nothing that connected an actual change or drop in AHI with an improvement in clinical outcomes, because the evidence was lacking.

Craig Blackmore: There was a paper that came out in the New England Journal in 2011, which is the first author was, I got to find it here.

Man: Sharma.

Craig Blackmore: Sharma. Was that included in your review or was that subsequent?

Ethan Balk: I'm taking a look now. We have an article from Sharma, 2006, but it must have been published after our review, after our last update. No, it's not included.

Craig Blackmore: Other questions on the content of the vendor report?

Man: One other question. So there's data like this on people with AHI of greater than 30. You don't mention a lot as to whether there's data on people 15-30, 5-30, and I'm not catching that there's no data. Then nobody's looked at people with lower AHIs or that lower AHIs are not associated with these findings.

Steve Hammond: The former.

Man: Nobody's looking at it. So people just picked AHI of 30 as a bar somewhere and that's what they're putting in studies, and we're not getting numbers on the lower AHIs?

Ethan Balk: If I can, I think, clarify. Again, key question #4 about the predictors found that people with high AHI, as defined as greater than 30, had worse outcomes than people with lower AHIs. So people with all the whole range of AHIs were included in those studies. The trials themselves, people definitely had a range of AHIs at baseline, mostly over 15, but actually I don't remember the details and maybe these can be added on. I believe the evidence is, as far as it goes, is applicable across baseline AHI below 30 and above 30.

Man: So these findings didn't become significant until patients had an AHI of greater than 30? So people under that…?

Ethan Balk: I mean there are 2 different questions. One is that the risk of clinical events is higher if you're baseline AHI is high. So, high AHI is a predictor of bad clinical outcomes. That says nothing about whether treating the high AHI is effective or not. The trials that look at treatment, for example that compare CPAP to no treatment, did not do a full report subgroup analyses to say whether people with severe sleep apnea did better with treatment than people with less severe or mild sleep apnea. We don't have information about that.

Man: In terms of health outcomes.

Ethan Balk: In terms of - well really in terms of any outcomes. We don't have adequate information about whether the treatments are more effective in one category of patients versus another category of patients.

Akran Khan: I just wanted to add one thing. One of the things is that the level of evidence that they are using requires a very high bar. So, a lot of the studies have been excluded, which would show improvement because they don't meet the criteria of a randomized trial or a sample size. So, that is why when you do an evidence-based analysis, a lot of clinical data that we use for

clinical practice that ESM has used in defining its guideline does not make it into the EHR queue review or into any evidence-based review because it does not meet the high threshold that we've set up. But that does not mean that…

Ethan Balk: I should - I mean you're right that my comments are in the context of the evidence that we reviewed. Especially given the large volume of trials that have been done, I'm not convinced in this setting that a randomized trial is that high a bar, but nevertheless, for the comparisons, I am only talking about the trial evidence. You're right.

Craig Blackmore: And that's our charge, as a committee, to look at the best available evidence, which certainly if it's available we randomize trials. If I might ask a question on slide 24, and this gets a little bit - maybe this is an editorial comment from me to the vendor a little bit, but I always struggle when we say things like there's no statistically significant improvement found in a trial. Of course, that can mean several things. One is that the trial wasn't big enough to find an improvement if there is one, and the other reason could be that the trial was big enough and there's no difference. So, when I see no statistically significant improvement, I need to know how much power the trial has to interpret that information. So, perhaps at the break if we could specifically the one study that looked at objective clinical outcomes, if we could understand whether or not that study had any power to detect objective clinical outcomes and if it didn't, then we could interpret this as sort of insufficient evidence rather than evidence to the contrary.

Man: I have, I guess, 2 questions. The first is for the vendor, and it's on slide 17. You commented with regard to hypertension there was no - that the evidence was equivocal but the 2 situations we're looking at there both say that it was statistically significantly correlated as a predictor of hypertension, the first with an AHI of it looks like greater than 15 and the first was just greater than 0. So that's a pretty low threshold and still seems to say that there was an association. So, can you comment? Did you speak properly when you said there was not an association or?

Woman: Which slide?

Man: On slide 17 with regard to hypertension.

Steve Hammond: I think you're talking - you were talking about type 2 diabetes?

Man: No, hypertension.

Steve Hammond: Those are saying that AHI is a predictor of hypertension, as a longterm outcome and the question was…

Man: You said the opposite when you gave the report and it's not - in the summary, it's not mentioned at all.

Steve Hammond: Oh, I'm sorry. Yeah, that is incorrect. That is definitely - you're correct.

Man: Okay.

Steve Hammond: That must have been my error.

Man: And then the second question is a more general one, and maybe this is directed a little bit to our clinical expert, but when most of the studies that were shown here had a followup period of about 3 months, and that seems quite short in terms of looking at outcomes like cardiovascular death and CHF and other things like that about having an impact, and so I'm just curious if there's any comment on whether there's been any longer term studies showing real findings with regard to those outcomes.

Akran Khan: So, I think one of the reasons it included the Marin Study is that it was probably one of the most well-followed samples that was followed for over a period of 10 years. They also had a subset of patients who had OSA who were treated with CPAP and their mortality was the same as the non-OSA group. So, we know that treating a group of people with OSA does, at least if you were to go by that, does improve longterm outcomes, and you are very rightly pointing out that the evidence that we are basing everything upon is insufficient because most of the trials are 3 to 6 months. We don't have longterm trial data. We have data from the Wisconsin and Sleep Heart Health Study cohort…

Craig Blackmore: I'm gonna break in here a little bit here. This is confusing, and it's technical, but the role of the clinical expert is to help us with content and technical aspects of the treatment and the interpretation of the evidence is the job of the vendor and of the committee itself. So I don't mean any disrespect, but we have to

have these boundaries for our process and to make it fair, etc. So, I can shift the question over to the vendor, if I may.

Steve Hammond: Would you mind asking that question, again?

Man: Well, it was in regard to the short duration of followup for most of these studies. It seemed to be only 3 months, yet the outcomes of interest seem to be things that would take place over a much longer period of time. So I'm curious, I guess one about whether these studies were inherently set up to not be able to find out a difference and then secondly if there were any studies that we're not seeing here or that are not well called out that might have had longer term followup to look at those important questions.

Steve Hammond: That's a good point. The answer is, when we put a comparator up, and I want to make a comment about 2 things. When the slide is showing a comparison between 2 different groups, we included an AHI range so that you could see the baseline AHI of a particular comparator and the same was true with followups, and you rightly point out some of the followups were weeks and some were 12 years, 14 years. So, I think that is just a matter of what the evidence base is showing, and if you're looking at a longterm clinical outcome, it needs to be longer than 3 months.

Ethan Balk: This is Dr. Balk. If I could just also add that the outcomes that were evaluated were all intermediate outcomes, such as the apnea-hypopnea index and at least an argument can be made that those do change rapidly. So that's why the studies are short-term. Even longer term studies that we found anyway were not reporting the clinical outcomes. It's not just that they were only 12 weeks and that's why we don't have clinical outcomes. That wasn't their questions of interest. They weren't looking for clinical outcomes.

Craig Blackmore: So, if I might ask, Dr. Gleitsmann, regarding the Sharma paper of 2011, which was brought to my attention and not necessarily to the whole committee, could you summarize that for us and try to put that into the context of what we have.

Ken Gleitsmann: I can try.

Craig Blackmore: Thank you.

Ken Gleitsmann: This, again, was outside the purview of this report. It was a subsequent study that I just was looking at. This was a total of 86 patients out of India, and basically it was a crossover study with a 3-month use of CPAP and a 3-month comparator was Sham CPAP. So, a group of patients total in of 87 patients had CPAP used for 3 months, washed out for a month, and then went to the comparator. The results actually were clinical results that were reported to include blood pressure, systolic and diastolic blood pressure, cholesterol, and they broke down a lipid panel. Triglycerides were looked at, and basically these were patients with a metabolic syndrome, and at that period of followup there was a significant improvement of those clinical parameters. There was a statistical improvement of the clinical parameters. Whether you think its clinical or not is a judgment call. For instance, I believe one of the systolic blood pressures improved by 3.2 mm between one group and the other. So, there were many statistical findings that were all significant between those 2 comparator groups in a fairly small sample size in the metabolic syndrome.

Craig Blackmore: Would you be in a position to grade this study by your scale or would you. I mean it hasn't been imbedded through the process. This was subsequent. This is an effort to make sure that we're as up-to-date as we can be, and you can decline, if necessary.

Ken Gleitsmann: I'm not in a position to do that, honestly. I just really had just seen this.

Craig Blackmore: I understand. I do have this paper available if anybody wants to look at it at the break.

Woman: So there were clinical outcomes with an improved blood pressure, improved lipids, and what else?

Akran Khan: (Unintelligible) non-HDL lipoprotein (unintelligible) LDL. I can give you the fourth table if you want.

Woman: Sure. And that was in that 3 months? You got lipid improvement in 3 months?

Akran Khan: Three months (unintelligible) study in which one group received CPAP and the other group received Sham and then they reversed it.

Woman: Thank you.

Craig Blackmore: Okay. I think…

Man: I have one more question, so.

Craig Blackmore: Oh, go ahead.

Man: I'm still - this whole issue of intervening with CPAP and making a difference in cardiovascular mortality. On page 66 of the evidence report you talk about the Marin Study, which, and there are a couple of sentences on here. One is talking about cardiovascular mortality and it says, "In addition, association was not seen in those people with CPAP," and one on nonfatal cardiovascular disease saying only patients with an AHI greater than 30 events per hour who were not treated with CPAP were at significantly increased risk for nonfatal cardiovascular disease. So, this implies to me that there were people in these studies who were treated - in this study - that were treated with CPAP who had different outcomes than those without CPAP and I'm wondering why that isn't considered significant or considered evidence that there is a benefit in terms of these morbidity and mortality measures from the use of CPAP in people with a high AHI. Does that make sense? I'm still trying to get at that same question of are we?

Steve Hammond: Would you - the page number again?

Man: Excuse me, it's page 66 of your report.

Ethan Balk: Is this page 66 of the original report or?

Man: The Washington Health Technology Assessment Final Report.

Man: No, in your - in the AHRQ, it's appendix D, tables 4.3 and 4.4.

Ethan Balk: Yeah, we have that.

Man: Yeah. Those 2 tables, yes.

Ethan Balk: If this is regarding table 4.4 - result of multivariable analyses of AHI as a predictor of cardiovascular events, and we're talking about the Marin Study, I believe. Yeah, again, this study was not about intervention. So, this is, AHI is a predictor. Okay, I'm sorry.

I'm looking at the table for the first time in several months now. So, forgive my confusion. So, it did have people at baseline on CPAP. So, that was included in this table. But again, going back to the point we made earlier, for the purpose of our review, for actual comparisons of CPAP to no CPAP, we only looked at the randomized trials. So, we may have mentioned this observation but we did not consider the CPAP at baseline with no CPAP at baseline in this observational study. We did not include that in our analysis of the comparison of CPAP to no CPAP.

Man: But we don't have randomized trials of CPAP versus no CPAP that are longterm health outcomes, correct?

Ethan Balk: Right.

Man: So the best we have are these sort of cohort studies where you have groups who are on CPAP and groups without CPAP. Some of them have a low AHI or 0 AHI and some of them a high AHI and they looked at various longterm health outcomes of those different groups. So it's a co-…

Ethan Balk: Yeah, I mean the decision could have been made that we would look for observational nonrandomized comparisons of further clinical outcomes, but given the scope of the whole report, that's not the approach we took. So, I can't really add much more to it without going back and reanalyzing the data. We did not look for all the observational studies that compared CPAP to non-CPAP. So, the Marin Study is really just one example, and we don't know if it's the largest or the best, or the only.

Man: You weren't looking for them particularly.

Ethan Balk: We haven't - we didn't look for other comparable studies for that question.

Craig Blackmore: Okay, I think we're at the point where we'll take a break, and then we'll come back, and then we get to sort of the meat of the decision making, which is open discussion among the committee around these areas, and we'll get continued input from all of our sources. It's now 10:18. Why don't we resume at 10:30? Thank you. So we have a quorum, so the meeting is back in session. I want to remind the committee members and anyone else who is speaking again to please use the microphone, as we are preparing a transcript and also to identify yourself so that the

transcriptionist can know who you are. Committee members a copy of the recent Sharma article, which was subsequent to the completion of the technology assessment has been distributed to you during the break. Okay, so this brings us to the decision making portion of the session, but first I want to give the committee members an opportunity to ask any of our presenters if there's additional information or clarification they'd like at this point?

Man: Can I, oh Margaret's not here. My incidents question.

Craig Blackmore: Okay. Any other questions while we wait?

Marie Brown: I have a question of our - is it, from what I can see from the appendices, the majority of evidence that's observational, or case control, supports the fact that there is some benefit to CPAP on a longer-term basis with some clinical outcomes, or at least there is - that's the first question. The second one is…

Man: Well, is that a question or a statement? Is that…

Marie Brown: Am I correct? Yeah, and most of it, however, doesn't stimulate whether it's CPAP or not. So, does that longterm data with prospective negative outcomes of sleep apnea exist as sufficient evidence to really look at sleep apnea outcomes. It seems to me, yeah. So that's the question.

Craig Blackmore: So I'm taking your question as kind of a discussion point for the committee rather than a technical question. I think we've heard what we've heard. Is that sort of your intent?

Marie Brown: Well, I would like some input about whether my observation is correct - that the data is suggesting that just without differentiating whether CPAP is effective. That there are longterm negative consequences of sleep apnea. So that's my question. Is that correct?

Ken Gleitsmann: I sort of heard 2 questions. The last part you were asking about CPAP and longterm clinical outcomes and was there evidence to support using CPAP to improve longterm clinical outcomes. There is not evidence for that. There is evidence for surrogate outcomes that are being used for that and AHI is the big surrogate outcome and intermediate outcome, and there is a very large effect size seen across the studies that CPAP actually improves the

sleep indices to the point that the next point is that makes up a moderate strength of evidence that CPAP is actually effecting the longterm clinical outcomes. The answer, as Dr. Hammond mentioned, is longterm clinical outcomes in CPAP treatment are unknown. The effects are unknown.

Marie Brown: Thank you.

Craig Blackmore: Do you want to ask Margaret?

Man: Yeah. So, my question got back to the incidents of this. What I was curious about is comparing the incidents of obstructive sleep apnea in the general population in these general population studies, which I believe are what 20% over the age of 60 and 2-6% under that and then what the rates are for the state, the frequency of this diagnosis in the population that is covered, if we could compare that.

Margaret Dennis: These numbers were calculated from the published data, the published state data. We just didn't put it together this way. So, I did it in the break.

Man: Hmm. It's fairly low.

Margaret Dennis: So you can see that the PEBB is about 3% and growing to 4-4.5%, and the Medicaid data is around 1% and drops off in 2010.

Ken Gleitsmann: Just my understanding on this, it dropped off. That means that it should have, if there were no new cases it should have been the same, correct?

Man: What if the population changed.

Man: It's a percentage of the population.

Margaret Dennis: Right.

Man: Or it means that the percentage of the population changed, so maybe you had more children, the percentage of younger people in the Medicaid group grew.

Ken Gleitsmann: Well, it would seem to me that the people from the year before are carried on to the next year so that every year there should be at least the same on average.

Man: I take care of Medicaid patients all day long and that's not true. They don't carry on from one year to the next. People lose coverage and get it back like a revolving door.

Ken Gleitsmann: Yeah, that's a good point.

Craig Blackmore: Okay, so I'm going to jump in again and let's all say our names. This is Craig Blackmore speaking. Give it a few times so the transcriptionist can get our discussion down reasonably well.

Margaret Dennis: This is Margaret Dennis. We have a dropoff in Medicaid data we believe because of a reporting anomaly for 2010, and so those patients may be delayed in reporting. We may see the cases in 2011 data.

Craig Blackmore: Does anybody else have any questions?

Marie Brown: This is Marie-Annette Brown. How are decisions made clinically about choices about who obtains a sleep polysomnography versus a level 2 or 3 type of sleep testing?

Craig Blackmore: That sounds like a great question for our clinical expert.

Akran Khan: This is Akran Khan. So, a good way to use outpatient testing would be to examine the patient, find out if they have, and this is how I would use it that if they had excessive daytime sleepiness, they are obese, their significant other was saying that they were having apneic episodes, things that are asked in the questionnaires so it kind of gives me an idea that this person has a high pre-test probability of having sleep apnea. In such a patient, I could send this person for an outpatient level 3 sleep testing and if that level 3 sleep testing shows that say the AHI is more than 15 and then they are sleepy, I would typically put them on an auto-titrating CPAP machine, send them home, and let them tell me if they're improving. Now, if that person comes in and they have an AHI of less than 15, the AHI could still be more than 15. So, they could still have moderate sleep apnea with sleepiness, except that the machine is not that sensitive. But certainly, it would pick up the majority of the machines that level 3 would pick up the majority of patients who have moderately-severe sleep apnea if your (unintelligible) probability was high.

Marie Brown: If you screened a specific population.

Craig Blackmore: Okay. Sorry, Dr. Standaert.

Chris Staendaert: One more question on the evidence. We talked a lot about CPAP. We didn't about surgery, and I guess I have 2 questions. On slides 31 and 32, so 31 is there a conclusion there? It's a comparison of surgery and control trials but doesn't actually say what the outcomes were, as far as I can read. It says what the studies were, but there's no statement on differences in outcome.

Marie Brown: Both RCTs found no difference in outcomes between RFA.

Chris Staendaert: But that's in the next slide. That's 12 other studies.

Marie Brown: Okay.

Chris Standaert: Like the next slide says both RCTs found no difference in outcomes. This one doesn't say.

Marie Brown: Okay.

Chris Standaert: Am I correct in reading that?

Marie Brown: Right.

Chris Standaert: So I just sort of read this and I went, whoa. What happened? My question about the surgery is going to come down to what one of our presenters said that if the surgeries are all lumped together you might get different data than if you - is there one individual procedure that stood out from the others, because have 3 level A studies, then this one, then the second. The next slide, 32, you have 12 studies comparing surgery with CPAP and you have 24,000 patients, which is a lot of patients, and 1 level A study, and was there any data in there that is helpful in terms of a particular surgical procedure? Was it just the effect to sort of put them all together that didn't show any benefit?

Ken Gleitsmann: Is Ethan still on the line?

Craig Blackmore: We'll give the vendor team a minute or 2 to look into that. Are there any other issues that we want to ponder while they can dig a bit?

Carson Odegard: I just have one procedural question, actually a clinical question, this is Carson Odegard, of Dr. Khan. When they do these studies, when they do the lab testing, because in these reports we're seeing a lot of different baseline AHI scores, are these machines - I don't even know how to ask this, but are these machines pretty much the same machines as standardized when you get an AHI score? Are the values going to be confident?

Akran Khan: So each level of machine should be able to be compared with itself. So, for example, a type 1 polysomnography and type home portable polysomnography would be similar because they're measuring sleep, and you actually know when the patient went off to sleep and when they did not. They are using the same technology and nasal cannula in an older thermistor and a chest belt and an abdominal belt to measure that. Now, there's a slight new one, so the level 3 and 4 portable polysomnography home testing machines actually, the patient puts it on to go off to sleep but it may not record the EEG. So, they give you the time that they were in bed rather than the time they were asleep. So, what we get is a respiratory disturbance index rather than an apnea-hypopnea index. So, because you may not be asleep and you may not have apneas when you're not asleep, the respiratory disturbance index might be slightly lower than the apnea-hypopnea index but clinically, it may not make a difference because they are fairly comparable. So, what you might end up happening is that somebody who may have apnea-hypopnea index of 20 in the sleep lab may have a respiratory disturbance index of 12 in the portable device but that would still end up treating the person who tells me that they're sleepy and they have hypertension. So, there's a slight difference.

Carson Odegard: I see. So, they're measuring different things? So those events - those are still considered events?

Akran Khan: Those are still considered events, except that the patients, if the patient was awake for 2 hours between 2:00 am and 4:00 am in the polysomnography, that time will not be counted. Unless the patient documented and we deleted that time out in the home portable test, we would not be able to know that.

Carson Odegard: I see. Okay. Thank you.

Akran Khan: But the actual technologies are the same; the nasal cannula is the same. Some portable devices have an oral thermistor, some do not.

Carson Odegard: Okay. Thank you.

Craig Blackmore: Dr. Gleitsmann, are you ready?

Ken Gleitsmann: Yes, this is Ken Gleitsmann. Yes, thanks for giving us a moment, and thanks for pointing out that the slide was missing a conclusion. The 6 trials were - each study used a different intervention, and that was the reason that there was insufficient evidence to recommend one over the other, and that should have been the conclusion of that slide. In the overall summary for the key question #5, that statement is explicit. The efficacy of surgical intervention of OSA versus control CPAP or MAD is insufficient strength of evidence. Does that answer your question?

Seth Schwartz: Mm-hm. This is Seth Schwartz. Just to follow that one up, so what you said was that if you're comparing one of those treatments to the other, or one of the surgeries to another, there was no evidence, but in those individual studies did they - how did they compare to CPAP or Sham surgery? In other words, was there - what you said is insufficient evidence. What do those individual studies show.

Ken Gleitsmann: So the 6 trials had each surgical intervention that were comparing them to controls and, um, the controls were Sham, conservative, or no treatment, and we can talk about each study result.

Seth Schwartz: See, that's what I'm getting at. I think if we're saying they each looked at a different surgical treatment and then you're lumping them all together. I'm saying do you have the data here to pull them apart. For instance, so if it was useful in one survey but not useful in another, that would be important to know.

Ken Gleitsmann: We have - each study is in the report and the sum total of the studies still comes up with an insufficient strength of evidence. I understand your question.

Seth Schwartz: So, can you refer us to a table or something?

Chris Standaert: I guess 521.2 and 521.3.

Ken Gleitsmann: Yeah, the tables that cover this would be tables 5.20 to 5.20.7. It's 5.20.2 to 5.20.7.

Craig Blackmore: And I'm hearing page 127? Is that in our report? Is that correct?

Chris Standaert: They start at 125.

Craig Blackmore: 125. Thank you.

Chris Standaert: It did mention 1 study, the Weaver Study of 2004 with 20,000 patients using CPAP and 2,000 patients who had surgery, and you mentioned a higher mortality on the CPAP than the surgical group at all time periods through the study. I mean, this is a huge number of people. I mean we're talking about, do we have big studies. That's almost 23,000 patients that did show a reduction in mortality with surgery compared to CPAP. Is that not significant or is that not?

Woman: What are you looking at?

Chris Standaert: I got page 128 - Weaver in 2004.

Akran Khan: If I could just interrupt you and say Dr. Weaver is here and maybe he can answer your question.

Craig Blackmore: I think if we have a specific question, we could certainly refer to Dr. Weaver as one of the authors. Right now, I'm not hearing a specific question.

Chris Standaert: I'm curious on the vendors comments on that study and why they felt that wasn't evidence that surgery showed some benefit, as it seemed to have been associated with a significant reduction in mortality over CPAP.

Craig Blackmore: Yeah. I mean, I think the committee.

Ken Gleitsmann: Yes. So the response would be the same as the discussion we had with Dr. Balk earlier and that is that this is based on one study, which defines insufficient evidence.

Chris Standaert: It's a cohort study. It looks like there's 23,000 people, which is enormous. So…

Craig Blackmore: But it's one study.

Chris Standaert: …but it's one study.

Craig Blackmore: So, Craig Blackmore again. The report we're getting is, you know, a summary of the evidence and the vendor has applied their own rules and levels of evidence around the topic, and that's all very appropriate, but our charge as a committee is obviously a little different. We're going to make a decision no matter what the level of evidence is. So, I think we're not constrained by insufficient evidence. We have to make our own decision based on the best available evidence, and that's our change. So, I don't think we need to necessarily be completely focused on what grading was applied by the vendor, though we clearly do need to focus on understanding the data that drove that grading, itself.

Marie Brown: That's a good summary.

Ken Gleitsmann: Just to clarify one thing about this particular study. The comparator was CPAP versus surgery, and the surgery group was 2,000 and the CPAP group was 20,000, and the net difference was 0.06, which was statistically significant. Whether it was clinically significant is, again, the question.

Craig Blackmore: That's the difference in death.

Chris Standaert: That's mortality, yes.

Craig Blackmore: I'm sure there's a minimally clinically significant…

Ken Gleitsmann: I'm sorry. The study quality was C, as well, which is important in terms of methodology, the study quality was - which is.

Craig Blackmore: Any other questions for any of our team? Okay. I want to press forward. This is a lot of information and it's a big topic. It has a lot of components you've provided us. You've all provided us with a great deal of great information and great summary of what's out there. Now we have to use it and we have to make a decision. I'm going to refer the committee to the decision-making tool, which is in your packet towards the end of the tab. If I didn't identify myself, it's Craig Blackmore again speaking. So, this is the HTCC coverage and reimbursement determination analytic tool, and the first several pages of the tool are kind of in the background and the principles by which we operate, and these are all very familiar to the committee, so I'm not going to go

through it. Determinations are evidence based. Determinations result in health benefit using evidence as the basis for coverage decision, and then there's a section on all of the - or not all but a sampling of other decisions made by other groups, as well as the Medicare National Coverage Determination, and we are required to either conform with Medicare National Coverage Determinations or to make explicit why we believe the evidence dictates otherwise. And then I would like to have the committee turn to page 21 of that analytic tool, which is - it's where we get into our decision making. This is the health technology evidence identification and the first part of the process is for us to go through and delineate the outcomes that are relevant that we believe are important for the topic, and those will be outcomes around safety and around efficacy effectiveness, also special populations and cost. So, staff, Josh and his team, have prepopulated this with relevant outcomes but we need to ensure these are the ones that we are going to be focusing on in our decision making. So, the first is under safety outcomes and mortality is listed. Obviously that's the most important, and then morbidity is listed, and I think we could break down, and I'm giving an opinion now and expanding or defining this list a little better - what are the morbidity outcomes that are of interest to us, and I'm hearing at least 2 categories. One is sort of a quality of life category. People are sleeping better. They are less tired during the day. Maybe it's effecting their lifestyle, their SF-36 or whatever it is as one category of sort of morbidity that we might think is important, and then - I've jumped ahead. I'm supposed to be starting on safety. All right. Forget what I just said. Focusing on safety, morbidity, and mortality.

Josh Morse: Excuse me, Dr. Blackmore, this is Josh Morse. We broke this out by - there's 2 - one for diagnosis and one for (unintelligible).

Craig Blackmore: Oh, I'm sorry.

Chris Standaert: That's what I was about to ask. How would we think about safety in diagnosis? Are we talking about the safety risk association with not diagnosing, or are we talking about…

Josh Morse: It's the safety risk of making the diagnosis.

Chris Standaert: (Unintelligible). So, you're talking about the risks of the actual tests.

Josh Morse: Right.

Chris Standaert: So, we're talking about a cardiac catheterization, we talk about the risk of doing the cath on somebody.

Craig Blackmore: I'm going to break…

Chris Standaert: So that's what I'm trying to straighten out.

Craig Blackmore: I didn't understand how this - but now I do, and I'm going to suggest and please provide me feedback. I'm going to suggest we focus on treatment first. So if we're not covering treatment, I don't think we're going to cover diagnosis. So, let's focus on treatment and if we elect to cover treatment then we can come back to diagnosis, because then it will be important. Does that resonate with the committee? Okay. So, jump to treatments. Now, I'm on page 23 and so first the safety outcomes. So, again, mortality, morbidity, you can see them on the list, surgical complications if we're looking at a surgical treatment obviously, speech voice changes, dental, claustrophobia, discomfort, epistaxis. Are there any other safety outcomes that we think should be added to this list? And then effectiveness. Again, obviously mortality would be the most important, all cause or cardiovascular, and then I'm going to just correct this to say nonfatal cardiac morbidity, and then morbidity could be cardiovascular or noncardiovascular, and stroke and hypertension are listed. Then, there's sort of more quality of life related outcomes. Quality of life, again, this sort of symptomatology. Are there are other efficacy or effectiveness outcomes that we haven't talked about?

Seth Schwartz: This is Seth Schwartz, and we've just had this new paper from New England Journal, which looks at this metabolic syndrome that's not covered there.

Craig Blackmore: Okay. So we'll drop that in here. Any other aspects that we should be considering? Okay, and then, we'll just populate this list and move on to discussion.

Chris Standaert: Neurocognitive function, I guess. Is that not in there?

Craig Blackmore: No, not explicitly.

Man: Under quality of life.

Chris Standaert: Is that considered quality - there are quality of life measures I think of 36, somewhere in there, neurocognitive studies.

Craig Blackmore: I mean, we can tease it out as well.

Man: The one other thing we heard about was car accidents.

Craig Blackmore: Car accidents, yeah. Okay. All right. So, what I need at this point, then, is I need the committee to help me bring this all together and summarize where we are and what we need to discuss and decide on. So, if I could get a volunteer from the committee to sort of give us the brief summary of where we stand, what we think, what they think the data tells us around specifically the question of treatment for sleep apnea. Does anybody want to take a first stab.

Kevin Walsh: This is Kevin Walsh and I need a point of clarification. Are we being asked to lump all treatments in that question, or are going to talk about each treatment.

Craig Blackmore: I'm giving somebody the opportunity to take the first stab, which might be that we separate, as well. Let me back up. We need to separate the treatments, right. There's surgical and nonsurgical.

Kevin Walsh: Okay.

Craig Blackmore: So, somebody give me a starting point on nonsurgical.

Kevin Walsh: This is Kevin Walsh. My interpretation in terms of nonsurgical treatment is that the evidence is insufficient to show benefit from CPAP.

Man: In terms of?

Kevin Walsh: In terms of well - in terms of mortality improvement, in terms of cardiovascular and noncardiovascular morbidity, but I think that the evidence is strongly suggestive that there is longterm benefit. That's how I look at the evidence.

Seth Schwartz: Seth Schwartz. I would say that there is fairly significant evidence that sleep apnea, at least in the severe classification of an AHI of greater than 30 does have evidence that there is longterm negative outcomes related to the sleep apnea itself. We have, I'd

say, moderate evidence that there is - that treatment, specifically CPAP, reduces the intermediate or surrogate outcomes of improving AHI, and a few of the other intermediate markers. I think there is not a lot of longterm evidence, but I think that that's probably partially limited to the nature of the studies where those outcomes weren't looked out due to the nature of the fact that those are longterm outcomes. There probably is some cohort evidence that we've seen suggesting that there is a benefit, but the quality of those studies is not as high based on the nature of the designs.

Craig Blackmore: Can I get some feedback on that from the other committee members? Do they agree with that summary or?

Chris Standaert: I would tend to agree. Also, with nonop we have 2, we have CPAP and we have the mandibular devices, so they're different things, but it's somewhat frustrating because it certainly looks like high AHIs are associated with bad health outcomes, and it seems like you can improve the measures on the test by the use of CPAP. You improve the things you find when you do the sleep test, and what we don't know, and I'm frankly not sure was deliberately looked for, was the effect longterm of doing so - of intervening with CPAP and changing these measures you find on the sleep test, CHI in particular so you can improve immediate sleep measures but does that translate into longterm health? We don't really know from the data we have. There is some data buried in here that is suggestive of it, but it wasn't at a level that they pulled it out to point to as significant. So, it's - that's where I am.

Craig Blackmore: I'm going to look over to this side here. I've been leaning this way too much. Carson, can you reflect for us.

Carson Odegard: I have to agree with Seth. The thing is though that the mandibular devices are just, the MAD devices, and if we compare both of those, CPAP and the oral devices to the control, they both showed benefit, as far as AHI levels. Comparing the 2, CPAP was a little better. I don't know if we want to include any of the other - basically that's what the study pointed out. It didn't even look at any other types of devices or any other types of treatment, but there was definitely a benefit when compared to control groups.

Craig Blackmore: Richard, do you want to give us your perspective?

Richard Phillips: Oh, I don't know if I have anything more to what's been said. My perception of this whole thing is that there's no real evidence that it helps in the long run, real clinical evidence. There's no evidence it reduces mortality, any of the treatments. So, what we're really stuck with from my perspective is dealing with surrogates and whether we believe that the surrogates are adequate to justify the treatment. In other words, knocking down AHI, is that enough to say that it works? I don't know the answer to that, but I think that's the kind of information I'm going to be responding to. I think it's all been said, and I think it's one of these situations where we, we're going to be asked to make a decision largely on a lack of evidence rather than on the strong evidence we'd all love to have.

Marie Brown: I think we're making a decision on intermediate evidence that in itself is fairly substantial. So, I agree with Seth.

Craig Blackmore: Okay. So, I'm hearing a consensus, or at least an opinion that seems to be shared that there is sufficient, perhaps, evidence that the noninterventional therapies decrease the rate of the intermediate outcomes and there are longterm sequelae associated with the disease and we're struggling with whether or not those intermediate outcomes are sufficient for us to endorse this and recognizing that there is some evidence but it's not good randomized clinical trial evidence for longer term outcomes, and I think we might assign different weights to what we think of the strength of that more direct evidence. So, and this specifically would be related to the MAD and the CPAP, the nonoperative interventions. So, how should we approach this, thinking out loud a little bit? I think we need to look at surgical treatment and nonsurgical treatment separately, and I guess we can render a decision on nonoperative treatments or we can come to a decision on nonoperative treatment and then move on to surgical treatment separately. Is that probably the more meaningful approach?

Group: Yes.

Craig Blackmore: Okay.

Chris Standaert: Yeah. If we do it that way, we can just do the safety, efficacy sort of conversation, yeah.

Craig Blackmore: Yeah. So, maybe that's the best way to proceed is to go through safety, particularly related to nonoperative sleep apnea treatment, and we're going to have to, in our nonbinding vote which comes up soon, we're going to have to make a decision about whether it's as safe as nontreatment, etc. We should touch on that here before we get to that nonbonding decision. We talked about effectiveness. What is the group's feeling on the safety of this nonoperative treatment?

Chris Standaert: I didn't see any documented safety concerns with CPAP and mandibular devices.

Seth Schwartz: This is Seth Schwartz. The only safety issue I saw was for the mandibular devices looking at TMJ and jaw pain and tooth problems.

Chris Standaert: 6% of crown problems.

Craig Blackmore: Any other concerns? All right. I think we can move towards decision making on this. The other question, I guess, would be to start to think about the spectrum of disease. We've heard about AHI greater than 30. We've heard about the 5-29 category, etc. Seth, I'm going to kick it back to you. The statement that you started us with, would you apply that to the over 30 event group or would you apply that more generally? What's the starting point for the discussion?

Seth Schwartz: I think that's a hard question. I think that the only outcome that I really saw data on lower than 30 was hypertension, and it looked like there was fairly significant data supporting the association of really any AHI with hypertension over a period of time, and I think there was a study - I can't remember which study it was, but it looked like there was actually sort of a dosed response. So, the worse the AHI the worse the hypertension, but that was really the only area where I saw less than 30 as being used in any of these studies, as a measure. So, I think it's difficult - it's going to be difficult to draw conclusions below 30.

Craig Blackmore: Anybody else want to chip in on the severity of disease, as measured by these tools? Okay.

Chris Standaert: Well, I had asked the question earlier. I mean, it's the same - it's difficult to draw conclusions on the people over 30, too. I had asked the question before and I want to make sure I got the

answer straight in my head. I had asked the question before, do they not mention data on people with an AHI of under 30 very much because there's no data. I mean it didn't show a relationship. I thought the answer I got, but again you can correct me, was that there did seem to be a cutoff in the studies, that the more severe sleep apnea the more severe the AHI the worse things were when you hit that cutoff of 30. So, it wasn't the absence of data below that. Actually, the data below that wasn't as supportive of the association. Is that correct? I just want to make sure I understand the data correctly.

Ken Gleitsmann: This is Ken Gleitsmann. That is correct, however, with the caveat that any elevation in AHI - let me say this correctly - there were findings in addition to AHI that, for instance, the Epworth Sleep Scale studies showed that there was excess sleepiness associated with lower AHIs than 30.

Chris Standaert: Okay. We didn't talk about the consistency of the testing, like the reproducibility of an AHI of 30. Say you do a sleep test and if you're going to make a cutoff, you have to have a reproducible cutoff, don't you?

Ken Gleitsmann: That's diagnosis.

Chris Standaert: That's diagnosis, okay.

Seth Schwartz: There's the one other question that we didn't really talk a lot about, and we're talking about mortality and surrogates of mortality essentially, but the other question here is the quality of life data, and the only slide I'm really seeing about that is, I guess, is slide 22 on the vendor report. I would also have that same question as this sort of says that there is no significant difference, and I'm curious whether that's because there was - that was because they were underpowered or because there just was really no difference.

Craig Blackmore: It looks like slides 21 and 22 from the vendor report, as I'm looking at this, CPAP versus Sham CPAP, ESS, there were significant differences favoring.

Seth Schwartz: Those are markers of sleep and those are sleep quality, but then on the next slide on 22, it's actually talking about the quality of life outcome measures.

Chris Standaert: Go to slide 25.

Ken Gleitsmann: Slide 25 would be most helpful. So there were 7 different quality of life instruments with inconsistent findings across the studies and that's in CPAP versus control.

Chris Standaert: And the other tables are CPAP versus Sham CPAP. What's the difference?

Ken Gleitsmann: So we're looking for quality of life.

Marie Brown: And so the quality of life, 29 quality of life comparisons using 7 measures and how many of those, or about how many of those supported improved quality of life. I understand it's inconsistent, but.

Ken Gleitsmann: We'll look.

Craig Blackmore: Okay. I want to move us then back to the tool, and what I'd like to do is I'd like us to go to the first nonbinding vote, and that'll give us a sense of where we all are and that'll help us focus on where we need to have further discussion. So, I'm on page 25 of the tool and we're going to vote, and again, this is nonbinding. This is to guide our continued discussion and what I would like to do is I would like us to address the question of nonsurgical treatment, and I guess we'll break it out into CPAP and the MAD mandibular device, and we'll treat those separately, and I want us to ask the question is there sufficient evidence that this effects - well we have effectiveness, safety, and cost effectiveness on here. So, the first is gonna be - we'll start with the CPAP. Is CPAP equally, less, more, or of unproven effectiveness compared with basically no treatment? This is effectiveness, and it's specifically CPAP and again nonbinding, and we're not quite all on board yet. You've got to be a little careful, Margaret's behind you. So, I had more than one card in my hand. So that's straightforward, and then same issue CPAP. Is it safe compared to basically no treatment? And then is it cost effective compared to no treatment?

Man: 7 unproven.

Craig Blackmore: All right. Now, I'd like to ask the same questions with respect to the MAD, the mandibular device. So, first dealing with effectiveness, is it compared to, and this will be again compared to no treatment?

Man: 7 more.

Craig Blackmore: Safety.

Man: 7 equivalent.

Craig Blackmore: And cost effectiveness.

Man: 7 unproven.

Craig Blackmore: Okay. That may allow us to lump them going forward. So, further discussion. Are we ready to move on this, or do we want to - we're still, I guess we need to get a little more input from Dr. Gleitsmann, maybe?

Ken Gleitsmann: Yes. So, there were 14 studies, quality of life studies, 6 showed an effect and 8 showed no effect with CPAP. So, the preponderance of the studies, if you want to say it that way, were that CPAP did not have statistically significant effect on the quality of life outcomes.

Seth Schwartz: This is Seth Schwartz. Were there any difference in the qualities or sizes of those trials that?

Ken Gleitsmann: Yes. The 1, 2 - 8 are B studies and the rest are C studies, and the ones that were significant CPAP positive studies on quality of life were 3 B studies, and the rest were C studies.

Man: What tools were used on those again? SF-36 or?

Ken Gleitsmann: There are different tools.

Marie Brown: 7 different tools.

Man: 7 different tools.

Ken Gleitsmann: 7 different tools, yes.

Craig Blackmore: So, based on the nonbinding vote, I'm thinking we're probably, and tell me with nods if I'm headed in the right direction, that we may be looking at a coverage or coverage with conditions. I mean, we're not voting, so I'm not going to hold you to this, but when we get to this point, we make a decision kind of do we try

to define conditions or do we vote first and then try to define conditions, and I'm thinking in this case it might be most productive for us to try to define conditions that we would use if we were to use conditions as the next step. Does that resonate with everyone? Okay. So that's the next step, is we have to look at conditions. One thing we've heard as a possible condition is the severity of score with one of the testing tools. Are there other possible conditions that we think we should be looking at here?

Chris Standaert: We do have a CMS NCD for CPAP for OSA in adults that either we can use or we at least have to look at and decide whether we don't want to use it for some other reason and specify what that reason is.

Craig Blackmore: I think it would be a very good idea to look at it, and now would be the time to do it.

Chris Standaert: It's on page 167 of the Tech Assessment Report.

Craig Blackmore: All right. So, is that national coverage decision around testing or treatment?

Chris Standaert: There's one for diagnosis from 2009.

Man: It's projected.

Chris Standaert: And one from 2008 for treatment for CPAP.

Seth Schwartz: And it's on the screen there.

Chris Standaert: Yes.

Seth Schwartz: And it's on page 3 on the decision board.

Chris Standaert: And we have an LCD for oral appliances and for surgery but not an NCD.

Craig Blackmore: Okay.

Chris Standaert: They use a cutoff of 15 and then say 5-14 if you have either symptoms or hypertension, ischemic heart disease, or stroke. So, they definitely had a lower bar in terms of frequency than 30.

Craig Blackmore: All right, well. Let's take this as a starting point and I'll solicit input from the committee on how or if we might make this better. Who wants to take a stab at it.

Chris Standaert: Just to clarify. This is a - the slide cuts out some of the language. It says a 12-week trial period to determine benefit. So, I assume that means that this is a covered benefit if it's found to be beneficial after 12 weeks. They're not just giving it to people for 12 weeks. They're giving it to them.

Akran Khan: This is Akran Khan. Patients need to wear it for more than 4 hours for 70% of their time to document compliance and document in a face-to-face encounter with a physician that they are benefiting from use. It's only then that they will cover positive airway pressure.

Chris Standaert: So we could cover this with some additional clause of - we've done it before with epidurals and other things where we require some objective evidence of improvement, seeing the people actually using it and benefiting from it. We've done it before.

Craig Blackmore: So, I'm sorry. The Medicare decision requires…

Ken Gleitsmann: They're requiring 2 hours, aren't they? Two hours?

Akran Khan: No, 4 hours, 70% of the time. So, at least 5 nights of the week they need to be wearing it for more than 4 hours, and they have to see a physician in a face-to-face encounter between the 30th and the 90th day and both patient and physician have to agree that the patient is benefiting from the device use. They are more alert. They feel that they have less cognitive decline, and their maybe blood pressure has improved.

Craig Blackmore: So, there is definitely - they have to comply with the tool and there's at least a subjective, if I'm hearing correctly, assessment of the effectiveness of the device for it to continue.

Akran Khan: The CPAP device has a card inside that records how many nights patients are using the machine. It's a very objective criteria of use of durable medical compliance.

Chris Standaert: Actually, what you're describing is under the LCD not the NCD, which is somewhat different for us. The LCD, for continued coverage requires benefit between the 31st and 91st day. This is

under the local coverage, as in the Pacific Northwest, not the NCD. The NCD doesn't have that language it doesn't look like to me, unless it's somewhere else. The language you're describing comes from the local coverage decision.

Akran Khan: But it's the same - it's across the various Medicare sub 9-10 units. It's the same.

Chris Standaert: I understand - but there's a distinction for our committee between an LCD and an NCD. So, if it isn't expressly in the NCD, we don't technically have to address it specifically.

Craig Blackmore: But we can.

Chris Standaert: Yes, we can. I'm just pointing out there's a difference, that's all.

Craig Blackmore: No, that's important. Okay. So we have at least 2 options. We have option 1, which is use the Medicare National Coverage Decision. We have option 2, which is use the National Coverage Decision amended by the language that we find in the local coverage decision around compliance and at least subjective effectiveness. The third option would be to craft our own, maybe based on those or maybe not. Thoughts on those 3 options.

Man: Well, when you look at the Wisconsin Study, the benefit in hypertension and other measures of cardiovascular morbidity, the evidence is not nearly as powerful for an AHI less than 30, as it is for an AHI of 30 and above. So, I'm not comfortable that the NCD is backed up by much here.

Chris Standaert: What's the Wisconsin Study?

Man: It was…

Chris Standaert: Maybe another name or author. I don't know this study.

Man: It was in…

Akran Khan: The author might be Terry Young.

Seth Schwartz: Page 66.

Craig Blackmore: So, what I'm hearing is, as another option, we instead of using 15, we would use 30, and we might or might not include the other

aspects of the National Coverage Decision. The National Coverage Decision says over 15, as I understand it, and somebody tell me if I'm wrong. Over 15 is covered after you go through this trial period, and between 5 and 14 might be covered if you have significant symptoms, which are defined. So, we might - we can do whatever we want. We could use 30 as an absolute cutoff. We could use 30 as a cutoff unless you have significant symptoms and then you might have a lower threshold. Those are options. What are you proposing, Kevin? What do you think the appropriate?

Kevin Walsh: I'm more supportive of 30. Okay. So would you advocate for 30 as an absolute cutoff or 30 unless you have symptoms and then you would have some lower threshold, which is the way the National Coverage Decision is structured?

Marie Brown: The latter.

Chris Standaert: The study you talked about (unintelligible) says an AHI of 15 to 30 and an AHI of greater than 30 are both significantly associated with incident hypertension. So, I'm looking at this page 67 and that second full paragraph. So, it basically says at 15 there is a relationship with hypertension, so they're giving you (unintelligible).

Kevin Walsh: Well, there's a relation but there was a slide that Dr. Khan showed me from that study…

Chris Standaert: Well then can the rest of us have - I'm just trying to find…

Craig Blackmore: I think…

Chris Standaert: It's getting information from elsewhere that everybody isn't seeing. I don't know what to do with that.

Craig Blackmore: But just to clarify that the way the National Coverage Decision is structured is they allow a lower threshold if you have hypertension. So, if we said 30 but you can go to 5 if you have hypertension or whatever, then that's covering the association with hypertension. So, if there were - if we believe that 15 was a better cutoff than 30 for other events, that might be more powerful.

Chris Standaert: I think what this is, though, is a predictor of incident hypertension, which means the occurrence of hypertension, not in patients who have hypertension. They're saying an AHI of 15-30 is significantly associated with incident hypertension. I mean, people who had an AHI of 15-30 are more likely to develop hypertension.

Kevin Walsh: Right, so let me amend my proposal. So absolute cutoff greater than 30 without symptoms. Cutoff 15 and above with additional symptoms.

Man: Including hypertension?

Kevin Walsh: Correct.

Craig Blackmore: Okay, so that's one - that's another option. I need to hear from other committee members on the 15/30, for 15 sort of thresholds where they are on these. Richard, help us out.

Dr. Nebohar: Can I add an additional comment? This is Dr. Nebohar from the HTA. The National Coverage Decision also includes the RDI, or respiratory disturbance index, in its language, not exclusively the apnea-hypopnea index. So, you might want to also include both measures in terms of distinguishing levels.

Chris Standaert: Can somebody explain the RDI for me.

Akran Khan: The apnea-hypopnea index is based on a sleep study where you measure sleep while respiratory disturbance index could be based upon a time period in which a patient goes to bed and the time period that they get up. They may or may not be asleep during that time.

Kevin Walsh: So that's referring to the portable.

Akran Khan: Yes. That is referring to the portable sleep study. A lot of portable devices don't actually measure sleep. You could save costs by not measuring sleep. You could just get a respiratory disturbance index.

Craig Blackmore: And I think we heard from you earlier that the RDI would tend to be lower than the AHI because it's diluted by nonsleep.

Akran Khan: Yes. So if you had an RDI of 30, you really will have much more severe sleep apnea than say if you had an AHI of 30, just because.

Of course, you could be sleeping during that whole time and they could be equivalent.

Craig Blackmore: Okay. So if we use either AHI or RDI at these thresholds it would be a little conservative for the RDI but it would still be covering the same issue.

Akran Khan: Yes.

Marie Brown: What does the data say about the level of RDI? Or just an equivalent? Would you just in the wording say an equivalent.

Craig Blackmore: It needs to be a little lower is what I'm hearing.

Akran Khan: So you are less likely to measure it. So probably, their AHI was higher.

Chris Standaert: So it looks like they're picking a relatively safe cutoff that it starts to hit a threshold at something for AHI and is taking a rough guess at the RDI and they're just trying to pick one number is what it looks like to me.

Kevin Walsh: Well, and the other important thing to consider that I didn't realize is that the AHI is referring to a polysomnogram, and the RDI is referring to the portable devices.

Craig Blackmore: We need to include both, I think, and do we use one level for both?

Kevin Walsh: And I agree with you. I think now I understand better why this NCD looks the way it does.

Chris Standaert: It makes more sense.

Kevin Walsh: And I'm going to withdraw my proposal and agree with this one.

Carson Odegard: This is Dr. Odegard. They're saying AHI of 5 events or less would be equivalent to 14 events of the RDI?

Marie Brown: Yeah?

Akran Khan: No. AHI of 5-14 with additional symptoms of hypertension, sleepiness, other cardiovascular markers or having an AHI or RDI more than 15 irrespective whether you have symptoms or not.

Carson Odegard: But the RDI of 15 - 14 - would be significant.

Akran Khan: It would be significant. So if an AHI was 15, more than likely your RDI is going to be within a 10% range. It might end up being 12 instead of 15.

Chris Standaert: Mm-hm. I see.

Craig Blackmore: So the relatively lower cutoff of 15 versus 30 would capture the RDIs of the patients whose AHIs, theoretically, would be higher and might be above 30, and I hope that sentence did make sense.

Chris Standaert: And if we use AHI, then we can only actually apply that to the formal sleep test, in-hospital deal, so.

Group: Right.

Craig Blackmore: Okay. So, I think we're back to the…

Marie Brown: Using the NCD.

Craig Blackmore: We seem to be coming back to the NCD and the question would be, do we want to include the more specific language from the LCD that we heard from Dr. Khan, and we have it for us, which is that - centers around how we measure compliance in this sort-of initial 3-month trial and involves a subjective, at least, assessment by the patient and provider that there is effectiveness, and we're going to have to get the wording of this pretty carefully.

Chris Standaert: I think we should have something in there about compliance and effectiveness. That seems like a useful measure to throw in, and we could define it or we could leave it up to the agencies to define it, as we've done before, in some manner that they can verify.

Marie Brown: The LCD coverage is adequately focused…

Man: Yeah, I think so to.

Marie Brown: …that we can use it.

Craig Blackmore: Okay. Other comments? Okay.

Chris Standaert: What is - on the LCD it says in those where CPAP is ineffective, a bi-level respiratory assist device can be used. Does that mean a - is that a…?

Craig Blackmore: I'm going to - I need - if we're going to vote on something, we have to have the language on a slide on the board that we can see. So, it's a lot of language and what I'm going to ask Margaret or, I can't remember your name, sorry. If I could - Christine, sorry. If I could get one or both of you to cut and paste from the NCD and the LCD and get it up on the screen and I'm going to give you time to do that, and we're going to move onto something else, and we're going to come back. But I want to have it on the screen before we start voting on it. So, is that - do you understand what we're asking so that?

Woman: Mm-hm.

Craig Blackmore: Okay, thank you. And while they do that, I want us to refocus on the surgical procedures and start to address the same issues and I'm going to pick on Seth, because he knows more about this than I do. So where are we, Seth?

Seth Schwartz: I think this is a little bit difficult for the reasons we've talked about in terms of the multiple different procedures being lumped together and probably having variable effectiveness, but from the data that we have seen looking at surgery, the surgery principally was compared to CPAP and what we saw was that there was essentially no difference, or no significant differences, between surgery and CPAP. So, from an effectiveness standpoint, I'm interpreting this data as that there's no difference between surgery and CPAP or that it's equally effective. I would love to tease out the different surgical procedures more, because I know clinically that there are differences in those procedures, but I don't know that we've seen the data today to be able to do that. So, I guess that's where I see us with regard to surgery.

Craig Blackmore: So I look for input from other members of the committee on the issue of…

Chris Standaert: What do you make of the large - that study of 23,000 people. I mean, you're looking at death, so you need a longterm. It's not going to be an RCT. They're not gonna find it doing level 1.

Seth Schwartz: No, I agree that I think that's very important data, but again, that's - the way I'm kind of interpreting this is we're really looking at surgery versus no surgery as the important measure. So, are we going to cover surgery for sleep apnea, not necessarily are we going to cover surgery over CPAP. So, if we believe that surgery is equivalent to CPAP, then our coverage decision is going to be that surgery is okay. Now we may have different views about when exactly that's offered, and maybe we can talk to our experts a little bit on clinically how that's done, but.

Chris Standaert: So do we have evidence of equivalency to CPAP or just - I mean, that's what I'm trying to know what to make of that study. It sounds like they're saying there's insufficient evidence all the way around, and you're going to have a question of surgery as a primary treatment surgery in people who fail CPAP surgery and all this other stuff.

Seth Schwartz: Yeah. No, I think - well I think those are good questions.

Craig Blackmore: So, if I could ask a question, Dr. Khan. I'm still a little unclear on the clinical context of surgery. Are we using surgery in somebody who CPAP works or are we using surgery in whom CPAP fails, or what's?

Akran Khan: So, there are, again - this has to be put in the clinical context. So, in the sense that if I have a patient who comes in whose 30 years old, has a BMI that maybe is 35, and has an AHI that's 35, if I put him on CPAP, he's going to be on it for 50 years, and should this patient have CPAP or should this patient, as a first line, go for surgery compared to a guy whose 60 years old, has an AHI of 30, has had a cardiac bypass and is on 15 different medications. So, that's a clinical decision that you have to make on an individual basis. Also, you have to kind of decide that if somebody has severe sleep apnea and they're intolerant to CPAP therapy, that the pressure is too much for them, they can't tolerate CPAP, they cannot tolerate bipap, should we just leave them unattended or give them an offering.

Craig Blackmore: So if I'm hearing correctly, it isn't simply that CPAP is used in people who - sorry. It isn't simply that surgery is used in people who fail CPAP or in people who had more severe disease, but there's really a lot of variability in its kind of built around content.

Akran Khan: Exactly, and I had done some research on it. So, just looking at it, what happens is if you lump every patient whose undergoing a UPPP, which is a uvulopalatopharyngoplasty, it looks very ineffective. The cure rate is like 50%. However, if you tease out the right patient then your cure rate becomes very high and becomes in the very high 70 to even 90% range.

Craig Blackmore: And that's something we hear a lot in these meetings, and that makes sense if you can define who those people are in a reliable manner and can show data on those people.

Akran Khan: Exactly, and these surgeons would typically examine them and do an upper airway endoscopy to look at the posterior pharynx, look at the (unintelligible) maneuver. So, they have clinical ways of determining that.

Chris Standaert: We have a problem that our data says there's insufficient evidence for surgery, which is a far cry from saying there is a group in whom it is highly effective that can be reliably identified. That's the dilemma there. Like you say, we hear this all the time.

Craig Blackmore: Okay. So, I'm looking for input.

Richard Phillips: My question - I guess my statement more than a question because I - it seems to me that if we have 15% of people who can't tolerate CPAP, if you don't allow surgery as an alternative, then you're basically saying we can give no treatment for you and we may think that it's equivalent, maybe we don't. We can argue that, because I don't think there's evidence that it's longterm equivalent. But, it seems to me that there are going to be people that either can't tolerate CPAP or people who fail CPAP that may be candidates for surgery. Am I wrong about that? Or is that?

Akran Khan: That's a good assessment.

Richard Philips: I guess I'm trying to understand.

Akran Khan: That's an excellent assessment that irrespective of what we do, if we are going to cover a treatment and there are going to be a bunch of people who'd be intolerant to it for various reasons. Then, their option is either a mandibular advancement device or surgery. Now Oregon Medicare that I deal with often, what they have gone for now is that if you are intolerant to CPAP, they first want you to try a mandibular advancement device before they

will send you for surgery. So, on the flip side, once you have surgery done, it's a permanent kind of a thing. You've cured the problem if you've chosen the right patient.

Group: Or not.

Craig Blackmore: You've certainly done something that's permanent. So, I mean this discussion would speak to limiting surgery to the slight patients who meet some criteria for failure, intolerance, or whatever of CPAP. So, this is sort of getting to if we were going to have conditions on surgery what they might look like.

Seth Schwartz: Unless I misunderstand the situation, it's even more complicated than that, because there are patients for whom surgery would be appropriate rather than CPAP, and there's also patients who fail CPAP who are not candidates whose problem would not be necessarily improved with surgery. So, there's almost like 6 subgroups here when we start breaking them down.

Chris Standaert: But we don't have a way to define either one - we don't have the data to define any one of those groups.

Seth Schwartz: Well, and my proposal would be that we either punt that back to the agency to define those groups or to define how that's going to be covered if we think that there's benefit to surgery.

Craig Blackmore: So, I think, again what I'm getting from this discussion is we're - well let's do it formally. Let's have our nonbinding votes on surgery because that helps us figure out where we are and whether we should be trying to define conditions or not. So, let's get to the right page. So, I'm looking at 25. I'm looking at page 25, apparently.

Richard Phillips: We're comparing surgery to nothing.

Craig Blackmore: So, we are - what are we comparing surgery to? That's a good question. Okay. Let's compare surgery to nothing, and then we'll compare surgery to CPAP, but first I want to compare surgery to nothing. So, is there sufficient evidence that in some or all situations surgery, not otherwise specified, is more, less, equivalent or unproven in effectiveness when compared to no treatment?

Man: 7 more.

Craig Blackmore: And now, safe compared to no treatment.

Man: 4 less, 2 equivalent, and 1 unproven.

Craig Blackmore: And then cost effectiveness.

Man: 7 unproven.

Craig Blackmore: Okay, and then I'm going to take another set of votes now. This is going to be comparing surgery not otherwise specified to nonsurgical treatment, meaning CPAP and/or MAD, but I want to be clear on this. This is - is there sufficient evidence under some or all situations. So, if you believe that there is some situation where surgery is more effective than CPAP, and we've talked about situations like you didn't tolerate CPAP. If you believe that there are some situations where it is more effective then you should vote more. So, if there's any situation where it's more effective, you should vote more. Is that clear? Okay. So.

Man: 7 more.

Craig Blackmore: Safety.

Man: 4 less, 2 unproven, and 1 equivalent.

Craig Blackmore: And then finally cost-effectiveness.

Man: 6 unproven, 1 less.

Craig Blackmore: Okay. Now, these are nonbinding votes, so we're - any continued discussion or continued thoughts you have, you don't have to comply with what you've tentatively decided. Nonetheless, based on this, I think we are in the situation again where we should consider what conditions we would use were we to vote for coverage with conditions. Does that resonate with everyone, again?

Group: Mm-hm.

Craig Blackmore: And then we can have a final decision on how to vote. So, now we're back to where we were, which is the proposal to how we might define conditions for coverage of surgery. Some of the things we might think about are which surgical procedures and we

don't have a lot of data on that. We might consider trials of nonsurgical treatment. We might consider - we've heard about patient characteristics that might be relevant. I'm looking for a straw proposal. A starting point.

Chris Standaert: There's an LCD on this one that reads reasonably and it sort of differentiates between the procedures. We have no data on granularity. This is on page 167 again. We have no data on individual procedures or any degree of granularity to agree or disagree with what they said, I don't think, is my problem. I mean, I don't, we didn't have any discussions, see any data differentiating UPPP from laser-assisted uvulopalatoplasty. I don't know how to - they drew lines somewhere. I don't know how to draw those lines based on what we got. The idea that surgery is covered for people who can't tolerate or fail CPAP and have at least relatively severe sleep apnea doesn't seem unreasonable. I don't know how to draw lines around the procedures, but it looks like they felt some of them weren't reasonable to cover.

Craig Blackmore: So there is no National Coverage Decisions, that's correct, with respect to surgery for sleep apnea? But there's a Local Coverage Decision. Okay, so one proposal is to follow the Local Coverage Decision.

Marie Brown: And it looks like it excludes bariatric surgery.

Craig Blackmore: It excludes…

Seth Schwartz: I don't know that it excludes it. It just doesn't mention it.

Marie Brown: It doesn't mention it.

Chris Standaert: No. Not as a treatment for sleep apnea.

Craig Blackmore: Okay. I'm not sure we need to mention bariatric surgery.

Marie Brown: No, but it looks like there's only 3, well more procedures here - laser assisted.

Seth Schwartz: I think this is based on the fact that some of this stuff has been teased out in the otolaryngology and the sleep literature. We just haven't really seen those studies. Maybe our expert can comment about the relationship between managing sleep apnea and things like the laser-assisted uvuloplasty and radiofrequency,

because I think some of those have been shown not to be terribly effective in terms of treating sleep apnea but are more for snoring. Can our clinical expert comment on that?

Craig Blackmore: I mean, I have - well sure. Go ahead.

Akran Khan: So, American Academy of Sleep Medicine says about say, for example, radiofrequency ablation in the statement should be considered as a treatment with mild-to-moderate sleep apnea who cannot tolerate or are unwilling to adhere to positive airway pressure therapy, or in whom oral appliances have been considered and found ineffective or undesirable. So, this is a statement from American Academy of Sleep Medicine. They did a review last year of all the surgical procedures and, again, a lot of their evidence is based on multiple things, including nonrandomized trials, clinical outcomes, other things.

Craig Blackmore: So, I mean I think sort of from a procedural standpoint, I think we should avoid trying to make judgments about things that we haven't really looked at the evidence on. So, I'm - I think we can justify if there's - we can justify, sort of, deferring to particularly the CMS to some extent, but I think if we really don't have - if we have really not looked at the evidence around something, then we should not make statements about it. So, I mean, I agree. We haven't really seen comparative evidence on the effectiveness of these different surgical treatments. So, we shouldn't go places where we haven't gone through our rigorous process.

Chris Standaert: So, do we address surgery as a whole then, or do we allow the state agencies to sort between different surgical procedures, or do we set up some other process to sort through them if we decide they should be covered in some fashion?

Craig Blackmore: Well, I think those are our choices. We can either say surgery is surgery. We can leave it to the agencies, but they've asked us to do it, so that's not terribly useful. Or we can defer to some other source, which we need to be very careful with, or we also have the option of having a subgroup, convening a subcommittee to look at this specific topic.

Marie Brown: Well in light of cost effectiveness, I'm concerned about that there may be data saying some of these approaches are not effective, and I would prefer we not cover those in here that are not effective, even though we haven't seen the evidence about it.

Richard Phillips: How do we know that?

Chris Standaert: There may be things that are clearly inappropriate, we just haven't had them called out as such.

Group: Right.

Craig Blackmore: So, just to ask our vendor, the constraints of the literature search that you've done - you've given us the randomized clinical trial evidence of these various surgical procedures compared with CPAP and compared with Sham or placebo, but you haven't shared with us more observational trials that we might - I won't say we might - but we haven't shared with us, necessarily, observational trials, nonrandomized clinical trials addressing those questions. Is that a fair statement?

Man: Between surgery and…

Craig Blackmore: I mean surgery and either, well it wouldn't be between Sham, but between surgery and no surgery - uncontrolled.

Chris Standaert: Or between different surgeries.

Group: Between different surgeries.

Seth Schwartz: Yeah, we're not so interested in between Sham or between control but between what the different surgical procedures are.

Richard Phillips: My interpretation of the evidence is that there's really nothing that says one surgery is better than another, short-term, long-term or whatever. The problem I personally see is that I think surgery needs to be a choice because otherwise we're basically saying that people who cannot tolerate CPAP, whatever, they need to have an option available, and surgery seems to be an option that works without really defining what surgery is. So, I personally would feel comfortable not getting down to the particulars of the surgeries, of the types of surgeries, and just saying surgery, and if the state wants to refine that further, that's fine, but you know, I personally don't - I think it - we get to a slippery slope here trying to define evidence - or trying to come up with evidence that I don't think really exists, at least what I read.

Seth Schwartz: Can I comment on that? This is Seth Schwartz again. I think that you're right. I mean, I think all of us, we've already demonstrated we all think that surgery is an effective option here so that probably we're going to come up with something that looks like surgery with conditions, but I think that just from a clinical standpoint there is a very significant difference between surgeries. So some of the surgeries they talk about are simply putting like a little probe into the soft palate, making a little burn spot, and I think versus others, which involve advancing the mandible and the maxilla of your face, or doing a tracheotomy. We're talking about very, very different things here, and I think there is substantial data within the sleep literature that there are big differences, that one may be completely ineffective. We know that a tracheotomy bypasses your entire airway. So, it's going to necessarily cure the problem. So, I think we have to make some comment. I think it could simply be accepting theirs, which we may elect not to do, or we can say we need a subgroup to tease this apart, but I don't think we could simply say that either all surgery should be covered, because I think we're missing something if we do that.

Richard Phillips: Well, I get to respond to your thing. I think that really getting down to the subgroup is the only way I personally would feel comfortable about it, because I have no idea, personally, how to respond other than t say, yes/no, and I would defer to your expertise, because that's where my knowledge base is.

Craig Blackmore: Dr. Hammond.

Steve Hammond: Yeah, this is Steve Hammond. I'd like to raise the possibility that due to the lack of available evidence, one possibility would be a coverage with evidence development decision or that condition.

Chris Standaert: I guess what I'm worried about isn't that there isn't data, but that this really wasn't the question that was part of this search. They really weren't looking to tease out are there surgeries that are more effective than others. They looked at sort of big populations and lumped it all together and I don't know that we got the answer that there isn't any data that tells us that some surgeries are ineffective completely and some may be more effective, and I don't think they give us any data to sort that out. So, the idea of saying coverage evidence development essentially means that we're only covering surgery in the setting of a randomized trial, which puts us in a very difficult position in this condition. So, I

guess we're trying to make - I'm just struggling with the idea that we don't have any data on these different surgeries, and if people are doing a radiofrequency thing versus a tracheostomy these are radically different things and if you just say they're all covered you could be opening up Pandora's box a bit. If you say they're not covered, you're running into a sort of an ethical dilemma there. So, I'm struggling with the lines.

Craig Blackmore: Well, I guess - Dr. Franklin.

Dr. Franklin: I think even a good observational study would be helpful here, since we heard that no matter what the surgery is we don't really know whose most likely to benefit. So, it's a patient choice issue not just which are the surgeries. So, I think one reason to consider a coverage of the evidence development would be how are we ever going to get better information on this if you don't do something like that?

Seth Schwartz: I think that the evidence of that level is there, we just didn't look at it. I think tons of observational studies have already been done. We just haven't looked at it. So, I think the subcommittee option allows us to actually look at the data that's there, which makes a little more sense to me.

Craig Blackmore: Procedurally, it's very concerning if we're trying to make a decision and the evidence exists and we don't have it, and I'm very comfortable that we've seen all the trials and I'm very happy with the comparison of surgery and/or medical nonsurgical treatments versus placebo, but I want to hear again, have we seen the randomized clinical trials and nonrandomized clinical trials where relevant for the intersurgical comparison group?

Steve Hammond: I think your point about the limitations of the ARC Review are pertinent. When they look for surgical intervention, they were only looking at those trials that had treatment arms of greater than 100 patients in each treatment arm. So that excluded a lot of studies, and I think that.

Craig Blackmore: So when we say there's been a trial…

Steve Hammond: There were retrospective and prospective cohorts and clinical trials, but they had to have very large treatment arms.

Craig Blackmore: So, I mean, it's sort of a - when you're looking at does the treatment work versus not work, on some level randomized clinical trials are the most important. You need those to say it works, but when you get into the idea of teasing out the variations of the treatment, you won't necessarily have a randomized clinical trial or multiple randomized clinical trials of every individual variation of treatment compared to placebo, and so we're kind of in the place where now we need to have that piece of data that allows us to compare treatment A to treatment B, because we know that at least something works. Are we - I think we're headed- I think we're headed in that direction. So, I think I'm sharing Seth's discomfort that we may not have the information we need to make a decision.

Akran Khan: Some surgeries, like maxillomandibular advancement like take 4 to 5 hours. They may be very expensive. Also, there may be a cost issue to the insurance program, as far as that is concerned. It costs somewhere up to $25,000 or maybe more.

Craig Blackmore: And this is the problem we encounter all the time. Usually, we encounter it as which subpopulations should we cover and not cover, but in this case it's which subprocedures, if you will, should we cove or not cover, kind of having come to some conclusion about the global picture. So, we as a committee need to decide if we want to - again our choices are cover everything, and we've heard some reservations for that, or to leave it up to the agencies, and that would seem to not be our job, or to defer to another source and one that's been suggested is the local coverage decision, although we don't know the basis for that decision, and the fourth is to convene a subgroup. Convening a subgroup - is the subgroup necessarily going to be able to make a better decision than we are, is the question we have to ask ourselves, but that is also an option. So, I mean that's sort of my take on where this is. Again, I'd like the committee's perspective on where we should - how we should proceed on those four lines.

Seth Schwartz: I have one suggestion, and you can completely shut this down. This is Seth Schwartz again, but we do have Ed Weaver here who is - knows a lot of this information and has been involved in some of those studies teasing apart some of these subcategories. It might be worthwhile to ask him to actually shed a little light on this. It might help guide us to determine where the LCD came from or what a subgroup might look like or what it might look at to make this determination.

Craig Blackmore: I'm happy with input on how the LCD was made. I'm not happy with our making a decision based on.

Seth Schwartz: No, I agree with that. I'm just saying, it might help us to frame the question for the subgroups or - and maybe shed some light on where this LCD comes from.

Craig Blackmore: So, Dr. Weaver, were you involved - you were involved in the Local Coverage Decision, is that correct? Would you share with us how that decision was made, not the details of the decision but the process that you all went through, and you have to come up to the microphone.

Ed Weaver: This is Ed Weaver. Again, I gave a (unintelligible) comment earlier. I was involved around 2004 in helping redraft the surgical treatment for obstructive sleep apnea and OCD through Meridian. I worked closely with Dick Witt. Is Dick Witt here? I worked closely with Dick Witt. We reviewed a fairly extensive amount of literature to try to tease out these issues. It included literature that suggests certain treatments that don't' work and if you look at the LCD, there are certain treatments that are actively excluded for coverage and certain treatments that did work based on in some cases randomized trials but mostly cohort data that hasn't been discussed here today. In addition, partly based on clinical experience and partly based on cohort subgroup data, we tried to specify conditions under which those things would be covered and that's written into the LCD. The LCD, I made a mention in my comments earlier that the LCD has been revised a few times since that first draft. I was involved in those revisions until the last revision, which I think was out of Wisconsin instead out of our region, and I don't know if the regions merged. I just learned this in the last week when I was looking up the materials, but the language is mostly the same. The one key difference is for some reason they changed the AHI criteria to no longer match the other treatments for sleep apnea, but I don't know the basis for that. Otherwise, previously it had matched the CPAP AHI criteria and then there are additional criteria, such as there had to be exam or test findings that would suggest for each given procedure. So for example, you would do a palate surgery if there was evidence of palatal obstruction or tongue surgery if there's evidence of tongue obstruction. Another condition was that the patient had to have been described the risks of the surgery. That's actually specific that that discussion happen to try to include the safety concerns

that go along with surgery. A third condition was that CPAP had been tried and that's not based on the outcomes data because the outcomes data suggests that surgery is as effective as CPAP in a lot of cases. So that was based on just the concern that surgery is permanent and CPAP is not. So, if a patient - if surgery is equivalent to CPAP and the patient succeeds with CPAP, and they made a permanent choice and succeed that's great. If they don't succeed, you can always go to surgery later. You can't reverse surgery back. You can go and have that surgery. That's still an option. It's just the risk - trying to balance the risks and that part was based on not as much data. It was based more on clinical judgment. I think that covers pretty well the gist of it.

Craig Blackmore: Thank you. Okay. I think, tell me if you disagree, committee members. I think we're down to - you know we need to vote. We need to vote to cover with conditions or cover or not cover. I think we kind of have 2 choices for our conditions is what I think I'm hearing. One is to base it on the Local Coverage Decision and the other is to having made that vote have a subcommittee convened to develop the conditions, which we would then subsequently vote on. That's what we basically did around the cardiac stent a couple of years ago. So that would be limited to the surgical. That would be to define the surgical conditions for coverage, including procedure, patient, factors, relationship to CPAP trial, whatever was sort of felt to be supported by the evidence. So I think those are our choices. Am I off on this?

Group: No.

Man: But you would be voting (unintelligible) voting coverage with conditions to (unintelligible).

Craig Blackmore: We would vote for coverage with conditions, and we would vote subsequently on the conditions at a followup meeting, and then we would have to adopt the final decision based on those conditions. So, it induces a delay, but there it is.

Carson Odegard: I have a question, Carson Odegard. Does the LCD cover those patients that don't necessarily - they tolerate CPAP; however, they have some anatomical variant, anomaly, that would require surgery at a younger age, like you were explaining, so they wouldn't be dependent on CPAP later on in life? Because a lot of this discussion is based around intolerance to CPAP and making a

surgery decision along with the risks involved. Does the LCD exclude those patients? I don't have it in front of me, so I don't.

Marie Brown: The LCD just includes - just includes 3 surgeries and laser-assisted LAUP somnoplasty, palatal implants, and submucosal ablation of the tongue base are not covered. That's on the LCD, so it's really 3 procedures.

Chris Standaert: It just says when CPAP whether or not invasive treatments are not tolerated in patients who have been informed of the benefits and risks of surgery. So, it's just purely when they can't tolerate it. Has our LCD changed?

Man: (Unintelligible).

Chris Standaert: To our vendor that's in the tech report. Did the LCD change?

Ken Gleitsmann: The LCD is in the appendix of the report.

Craig Blackmore: Okay. So, Margaret, you've got us prepped for the next step.

Ken Gleitsmann: There's a summary in the text and then the…

Heidi Criss: (Unintelligible).

Craig Blackmore: Let me go back to where we were.

Heidi Criss: (Unintelligible).

Craig Blackmore: So, I'm sorry. For the surgical or the nonsurgical? You guys were going to try to compile the nonsurgical for us.

Heidi Criss: I have the NCD wording that's in your decision statement.

Craig Blackmore: For the nonsurgical?

Heidi Criss: (Unintelligible).

Craig Blackmore: Okay. Can you - putting that aside for the moment, can you put the LCD for surgery up on the screen or not?

Man: Where is it? Does anybody know where it is on the…?

Heidi Criss: What we do is in the appendix of the report is we…

Craig Blackmore: Can you tell us…

Heidi Criss: Sorry.

Craig Blackmore: …who you are?

Akran Khan: 167 is the summary.

Craig Blackmore: If you haven't spoken yet, you need to identify yourself.

Heidi Criss: Yeah. My name is Heidi Criss. I was part of the evidence team that put this report together. So, when we do our reports, typically what we do is a full summary in the appendix of the report. I don't have the page number in front of me. I apologize. And then what we do is kind of a contextual summary that's actually in the text of the report. So, there is an appendix that will have the full NCD and LCD in the back. There should be.

Chris Standaert: Page 397 - 397 gets to be surgery. (Unintelligible).

Craig Blackmore: Okay, so we're looking at 307.31 updated, 324.11 for 40 states including Washington on page 397 of the final report.

Marie Brown: This looks reasonable.

Craig Blackmore: Page 397 of the final report, which I think is an appendix. It's one of the appendix. Appendix N.

Chris Standaert: It's funny they talk about an NDI not an AHI. RDI - sorry - RDI.

Marie Brown: This is fairly detailed.

Group: Yes.

Marie Brown: And looks reasonable, and it still says several of the procedures are not covered, so.

Craig Blackmore: What do you think, Seth?

Seth Schwartz: I think this probably more effectively does what we would do less well in a subcommittee.

Marie Brown: Yes. I agree.

Man: Well put.

Chris Standaert: And the amount of - for what it would to take to come up with something better than this I think would be maybe not worth it.

Marie Brown: Not worth the resources, yes.

Craig Blackmore: Okay. I am going to bring us to vote, then, unless somebody has other comments, and so we are going to do our first binding vote, and this is around surgery for sleep apnea. So this is only related to surgical treatments. This does not relate to our nonsurgical treatments, and members of the committee will have 3 choices. One choice will be to cover unconditionally all surgical procedures in all patients for sleep apnea. The second choice will be to provide no coverage for any surgical procedure in any patient with sleep apnea, and the third choice will be to cover with conditions. The conditions are defined as we said by Local Coverage Decision, which is on page 397 of the report in appendix N, and it is listed as L307.31 updated March 24, 2011, and it is the local coverage decision that includes the State of Washington and that would be what you indicate when you hold up your card that says cover with conditions. All right. So, before we vote the question has been brought up that this is a Medicare policy. Therefore, it's designed for primarily for individuals over the age of 64 and yet there may be children who are considered for the surgery. So, the issue is should we add an additional condition that this is only for adults.

Seth Schwartz: Yes. I think that sleep apnea is a very different disease in children. None of the literature that we've looked at or that we've talked about has included children at all, and I think we need to - that's a good point that we should specify this is in adults.

Craig Blackmore: Okay, and how are we going to define adult.

Chris Standaert: Were the search criteria people over 18 in all these? Yes. So, if the only evidence they looked at was people over 18, that's all we're commenting on.

Craig Blackmore: We have not looked at…

Man: The title of the report is adults.

Craig Blackmore: Okay.

Chris Standaert: So It's 18 and older - or older than 18? 18 and older.

Marie Brown: So taking up children is another…

Craig Blackmore: That's beyond the scope of our discussion.

Marie Brown: It's another review though.

Group: Right.

Craig Blackmore: Okay. So, we're going to specify 18 and older.

Group: Yes.

Craig Blackmore: Okay. So, amend that and if you list your card as cover with conditions that will be Local Coverage Decision as we talked about before on page 397 with the additional criteria in that it is only covered for individuals ages 18 and older. Any other comments or clarifications? Okay, let us vote.

Josh Morse: 7 cover with conditions.

Craig Blackmore: Okay and when we're making a decision we are required to determine if that decision is compliant with Medicare National Coverage Decisions, and in this case there isn't one, so that's not relevant. Second point we are to vote on is the nonsurgical CPAP and the MAD, and now I'll get back to Margaret. So we're not able to, in sort of real time, pull all that onto one piece of paper. Is that right, Margaret? We don't have it in our…

Margaret Dennis: We have the whole National Coverage Decision but it's so big it won't (unintelligible).

Craig Blackmore: We can leave…

Chris Standaert: NCD is really that big?

Craig Blackmore: Okay, so let's go to that in the appendix. That's where it is? It's in the decision tool. That's correct. Okay, so we had, to get back to our discussion, we had discussed using this as a starting point but adding additional criteria, which we found in the Local Coverage Decision, and do all the committee members, they all have this,

because it's in your packet. So, committee members are going to turn to page 3 and then Margaret can you put the Local Coverage Decision on the screen? You don't have it. Do we have the Local Coverage Decision anywhere?

Chris Standaert: Well the National Coverage Decision actually says CPAP is only covered for those beneficiaries diagnosed with OSA who did benefit from CPAP during this 12-week period.

Craig Blackmore: Right, the Local Coverage Decision specified how you determine benefit, that it was usage for 70% and physician and patient agree that it is effective. So, it's just - we talked about adding that language and I was just looking for the language that we could decide to add, so.

Margaret Dennis: I can't copy and paste.

Craig Blackmore: No, I don't need you to copy and paste. I just need you to put it up.

Margaret Dennis: (Unintelligible). I wasn't sure where I'm supposed to be. This is page 397 (unintelligible). Was this the right thing?

Josh Morse: It's 391 of the report has the national.

Margaret Dennis: Aren't we looking for the local.

Man: They don't know what you're asking.

Craig Blackmore: I'm asking for the local. See, the committee members have the national. If you show me the local, we can take the language we need out of the local.

Margaret Dennis: Right, so can you help me find it.

Josh Morse: Yes, it's under Noridian. You have to go to the Noridian site.

Craig Blackmore: Go to the Noridian site.

Akran Khan: If you google CPAP for sleep apnea, you do a search in google, the second link is the Noridian website.

Craig Blackmore: Okay. The team is going to try to find that for us. So, why don't we do this. Why don't get our lunches, since they're here, and that'll give - unless you've got it ready. Why don't we get lunch.

Josh Morse: Yeah, I agree.

Craig Blackmore: So, what we're going to do is we're going to adjourn the meeting for lunch. We're going to reconvene quickly and have a working lunch, because we're behind, and while we're doing that I'm going to ask the team to find that local decision so we can integrate it as we see fit. So, we will re-adjourn in 20 minutes, 12:45.

Craig Blackmore: So we have a quorum. We are back in session and so we've worked our way through the surgical, and we're on to the nonsurgical and we are the point of deciding what conditions might look like, should we elect for coverage with conditions, and the discussion on using the National Coverage Decision and amending it based on some of the language in the Local Coverage Decision. So the National Coverage Decision committee members is in your handout, and have we got the local?

Margaret Dennis: We do, but the 2 back screens are warming up still.

Craig Blackmore: Ah. Okay, so the issues that we were concerned about are how success or compliance is addressed in the Local Coverage Decision. So, if we could try to find the language related to that and documentation. Keep going. Is the Local Coverage Decision for surgery or for CPAP - or both.

Seth Schwartz: It's describing both.

Craig Blackmore: It's both, okay.

Seth Schwartz: It started off with CPAP and went to surgery.

Craig Blackmore: So I might need to help out Margaret, here.

Margaret Dennis: I can make it larger.

Craig Blackmore: Do you know what it is we're looking for, Margaret? Do you understand?

Margaret Dennis: No. I have no idea (unintelligible).

Carrie Lynn: Hey Margaret?

Margaret Dennis: Yeah.

Carrie Lynn: This is Carrie Lynn. I e-mailed you the WAC, which is a combination of the NCD and LCD for CPAP. So most of the language is based on the NCD, except for the 3-month review, which is more based on the LCD.

Craig Blackmore: Okay. So we're going to have to help Margaret. Okay, so page 394 is the LCD for the MAD - Mandibular Advancement Oral Device, and it looks like page 395 is the LCD for the CPAP device. So on the bottom of page 395, it says, "Continued coverage of a PAP device beyond the first 3 months of therapy requires that no sooner than the 31st day but no later than the 91st day after initiating therapy, the treating physician must conduct a clinical reevaluation and document that the beneficiary is benefiting from PAP therapy. Documentation of clinical benefit is demonstrated by face-to-face clinical reevaluation by the treating physician with documentation that symptoms of obstructive sleep apnea are improved and 2 objective evidence of adherence to use of the PAP device reviewed by the treating physician, adherence to therapy is defined as PAP greater than or equal to 4 hours per night on 70% of nights during a consecutive 30-day period at any time during the first 3 months of initial usage. If the above criteria are not met, continued coverage with a PAP device and related accessories will be denied as not reasonable and necessary, and then there's some more detail. So, we're not going to get it on the screen, but you have it in front of you. So, I would suggest we, as our draft conditions include the National Coverage Decision excluding the component in there for coverage with evidence determination, which is out of our scope, and add the section on continued coverage is dependent on conditions or the specifications that were in the Local Decision, which I just read.

Man: Mm-hm. I agree.

Craig Blackmore: Okay?

Group: Okay, yes.

Craig Blackmore: So, that brings us to a vote and the vote will be - this is the final voting, binding vote, on coverage of either CPAP or MAD and the - your choice. I'm going to back up. I didn't read the MAD portion

of this. I only read the CPAP portion of this, and there is an MAD portion, is there not? Is it the same? Here it is. Okay, so there's no criteria for compliance with the device.

Chris Standaert: I think once you get it, we can't take it back.

Marie Brown: Right, and it's much harder to be accountable in terms of adherence.

Craig Blackmore: Okay. All right, so the vote is now to cover unconditionally - you have 3 choices. One is to cover unconditionally the use of the CPAP and/or MAD device. The second is that we will not cover either device under any circumstances, and the third is that we will cover under conditions, and the conditions are age greater than or equal to 18 years, as well as the specifications laid out by the Medicare National Coverage Decision excluding the section on coverage for evidence determination and adding the section from the Local Coverage Decision specifically for CPAP only describing the continued use beyond a 3-month trial, and that's in your handout or in the appendix, and you've all seen it and heard it read. Is that clear before we vote? Does everybody understand what I'm asking? Okay. Then we will have the vote, as stated.

Man: 7 cover with conditions.

Craig Blackmore: Okay, so it's going to be a challenge for staff. We are charging you with producing a formal text of that decision that we will have a final vote on at the next meeting.

Man: So, is it okay if I state back to you what I (unintelligible)? It's the NCD less the coverage with evidence development portion or persons of age 18 or older plus the Local Coverage Decision for CPAP with the mentioning of the continued use of beyond 3-month factor.

Craig Blackmore: Yes.

Man: Okay.

Craig Blackmore: Thank you. Okay. We will put that aside and we will move on, and the next aspect of this is the diagnosis portion. So, if we go to our decision-making tool - so the determinate analytic tool starts on page 21. Okay, so the first thing we do again is we delineate the outcomes that we're looking at, and our staff, Josh and his

team, have prepopulated this, safety outcomes related to and we're going to consider both forms of testing or all 4 levels of testing, I guess. The safety outcomes are morbidity and mortality. Are there any other safety outcomes that we should be discussing? Okay. Then there are the effectiveness, efficacy outcomes, and this is a diagnostic test. So, primary outcomes are accuracy or sensitivity and specificity. Then, they have validity listed here. I might look at reliability, which is right here, repeatability and reproducibility, sorry. Then, clinical utility, which supposedly referred to usefulness of the test in clinical decision making. Any other outcomes where we should be paying attention to that aren't delineated? Okay. And then, special populations. Is this test more or less effective in these populations, and these are the populations we might think about. Are there other populations that we haven't teased out? I'm not sure there are, and then finally cost. We looked at total healthcare costs from the standpoint of the payer or payers and we looked at what evidence there is pertaining to cost effectiveness. Are there any other cost outcomes we haven't looked at? Okay. So, it sounds like we're in a pretty good agreement on this. So, I think the next step is to get a sense of where we are and to have the committee put into words a starting point for our discussion of this and sort of not necessarily a proposal but sort of a summary of where they think we stand, and I see Chris reaching for his microphone, so I'll ask him to take an initial stab at where we are.

Chris Standaert: So, in some of these things, the issue of diagnosis is there. It seems like the diagnosis is made by the test. So, I don't know what the gold standard is. So, if you're looking at a formal sleep test, I'm not sure what the gold standard is. Most of the things in the report seem to look at questionnaires versus PSG or different types of portable devices compared to PSG, but PSG is assumed to be the gold standard, and that's how the diagnosis was made. So, I don't know if there's another gold standard.

Akran Khan: That would be the gold standard…

Craig Blackmore: Microphone. Thank you.

Akran Khan: This is Dr. Khan. PSG would be the gold standard. The disadvantage of the PSG is the costs and the stay in a facility overnight versus a level 3 or a level 4 device that is easier for the patient and costs one-fourth as much.

Chris Standaert: So, if we're talking about PSG as a diagnostic test, we don't have a comparator, because that's how the - this is assumed to be the gold standard.

Akran Khan: Exactly.

Chris Standaert: So, all these issues of sensitivity specificity we don't have any of that data for PSG because that's how - the diagnosis is defined by the test.

Akran Khan: Right.

Chris Standaert: There's no other way to make a diagnosis. So that one's sort of problematic.

Craig Blackmore: I mean it's relevant for the do we approve - we're going to have to approve some form of testing if we're going to approve the therapy, right?

Man: Right.

Craig Blackmore: And so, the level 1 testing is the gold standard so you pay more for it, but it's the gold standard and then these other metrics would be the relation of the other levels of testing to the gold standard.

Chris Standaert: Plus the issue - and then whom you perform the test - I assume, because they have these other studies looking at the predictability of different questionnaires to the results of PSG, I assume, and screening tools, which sounds like some way to find the people who should undergo PSG. I didn't see that really clearly either. Like, who are the patients who should get this? I didn't see anybody had defined the population of who should be receiving the PSG. I mean, they're not going to do it to everybody. I didn't see that anywhere.

Akran Khan: Okay, so.

Chris Standaert: Can we start with the vendor?

Ken Gleitsmann: That is the limitation of comparing the PSG to others is that it's being compared as if it were the gold standard. There's a section in the text that discusses all the reasons why that's not exactly the

case, but it is the clinical standard, and it's the consensus start, and it's the thing that's being used.

Chris Standaert: So there's no study of a PSG in essentially a normal - you take a normal population of people who are completely asymptomatic and do a PSG on them and see how many of them - I mean. Does such a study exist?

Akran Khan: That's how they made the determination of apnea-hypopnea index by - essentially they took a bunch of people in Wisconsin, middle-aged government employees, and did PSGs on everyone, and they found out that 2% of women and 4% of men had an apnea-hypopnea index more than 5 and around about - and were sleepy during their time. That's how the determination was made.

Chris Standaert: So it's a bit of a self (unintelligible) sort of thing isn't it?

Man: Toward the OSA.

Chris Standaert: Yeah.

Man: (Unintelligible) from that sleep study.

Craig Blackmore: But if - if I understand - if we're to try to define a group on whom we might pay for this test, the literature review did include trying to find any validated clinical prediction rules that we might use to say these are criteria for use of the test.

Steve Hammond: Exactly, and to try to say are there ways of - one of the questions was phase testing - was there anything, short of the sleep study, to avoid having to have everyone have a sleep study, and there was no evidence sufficient for that.

Marie Brown: In slide 10, it talks about the comparison of portable devices, and the type 4 devices, there were 70 studies. I see the specificity is very variable. The sensitivity is 75-82%.

Akran Khan: So, AASM - American Academy of Sleep Medicine - has some recommendations for, as a clinical guideline, that should be considered in patients who are obese, who are retrogonadic, who are complaining of daytime sleepiness, snoring, and hypertension. Those are the people who are at risk of sleep apnea, and again, this is a general guideline. I don't think it is based on any kind of…

Man: So I would ask…how…

Craig Blackmore: So we've seen the evidence, and there are no validated clinical prediction rules. So, there's no evidence to drive the determination of when we might use this test. I'm sure there's all sorts of people of rendered opinions, and that might be all we have, but at least from the evidence standpoint, we don't.

Seth Schwartz: I mean, can we do something like it's applicable for patients with signs or symptoms of, suspicious for obstructive sleep apnea or something like that. I don't know how - I guess the point is I don't know how specifically we're going to be able to define this, and it seems like there is a whole host of different criteria that the sleep medicine doctors and other doctors are using to assess who to offer this test to, and we have no data that's any better than that.

Craig Blackmore: Yeah, I agree.

Marie Brown: Given the cost of the tests, it's worth investing in looking at whatever kind of data this is.

Chris Standaert: Right, $1000 is an expensive screening if this is what it's being used as, essentially.

Seth Schwartz: But I think that's why the portable devices were developed, in an attempt to deal with that.

Craig Blackmore: Richard.

Richard Phillips: But isn't it true, we do not have any information on the costs of the tests. We have - the cost effective studies basically took the cost of the tests and the variable costs of different treatment options. We don't have the individual tests. Therefore, we can't make a statement about the tests.

Marie Brown: Well, I thought we had Washington data on how much the testing costs versus how much the treatment cost, and the testing - it's almost equivalent to the treatment.

Richard Phillips: We don't - I didn't know that we had that by different procedure. Now, am I wrong?

Chris Standaert: Well we had, I thought we had data on the cost per patient for the diagnosis and treatment, which was a little over $1000, if I remember my math.

Richard Phillips: Yeah, and treatment.

Chris Standaert: And we had a cost of a CPAP machine, which was relatively low in there, about $100 for the machine, and so my assumption in there was that there - so I'm just doing the math here.

Craig Blackmore: So let's ask - let's ask the agency directors. How much do we pay for testing for CPAP?

Richard Phillips: Margaret, if you can help bring that up. I'm looking for it myself.

Craig Blackmore: Or for that matter, how much does Medicare pay? I mean, is there a standard national reimbursement?

Akran Khan: Roughly $800 for diagnosis, around $120-$140 for interpretation, so roughly $900 by Medicare. And then, for a portable level 3 device, we'll get $145 for the whole thing - diagnosis and interpretation. An average CPAP machine costs, the cheapest one costs $450, but an average cost to the patient is roughly at least $800-1200 by the time they go through the BME because they charge them for masks and supplies and other things.

Richard Phillips: I think that's better information than what we can determine from our administrative data.

Craig Blackmore: Okay. So where's that get us?

Marie Brown: Well, there's built-in incentive to do the most expensive tests.

Craig Blackmore: Well, it depends on your perspective.

Akran Khan: For the sleep physicians, it is, and only - I was also reminded by Medicare that I'm only ordering 2% of my tests as portable, because OHSU does not offer it just yet, but only nationally around 10% of the testing is done as home portable testing, 90% of the testing is done as in-house, and I personally feel the reason is because people are making more money off the tests so they are doing.

Craig Blackmore: There's been - I mean I can say from my own experience that there's been a recent trend in commercial insurance to stop paying for the in-house and only pay for the portables because they're much less expensive, but again that's - different payers have completely different practices, and that's not evidence based. That's based on perception of what's going to cost the money, which is part of the deal. Okay. So…

Richard Phillips: Can we be generic about It to say that we will not, say, allow phase testing. We won't cover phase testing?

Craig Blackmore: We could say that.

Richard Phillips: And, well I mean based on the evidence and really not specify which tests we want to use because we don't have the sensitivity, specificity information and leave it to the physician.

Craig Blackmore: We certainly have that option. Did somebody over here have - do you have a question?

Man: I was just curious. If you look on slide 12 when it talks about the diagnostic abilities of the other. I guess those are just the questionnaires. So, I'm curious - I guess I'm curious in terms of the class 3 and 4 testing. Do we have a sensitivity and specificity of that, and is that - do we have that relative to?

Richard Phillips: Yes.

Man: And that's relative to inpatient assessment.

Craig Blackmore: Inpatient is the gold standard.

Man: Okay.

Craig Blackmore: Or the reference standard. I'm not sure it's a gold standard. So, I guess I will ask the question of the directors, what is the current policy regarding sleep testing in-house, or in hospital and portable? I don't know if it's - it's probably in the handout somewhere.

Steve Hammond: The only agency that has a policy that address that is Medicaid, and they require the testing to be done in an approved sleep lab.

Craig Blackmore: So they only pay for the in-hospital - the gold standard.

Steve Hammond: Correct.

Chris Standaert: It looks like the portable tests are cheaper, but sensitivity and specificity is in a pretty big range for some of these.

Akran Khan: But you have to remember when you are looking at the specificities, the variation is in the lower end. So, looking at somebody it is difficult for the home testing to determine if somebody has an AHI less than 15, but they are fairly good in determining somebody with an AHI more than 15 on a fairly consistent basis. So the kind of people that you're interested in covering, you would get good data based on a portable test where cost saving was criteria or a consideration.

Chris Standaert: What I was going to say is that if you have the specificity and sensitivity are off, you run the risk of under and over diagnosis if you're doing a test that isn't as sensitive or specific, and then again we - I appreciate your comment - but we are given no data whatsoever on subpopulations for testing, which would be nice if it exists. That would be very nice to know, but that's the only - my point was a concern as sort of using under sensitive and under specific tests is that you either under diagnose something we think is a significant medical problem if it's bad, but maybe it doesn't under diagnose those people because it finds the people who are bad. Or you over diagnose and then you lead to expensive treatment, CPAP and surgeries and all sorts of other things that get expensive. So, is there data on subpopulations where you think they're more effective in the people - the clinically relevant patients. Is that?

Steve Hammond: I think it's - the answer is no, not in subpopulations. However, in slide 13 the overall summary of the diagnostic ability of these tests. Where it says moderate strength of evidence for portable monitors type 3 and 4 where the strength of evidence is moderate what that means is what Dr. Khan was saying. In terms of the predictability of that test to agree with an elevated AHI suggestive of OSA, as compared to the gold standard, if we're calling it that, but the actual estimate of what the AHI is, is where the variability is. So, it can't tell you exactly whether the AHI is 17 or 32 and compare that with the PSG, but it can tell you that it's an elevated AHI, high enough to be suggestive of OSA. So, that's what that moderate strength of evidence is. So, I think you do

have evidence to show that those type monitors can predict those folks who have an elevated AHI, which is what we're after.

Chris Standaert: They can predict them. So, is there - I assume if you did a portable monitor you would - on some people who are wildly abnormal, some people who are minimally abnormal or normal, and some people who are somewhere in the middle, and do you want a situation where you do that and then you turn around and go do a sleep study?

Akran Khan: No, not necessarily. So, I think the role of the portable sleep monitor is that you have a high pretest suspicion that your patient has sleep apnea. Your pretest probability…

Chris Standaert: Based upon?

Akran Khan: Based upon your history and symptoms, history of snoring, wife noticing apneic episodes, history of hypertension, patient telling you that he is having snoring/snorting related arousals. His BMI is elevated.

Chris Standaert: But we were just told there is no validated clinical predictive rule.

Akran Khan: I know, but these - but this is not a prediction rule. This is using a clinical - the same rules but using them as a clinical suspicion. So, somebody comes in with all these features, I have a high suspicion that this person has sleep apnea. Another person comes in whose 20 years old, does not have enlarged tonsils, doesn't have any anatomic abnormalities, and is telling me that he's tired and fatigued and does not have hypertension. I have a low clinical suspicion that this person has sleep apnea. So, I may still want to do a test, but doing a portable sleep monitoring on this second person is not going to be very helpful. I have a high suspicion there, I can send that second guy for a portable sleep testing. My suspicion is more than likely to be fulfilled, and at that point in time I can decide that I can put this person on a self-adjusting CPAP machine and see if he improves and has symptomatic improvement.

Craig Blackmore: Okay. I need to hear from other members of the committee about where they are on this issue.

Man: I guess I'm trying to figure out exactly what we're trying to say here. Are we trying to determine cover sleep testing or cover

sleep testing under conditions, meaning cover inpatient or cover outpatient, or cover either of the 2 under certain conditions? So, I guess that's - I'm trying to still sort of figure out where we are. I think that's - I guess I should say I'm summarizing. I think that's where we are. Do people agree that's where we are?

((Crosstalk))

Man: So, we're talking about the conditions, and it seems like we don't really have any data to support what the conditions are other than the concept of increased clinical suspicion might make home testing more likely to be reflective of the true disease.

Man: You're right, but to further complicate it, we've already voted to treat. So, we have to use some kind of diagnostic tools. So, we either use the PSG as the only tool or we use the PSG with home monitoring, or the questionnaires. There's - so we have to choose. We've essentially by default chosen to use the PSG, because there isn't - I mean that's what we're calling the reference data.

Craig Blackmore: Yeah, but we can put conditions on the PSG and say only…

Man: Right, so the question then is do you want to apply conditions to it or not?

Seth Schwartz: And there's 2 different ways the conditions could look. The conditions could look only PSG, only the other things, or sleep testing yes, but only when there's clinical suspicion.

Craig Blackmore: Only certain patients.

Seth Schwartz: Only certain conditions.

Craig Blackmore: We can condition patients or we can condition the test.

Seth Schwartz: Right, and so I guess what I was getting - what I was kind of getting to is that I don't think we have the data to be able to make a determination of whether they should have an inpatient PSG or home testing. It's not quite equivalent, but I don't think we have enough data to make a call. So, I would be inclined to think - we talked about, okay sleep testing is okay, and then we just decide whether or not we want to put clinical constraints on that or clinical conditions on who can get it, or we simply can say cover it

and then assume that providers are going to use it when they have a clinical suspicion that they should use it.

Marie Brown: Well, the sensitivity of the 4 questionnaires they've included here range from 74% to 93%, so.

Seth Schwartz: I think the questionnaires are a different issue, because there's no cost in a questionnaire. That's the doctor hands out a questionnaire and says fill this out. I think the question, unless you're getting at should we be using these to set a condition.

Marie Brown: Well, it seems like to include them. Dr. Burke's talking about his clinical judgment or assessing a higher-risk population. That's what these questionnaires do is assess a high-risk population.

Craig Blackmore: It's what they try to do.

Chris Standaert: The sensitivity and specificity are about the same as the outpatient sleep studies and so I mean you could just sort of do a questionnaire and stick somebody on CPAP, I mean clinically I suppose. Or you have a clinical suspicion and you stick them on CPAP if it's high, but I'm going back to what Seth said. We have an NCD that we have to respond to in some way that basically covers all the things. It has more criteria for 4 than 1, 2, and 3, and we could just do a cover and people do sleep tests and not get into the granularity of it all.

Richard Phillips: That was my thinking, because I don't know what the real difference is between cover with conditions and cover ultimately. There's only so many tests you can do, and I think what it would be coming down to is, okay, we're going to be deciding between 3 and 4 and 2, and I don't think we've got that kind of…

Craig Blackmore: I think the risk is if we cover without conditions, then anybody could get a sleep test whether they need it or not, and the state would have to pay for it. Now, do we have the ability to say who needs it? There's no valid clinical prediction rule. In my opinion, in absent data I don't know how we can be better than the clinical judgment if we think the test for the treatment is effective and we need to identify their (unintelligible). I mean that's where I'm struggling.

Chris Standaert: But a condition of clinical suspicion - I mean, it's sort of a useless condition to put on something, because why would? I mean, you

say you can only cover it if you think they might have it. Well, why would you possibly go do it if you didn't - and if you were doing it just to make money you would just say I think they have it. I mean, that's not - there's nothing objective in there. So, purely subjective conditions on when you can do it are very difficult, as a condition. It doesn't make a lot of sense to me.

Seth Schwartz: I guess the only other condition I would think about is, how many of these do you do? Should we put any constraints? I mean, I've heard of there are circumstances where you kinda you do one and then the person goes on a diet for a week and you do another one and then you put them on CPAP at one level and you do another one. I don't know if that's a major concern about repeated studies, or if it's not then obviously we should just drop this, but might we want to put some constraint on the number of studies?

Craig Blackmore: Do the agency directors have a sense for how often people are getting more than 1 test.

Man: (Unintelligible).

Craig Blackmore: And they may not, because they're not covering portable anyway.

Dr. Nebohar: Dr. Nebohar from Washington Medicaid again. We use a Center of Excellence Model and part of the reason was to accommodate these types of concerns. So, in our verification process for Center of Excellence, the sleep lab needs to be accredited by the American Academy of Sleep Medicine, the physician needs to be board certified, the PSG technician needs to be accredited, and the sleep center itself goes through a 5-year review process. Part of that process includes not only an on-site evaluation of the physical structure of the lab itself but also includes things like the physician maintaining board certification, the techs maintaining their certification, and they also evaluate integrator reliability for their results. So, that's why the COE model was adopted by Washington Medicaid.

Craig Blackmore: So, your COE Model applies to a sleep lab basically, in-hospital testing.

Dr. Nebohar: Yes.

Craig Blackmore: But is that applicable…

Dr. Nebohar: It includes hospital testing as well.

Craig Blackmore: Is that applicable…

Dr. Nebohar: As long as they're accredited.

Craig Blackmore: Is that applicable to the outpatient home testing paradigm?

Dr. Nebohar: No, not at all. That's a noncovered benefit.

Craig Blackmore: Currently.

Dr. Nebohar: Yes.

Chris Standaert: But in an odd way, that can be driving up the cost because you're not letting people do a cheaper test that might be just as good in patients who have a high clinical suspicion. So, that may be counter-productive. I mean, I guess you could put conditions like has to be ordered by somebody board certified in sleep medicine.

Dr. Nebohar: It's ICD-9 restricted to narcolepsy and sleep apnea.

Craig Blackmore: Other committee members have input?

Marie Brown: Well, there is moderate strength evidence for the portable monitors, type 3 and type 4, in the evidence report, slide 13.

Chris Standaert: Yeah. That was actually very confusing, because type 2 seems like they are closer to an in-house sleep study if their level is lower, which may just be the number of studies, not actual clinical benefits. So, that's - the moderate to low is actually an artifact of the amount of studies as opposed to the clinical relevance of them, which is sort of interesting.

Seth Schwartz: Am I wrong, or it seems like we can either - we all agree that it's confusing that the literature around the portable devices is not very helpful. So, we either choose to make a blanket affirmation and cover all of them, really, because none of us feel comfortable with what we know to put conditions. So, that at least allows the home monitoring, or we don't allow the home monitoring. So, if we look at the level of evidence being moderate for the 3 and 4, type 3 and 4 portable monitors, really the question is do we feel that's acceptable or not?

Richard Phillips: Rather than get into the particular tests, would it not be reasonable to insist that the study either be done at a particular location, as is done with the LCD, or that it be interpreter by someone whose certified in whatever the certification is and leave it at that and then let the decisions be made to try to - my point being just try to do that to eliminate the abuses that might potentially occur without getting into specifying the tests.

Craig Blackmore: It gets back to, the current Center of Excellence model wouldn't seem to apply to home testing, but is it possible to adapt that to a sort of Center of Excellence model, if you will, that also includes portable testing.

Akran Khan: Can I point something out?

Chris Standaert: The NCD includes portable testing.

Craig Blackmore: The NCD includes portable testing. Let's look at the NCD.

Akran Khan: It does.

Craig Blackmore: Okay. So, I think again we're back to the scenario of it looks to me like we're either going to be in a cover or a cover with conditions mode. So, what we need to do at this point is come up with cover with conditions might look like and one proposal is that we simply, again, adopt the National Coverage Decision, which is on page 383, as are conditions for coverage. The other choice is that we come up with our own or amend that in some way. So, I'm looking for feedback from the committee specifically on the National Coverage Decision, as choice for conditions or if we wish to come up with an alternative.

Chris Standaert: My only big issue with the National Coverage Decision is that it doesn't mention anything about the qualifications of the people ordering or interpreting the tests. So the treating physicians or anybody could order one and interpret it, which is quite contrary to at least the Center of Excellence model, but.

Seth Schwartz: But the Center of Excellence model also doesn't allow portable testing. It doesn't cover portable testing.

Chris Standaert: No. It's probably not…

Seth Schwartz: I mean, let me spin out the scenario for you. You come in my office. I'm suspicious that you have sleep apnea. So, I can either send you to the sleep lab. You will never get a portable test at home from the sleep lab. You'll get a PSG. I can order the portable test and then I can decide that I'm comfortable using some cutoff or not, and order CPAP for you independently of the sleep lab, or I can order the portable test and send you to the sleep lab, in which case you'll get a PSG. That's the reality of the situation.

Chris Standaert: Right. I guess, could you allow it all but only PSGs in certified sleep labs?

Marie Brown: They're still going to be doing it every time.

Chris Standaert: No, no, no. So, you can't have people who don't have - you have all these certifications for a sleep lab, and I guess our issue is that it looks like the non-sleep lab testing, outpatient testing, actually is relatively good in the right patient. So, if you just follow the sleep lab model and say you only want COE sort of sleep labs, you're restricting people from getting a cheaper test that might actually be more effectively used in a primary care setting, but what you also don't want done is you don't want uncertified sleep labs popping up all over the place for sleep apnea. So, you may need both. You need a mixture of sort of letting people do the outpatient testing, but when you have a sleep lab, have a certified sleep lab that has to go through all these criteria to actually be a sleep lab.

Seth Schwartz: In order to do PSGs.

Chris Standaert: In order to do PSGs, yes. But allowing patients treatment and diagnosis either way.

Craig Blackmore: Dr. Khan.

Akran Khan: I think - I just want to point out one thing. A lot of private insurances have moved to portable testing as screening, and we talked about an accredited sleep center being the gold standard that we are using. So, American Academy of Sleep Medicine does realize that people have moved to out-of-center testing, and they are providing accreditation to out-of-center testing centers. So, now they have in-lab testing centers that are accredited, and they are willing to accredit your program to an out-of-center testing

accredited center where they will be done by certain standards, and the sleep studies would be read by a sleep physician and not by a dentist or anybody else in prescribing any kind of a device. So, you could ask that the sleep out-of-center portable testing be done at a center that is certified by American Academy of Sleep Medicine, which is the same criteria that you are using for the in-center testing.

Chris Standaert: We're requiring certification by the American Academy of Sleep Medicine currently for the inpatient settings from an independent medical society?

Craig Blackmore: Who certifies the Center of Excellence currently for Medicaid? I can't remember her name. Dr. Nebohar, are you…?

Dr. Nebohar: It's the American Academy of Sleep Medicine criteria, but the sleep labs have to actually apply for COU recognition from the agency.

Chris Standaert: Right, the state certifies them, not some other independent body. I think it'd be unusual to bother.

Craig Blackmore: I mean, we could have a condition that the portable could only be performed by approved providers by the state and then leave to their discretion what criteria they used for that determination, because I'm not sure we're equipped to judge the merits of various schemes, and that would leave them some leeway they could approve everybody or they could have a more stringent criteria, as they saw fit. Is that enforceable for the directors?

Man: Yes, I think that would definitely be enforceable.

Craig Blackmore: I don't want to get us into defining credentials but allowing for the mechanism to exist. Okay. So, I think we're honing in on the draft conditions being the National Coverage Decision, which is somewhere, with the additional criteria of ages 18 and over and with the additional criteria being that sleep testing is only covered when provided by a state-approved provider. Any thoughts on that?

Seth Schwartz: No, I think that sounds great. I just had one question for the agencies. So, if we specify 18 and over for this situation, if you have a child who needs a sleep study, does that mean that we'd

be saying that they're not covered for a sleep study, or does that mean that this particular decision is something that doesn't apply.

Richard Phillips: Yeah, I would interpret that as just being silent on the coverage question, but Dr. Nebohar, would you agree?

Dr. Nebohar: For the children, the process is slightly different for children who are 10 years of age or younger, we actually ask that they be evaluated by an otolaryngologist prior to their sleep study, frequently because they're being evaluated for congenital anomalies of the airway, and that's basically the only other criteria that's different when the children are evaluated for a sleep evaluation.

Chris Standaert: But if we say 18 and over, that means the state can do whatever they want in terms of their existing policies or new policies to cover. We just aren't commenting on that.

Seth Schwartz: I just want to be clear that we're not excluding sleep studies for children. We're just simply saying that this - that we're not talking about them here.

Craig Blackmore: We'll make that explicit in our decision, because I think we want to make sure there's no ambiguity. So, the minutes will reflect that we are only discussing individuals 18 and over, and we'll make that explicit. Okay. So further discussions?

Chris Standaert: She needs to put up the sentence about state-approved providers.

Craig Blackmore: Okay. I'm not sure we did the nonbinding votes, and just on a procedural note, we should - did we do that?

Seth Schwartz: Yeah, we did.

Craig Blackmore: We did that. Good. It's been a long morning, already. Okay, so now we're getting to the…

Chris Standaert: But we didn't did the nonbinding vote for…

Craig Blackmore: All right. We're going to do a nonbinding vote. So this is going to be procedural. Is sleep testing, either portable or in-hospital, more, less, or equally effective to not using the machine to test? So, clinical exam is your comparator and the answers - votes.

Josh Morse: 7 more.

Craig Blackmore: Okay. And is it of equivalent, more, or less…

Josh Morse: Can I change my vote? I want to change my vote.

Craig Blackmore: Okay. You want to revote?

Josh Morse: 6 more, 1 equivalent.

Craig Blackmore: Okay, and the next is on safety. Is testing with either in-hospital, in sleep lab, or portable equipment safer more, less, equivalent, or unproven safety compared to…

Seth Schwartz: Compared to what? Nothing?

Craig Blackmore: Compared to not testing. Compared to the clinical exams.

Josh Morse: Let's see, 6 equivalent and 1 more.

Craig Blackmore: Okay, and then finally, is testing with equipment in either the lab or portable more, less, equivalent, or unproven cost effectiveness compared to using just clinical exam?

Josh Morse: So 7 unproven.

Craig Blackmore: Okay, having done that, now we can move onto the binding vote, and the binding vote is State of Washington will cover sleep testing unconditionally, meaning anybody - and this decision - this is the State of Washington will cover in the age of 18 and over, that's all we're voting on. The state will cover sleep testing for everybody, which would be cover; it will not cover sleep testing for anyone, which would be no cover; or it will cover with conditions and the conditions we've defined as the National Coverage Decision with the additional statement that the decision only pertains to individuals 18 years old and over and coverage will only be provided for state-approved providers, and it's explicit that this refers to all - that it refers to both home testing, as well as testing in a sleep lab. Is that clear before we vote? Okay.

Richard Phillips: This should be for obstructive sleep apnea. Only obstructive sleep apnea?

Craig Blackmore: Yes. This is specifically to obstructive sleep apnea. Thank you.

Josh Morse: 7 cover with conditions.

Craig Blackmore: We charge staff with formulating that into a final document, which we will vote on at the next meeting. We are charged with either agreeing with National Coverage Decisions or delineating why we disagree. In this case, we did amend the National Coverage Decision. We amended it to ages 18 and over because that is the area of literature that we've covered, and our decision is not applied - we've made no decision on individuals under the age of 18. We also made the addition that only state-approved providers would be covered - would be reimbursed for this procedure, and I'm going to summarize where I think the group is on that and that's that we believe that some level of expertise is required in interpreting this test, which shows some variability in the literature, and we believe that having a process where the state approves providers improves the quality of the testing. This portion of the meeting is complete. We move on to the second portion. Want a 5-minute break?

Group: No.

Craig Blackmore: All right. We're moving on. We welcome Michelle Simon. Our quorum has grown. Second question for discussion. The topic is bone morphogenic proteins for use in spinal fusion. We will start with scheduled open public comments. I will apologize to anyone who is coming here expecting us to be on schedule, but there it is. We have had several people approach us in advance and will be given the opportunity to talk. Do you want to run this part, since you're much more.

Josh Morse: We'll start with the scheduled public comments. We have 5 commenters right now. We'll start with Dr. Rooney. Is Dr. Rooney present? Okay, we'll move on.

Craig Blackmore: Not everyone was here this morning. So, if you are going to speak to the committee, we ask that you identify yourself, tell us who you represent, if it's just yourself or if it's an organization, a society, a company, whatever, and also tell us if you have any conflicts of interest, if anybody has paid for you to come here, your travel, or whatever it is. I guess I should also say that we will limit the time, because we have only a limited amount of time,

and Margaret will be holding up signs that tell you when you have 1 minute to go and 30 seconds to go. Thank you.

Josh Morse: Is Dr. Rooney here? Dr. Ratliff. Thank you, Dr. Khan, and Dr. Ratliff has 5 minutes.

John Ratliff: (Unintelligible).

Marie Brown: Mm-hm.

John Ratliff: My name's John Ratliff. I'm a neurological spine surgeon at Stanford University. (Unintelligible). Can everyone hear me okay? Is the mic on? Very good.

Marie Brown: You could get a little closer to the mic.

John Ratliff: I have no conflicts of interest to disclose to the group. I come here representing the American Association of Neurological Surgeons, a group of 8000 neurosurgeons. We were identified in 1931, and also the Congress of Neurological Surgeons, which also represents 8000 neurosurgeons throughout the world. They're the same 8000 neurosurgeons that are in (unintelligible). My travel will be reimbursed by those groups. Otherwise, I have no conflicts. We prepared a written statement, which I believe is included in your packet. I will limit my discussion t that written statement and to the health technology assessment itself. We'll respond to the 5 key questions provided by the HTA and/or bone morphogenic proteins primarily infused is what we're discussing today, we feel are comparably safe and effective bone graft alternatives, appropriate in patient with medical indications, as determined by their treating surgeon. The FDA has approved nonlabel use of BMP based upon equivalent or superior fusion range, shorter operative times, and also decreased bone donor site complications. We believe (unintelligible) of our bodies (unintelligible) that the literature supports the use of bone morphogenic proteins in these select patients from anterior lumbar interbody fusions, posterior interbody fusions, and they're an appropriate substitute for single-level posterolateral fusions. Now I'll move through the questions as quickly as possible hopefully to stay within my 5 minutes. Question number one looks at validated instruments and outcome measures in the spine surgery, and it touches on a lot of controversy that our specialty is faced in assessing and determining many different pools, only one of which has been validated. We have a national

registry the neurosurgeons developed where we are polling patients who have undergone spine surgery where we could have chosen this group of outcome measures, but that's on ongoing research effort. We're not sure that we have the answer yet, as to what the best outcome measures are for assessment of spine surgery patients. That's an ongoing process. The second question that the HTA puts forth is efficacy and effectiveness of BMP. We believe the HTA review is quite thorough. The conclusion of HTA is that there is evidence in the literature to support efficacy and also effectiveness of on-label and off-label use of InFuse or rhBMP-2 in lumbar and cervical fusions. There is also off-label support for a separate product, BMP-7, which is no longer on the market. These conclusions echo the physicians held by the American Association of Neurological Surgeons and also the Congress of Neurological Surgeons. One issue has become that the original studies supporting InFuse, or supporting BMPs, were non-inferiority studies. They were designed in such a fashion because they were based on the gold standard use of autograft. We have developed better understanding and more information of BMP, as we've used them more. At present in 2008, BMPs used are approximately 40-50% of degenerative thoracolumbar spine surgeries. It's a very widely used product. We have much more information about InFuse than has been reported in the original studies. There have been approximately 70,000 patients who have had InFuse used for degenerative thoracolumbar conditions. So, again, this is a very large group of patients who these decisions may impact. Safety is one of the primary issues brought up by the HTA. The very thorough review of the safety issues on BMP within your group. Now, and the document has been prepared. There are significant potential complications with autograft bone harvest. There have been well-reported what they feel the use would be in the cervical spine. This has led most spine surgeons to move away from using BMP, especially in anterior cervical spine procedures. The surgeons who do use that, they're using additional medication, and (unintelligible). With the exception of anterior cervical procedures, the literature does not support that the complications raised either on-label or off-label are significantly higher than in those patients undergoing autograft harvest. The reports of Dr. Carragee, and Dr. Carragee's one of my colleagues from Stanford, very reasonable guy. The points made by Dr. Carragee and brought up by the HTA reflect more on the peer review process and decisions made by physicians receiving large amounts of funding from Medtronic with regards to their investigation of BMP. They do not, however,

impact BMP itself, which is still a useful product and a product that has significant efficacy. Exclusion criteria left out smokers. Other patients where a different (unintelligible) couldn't be seen and only had limited reports, but we do see some health (unintelligible) with smokers and other complex patients, and again, cost effectiveness is the real issue. Multiple cost analyses have shown cost savings through quicker rehab, peer revisions, and decreased narcotic use. We appreciate the opportunity to review these findings with the Washington State HTA and we appreciate your interest. Then again, we believe BMP is a viable alternative to autograft in clinically appropriate cases, as chosen by treating surgeons. Thank you for your attention. This is a group that developed this report and again thank you very much.

Josh Morse: Thank you, Dr. Ratliff. Dr. Tredway. Ok, Dr. Shuster. Hello. Thank you. If you could please state any conflicts.

John Shuster: Sure. My name is John Shuster. I'm a spine surgeon from Spokane, Washington. That's on the east side of the state. I stand before you pure as the driven snow with no conflict of interest, whatsoever. I did get a ride here yesterday and got to sit blocked by an avalanche at the pass for 3 hours yesterday. So, I'm coming here basically as an engineer whose now a spine surgeon whose been in practice for 15 years. I was in academics. I now am private practice for 15 years with a group of 25 orthopedic surgeons in Spokane doing 80% spine and about 20% trauma work. Basically, I would judge safety, efficacy, and lastly and least importantly cost, but I think it's wrapped up by probably best summarized by the other members of the committee. Why do we even do spine fusions? Some of you may or may not know. Most basic, a spine fusion is done to stop abnormal motion, because some people have abnormal motion that causes pain, some people have pain without abnormal motion at all, and some had the abnormal motion without pain. You can see this patient has painful slippage called spondylolisthesis. We all know that achieving a solid fusion is critical would sound most and so basic, but Jeff Cornblum figured that out and documented very well that a solid fusion performed radiographically correlates with a better reduction in pain. We used to use bone from various sources, the tibia, the iliac crest shown here, and using the bone that we harvest when decompressing nerves. We also have expanded that t get more volume with good cadaveric allograft. There's limited supply but an unlimited supply of this, and there used to be disease transmission worries, but there aren't anymore, and it

even goes to using Xenograft, such as porcine, bovine, and elephant ivory as bone graft. We don't do that anymore; rhBMP2 is fortunate to be in my spine fellowship with Scott Bodin when it was used in the United States for the first time. It was approved in 2002 to be specifically used in 2 different kinds of metallic cages at the bottom 2 levels of the spine for what's called an anterior fusion, going through the front. This is the safest place to put it. It's away from the nerves. It's in a confined space. It's what's called osteoinductive. It actually causes the stem cells to change into bone-forming cells and form bone. Well, what do we do with it? We solve problems. We previously had a lower fusion rate, a greater complication rate because we harvested bone graft, or we used xenograft. We eliminated the donor site morbidity that's normally harvested from the pelvis, which has about a 15% risk of cosmetic divots or chronic donor site pain, and we saved hours, in most cases, of operating room time to harvest and - to open, close, and harvest bone graft from a patient's pelvis. I, personally, only use it in relatively on-label applications, placed anteriorly inside a cage so it's confined. So, here's a 60-year-old male with disabling back pain, can't sleep, can't drive, can't sit. It's my secretary's husband. So, we did his fusion with BMP in an on-label application through the front in a plastic cage, so this is a little bit of an extension of it. There's a plastic cage there propping those 2 vertebrae a part, and then within about 6 months, here we go. Wipe the grin across his face, using it, and he goes back to playing golf with no pain. Here's the toughest thing. This is a HVAC laborer who has this slippage and collapse on the left, can't work. We do a 2-level fusion on-label with screws in the back to back it up and he goes back to work lifting heaving HVAC equipment, 100-pound lifting and ladders, 8 months after surgery. So what is it? It's the most effective way we have to grow bone. In patients who have prior laminectomy, hardship cases, we don't have the ability to harvest local bone. We could do it over the pelvis, but this gives us an unlimited supply of something that will actually grow bone, in most people's opinion, better than even the patient's own pelvis bone will do it. Here is someone who had 4 prior surgeries, got to me can't walk a block, can't brush his teeth because he can't stand up. We did a fusion with BMP. Two months later, he walked the US Open. Eight months later, he has a solid fusion and no back pain. Done through an incision on-site like that with all that hardware, you know, it looks horrible, but this takes a guy who can't work, can't do anything, and now has resumed to a relatively normal life in his late 60s. So, BMP-2. It's been proven to be safe. This is the FDA

clearance trial. This is level 1 evidence, BMP compared to bone graft with a higher fusion rate. This led to the FDA approval of 2002. In 2009, the same patients were looked at, and the results were very durable; 68% of the people were still working at 6 years after their surgery. Even though maybe other levels had gotten into trouble, 68% of those people were still working. Is it safe? Well, I think if we use it essentially on-label in a confined space, it is. So, every spine surgeon that I know doesn't believe that it results in a higher risk if they're retrograde ejaculation, but that's approach related. How you get there causes the side effects, what kind of decision to make. We don't think it causes tumors of any sort. In fact, it's thought to be protective against certain kinds of GI tumors. We understand that it has a greater risk than using an anterior cervical or posterior lumbar application because it causes such an incredible inflammatory response. Because it's making bone, it can cause a bone to form on the nerves or around the esophagus. So, when summarized, when you use it in a confined space, it solves the problem in a way we've never had before. The intangibles are hard to measure, but the feeling among those spine surgeons is, this is a useful technology to help us grow bone in either routine or hardship cases, and the loss of BMP is an option for our patients, which significantly impacts spine care in Washington State. Thank you.

Josh Morse: Thank you, Dr. Shuster. Our final scheduled presenter is Julie Bearcroft. Dr. Bearcroft has 10 minutes.

Julie Bearcroft: Thank you. Good afternoon. My name is Julie Bearcroft. I'm employed by Medtronic. I have a Ph.D. in biomedical engineering, and I want to thank you for the time today to be able to talk to you about the evidence demonstrating the clinical efficacy of BMP-2. All of my comments today will focus only on the FDA approved recombinant BMP-2 product, that's InFuse bone graft. As your report clearly identifies, amplified rhBMP2 matrix is an investigational product and not commercially available today. It's substantially different from the approved InFuse product both in its concentration and it's carrier composition. So, therefore we request that the deliberation on BMP-2 focus on the evidence related to InFuse today. InFuse bone graft has received PMA approval from the FDA in 2002 for the use in anterior lumbar spine fusions to treat patients with degenerative disk disease. previously iliac crest graft was the only reliable bone graft available for challenging spinal and orthopedic procedures. However, autograft harvest is associated with significant amount

of pain that can persist in many patients. InFuse completely eliminates the harvest procedure and the associated complications and the significant pain while improving fusion rates in longterm clinical outcomes compared to preoperative status. InFuse has been evaluated in multiple clinical trials and has more level 1 evidence than any other commercial bone graft. Outside of the spine indication, InFuse has received separate FDA approvals in 2004 for use in acute open tibial shaft fractures, as well as in 2007 for certain oromaxillofacial reconstructions. InFuse has also undergone other agency reviews. In 2002 where it was regulated as a drug in Europe, the European Commission granted marketing authorization first for the use in open tibial fractures and then later for the use in ALIF procedures. We place high confidence in these extensive external reviews of InFuse's clinical data conducted both by the FDA and the European Commission. Moreover, the evidence for InFuse, as used in the lumbar spine, supports positive and durable health outcomes. When followed for 6 years, InFuse patients reported improved patient function and decreased pain from their preoperative status. All adverse events are collected prospectively during the course of regulated trials, regardless of their potential relationship with the treatment, and regular monitoring is conducted to ensure that all events are captured. Usually, adverse events are later grouped for purposes of interpreting the comparative data from the control population. Regulatory agencies are also provided with this information, as part of the review process. Unfortunately, the categorization of adverse events under a broader, more general heading may cause confusion in the interpretation of the safety of the product. When grouped together in this fashion, it can be difficult to discern whether the adverse event is directly due to the surgical procedure, the device, or some other unrelated factors. We would note that in the final report, 2 such examples exist. First in the report, incidences of radiculitis were derived from the adverse events reported for back and/or leg pain. To clarify, this category does not include complaints for radiating pain. Secondly, adverse events labeled as an infection were interpreted as events involving infection and seroma and/or hematoma. However, this category includes many types of infections, such as even bladder infections, but does not include either hematomas or seromas. Unfortunately, without access to the patient level data from these trials, it's impossible to discern this level of detail from these broad categories alone. We request that the committee recognize the importance of these nuances in its evaluation. Again, Medtronic places high

confidence in the external reviews of the InFuse's clinical data by both the FDA and the European Commission and their subsequent determination that InFuse is safe and effective for its approved indications. We understand the concerns articulated in the report about off-label use and its associated side effects. We do not promote off-label use, and we work closely with the FDA to ensure safe and appropriate use of this important therapy. Although we do not support off-label use, we've taken proactive steps to ensure the safety of its use. We recognize the safety concern associated with the off-label use in anterior cervical spine surgery. We began proactively addressing these concerns in 2004 when we took early voluntary action to notify surgeons of the potential risks and worked with the FDA and other government agencies to add appropriate warning. Medtronic has and continues to work with the FDA to ensure that product labeling provides accurate information to the providers regarding the safety and effectiveness of the product. We understand the concerns about potential conflict of interest, and we have been proactive in ensuring that our evidence is free of any conflict of interest. In recent years, Medtronic has implemented policies placing restrictions on participation by royalty earners in clinical trials. We were also one of the first companies to voluntarily disclose payments to physicians on our company website. In demonstration of our commitment to the integrity of our evidence, we have provided Yale University with all available patient level data from all of the completed spinal trials sponsored by Medtronic, both published and unpublished, in order to coordinate an independent evaluation of rhBMP-2 and its clinical data. This is a seminal study, and we request that the committee consider deferring any coverage decisions on InFuse until after the release of the Yale findings. This comprehensive evaluation of its safety and effectiveness will also evaluate the quality of the studies, including the assessment of the risk of bias associated with the design, conduct, and report of each clinical study. The results will be released later this year in the fall. I appreciate the opportunity to address this panel and speak to the coverage decision on InFuse. Given the strong evidence demonstrating unique benefits of rhBMP-2, and the fact that it eliminates the need for the painful harvest procedure, we strongly believe it is important for the appropriate patients to have access to it. Please allow me reiterate and say InFuse and is complete data from the pivotal trials, has undergone multiple independent reviews, as part of the approval process in the U.S. and abroad, and each time the agencies have found InFuse to be safe and

effective for the approved indications. Thank you for your time, and my colleagues and I welcome any questions you may have. These are the (unintelligible). Thank you.

Josh Morse: Thank you Dr. Bearcroft. We have no one signed up to come.

Craig Blackmore: Is there anyone who has called in on the phone that wish to address the committee. If so, if you could just hang on for a minute, we're going to un-mute briefly and give you an opportunity to identify yourselves. So is there anyone there now who can identify themselves as interested in speaking to the committee on this topic? Okay. We're going to close the public comment portion and move on. Next on the agenda, I believe, is the Agency Utilization and Outcome data.

Bob Mootz: I'm Bob Mootz. I'm one of the associate medical directors at the Department of Labor and Industries in Washington State, and let's make sure I've got the stuff working right. Okay, as you've heard, bone morphogenetic protein is an alternative to bone graft, and it's aimed at eliminating morbidity associated with iliac crest bone graft. There are 2 products that have been approved by the FDA for use and one is BMP-2, one is BMP-7, and they have on-label indications, as outlined here; anterior lumbar fusion for degenerative disk disease, and there's a humanitarian exemption for another product. During the scoping for this technology assessment, a lot of the safety concerns came to light, and although we originally thought safety concerns were for all bone graft products were relatively small, some of these issues have come up, and so we have a lot of questions regarding safety. Currently, BMP is covered in Medicaid with no policy whatsoever. However, spinal fusion does require preauthorization beginning this year. At Labor and Industries, we essentially have on-label coverage for BMP and at the Uniform Medical Plan, there's no specific policy coverage, although spinal fusions are preauthorized. Agency perspective, we spent, as a state, over 4 years about $140 million dollars on spinal fusion surgery. We do know that in our Worker's Comp population, about 70% of workers who have fusion are still disabled 2 years postsurgery compared to only 35% of people matched comparisons who have avoided surgery, injured workers who have avoided surgery at 2 years later. BMP may be associated with serious complications and adverse events, as has been described in FDA notices. BMP use may involve higher complication rate in lumbar fusion than initially reported in the industry sponsored trials; overgrowth,

uncontrolled bone formation, and fertility issues. We may learn from longterm safety and better designed trials and evaluations that safety and efficacy are comparable to alternatives, and there is in fact a recent study published just last month on a Medicare population showing that it was pretty equivalent. From our limited agency data, this is a bit of a challenge to try and find out how much BMP is being used. It looks like we've got about 15% of spinal fusions do report through billing mechanisms that BMP was used, 10% in the cervical spine. That's just what we can discern from our agency data. The off-label use in the lumbar spine, we can't determine from our data. There is no specific billing requirement for BMP, so it's likely an underestimate. It looks like in the literature, anywhere from 50 to 80% of spinal fusions use BMP currently, and part of the reason is fusions are reimbursed usually through diagnostic-related group methodology. So, the specifics aren't visible. So, we're unable to determine how BMP is used in multilevel fusions, which would be an off-label use. This just summarizes the expenses, and I can call your attention to a couple of things. In the Public Employee Benefits Plan, these are our expenditures, and when we look at the ones where we can determine if BMP was used, it's about 10 to 15%. In cervical fusions, it appears to be about a similar amount during these years. At Labor and Industries, we also see a similar situation; about 8% of cervical fusions, 15% of lumbar fusions, and that's where we get that data from. Again, we have that data artifact in Medicaid where the dropoff occurs. They've had a data reporting change there, so the recent data - there's a problem there. So, our questions and concerns really focus on safety. We do know there are serious side effects associated with it; uncontrolled bone formation, problems especially in the cervical spine. There are concerns of cancer risk. We know that BMP stimulates osteoblasts, but it also stimulates some certain kinds of tumor cells, and that's been established in basic science research. There is one trial, on amplify, which is a higher concentration, as you heard earlier, where the cancer event within a 4-year period in the group that had the BMP was 4 times higher than it was in the comparison group. So, I think there's some unknown issues associated with that, and then there have been various local effects discussed. Recent issues raised in the literature include the funding bias and controversy associated with conflicts of interest and very large payments in the millions to researchers. There has also been discussion that there is over-reporting of complications in the control group treatment for iliac crest graft donor site and under-reporting of the adverse events, as was

mentioned previously. I think one of the things to keep in mind when we're talking about safety issues is, we need to have high confidence that there is not an adverse event, not high confidence that there is an adverse event. There are also some issues associated with efficacy, some trial designs. A lot of the early trials were open label, so it was not blinded. The typical style has been a noninferiority trial design where there's no expectation to show that there's better efficacy. This kind of trial design may be appropriate when we're talking about withholding treatment that would be shown to be absolutely essential, but with chronic back pain and degenerative joint disease, that may be a bit more nuance than that. The primary outcome of bone graft, iliac crest bone graft, was overestimated in these trials. All pain in the gluteal region was attributed to the donor site harvesting, but that kind of pain is common in all low back pain and so that was not clearly discerned. Another key issue here is the BMP was studied through the FDA device classification premarket analysis process where BMP is a biologically-active agent, and so it wasn't subjected to some of the more rigorous design issues associated with pharmaceuticals. There has also been some critique of some of the control groups in studies where the comparison group did not receive optimal care. There is inadequate information for off-label use, especially in the cervical spine. Cost is much less of a concern for us. The direct cost for BMP, per se, is not determinable in our agency data. The direct cost to hospitals appears to be more than alternatives, and there's been some reports on that. Cost effectiveness can't really be assessed due to the questions remaining on efficacy and BMP might be worth the cost if it's demonstrated to be safe over the longterm. Other centers have reviewed this. The ARC Report is probably the best HTA. It's a little dated now, and it essentially concluded the on-label use as appropriate. CMS has done a review, as well, and their voting and recommendations followed the ARC findings. I think kind of a summary in our thinking here is, there are still some uncertainties and controversy about BMP. There are safety issues because of the under reporting of severe adverse events. There has been over reporting of control complications, as well as the conflicts of interest, efficacy, the control treatments may have been compromised, inadequate blinding, and the noninferiority design issues. So, our summary review is that we do have these serious safety questions. Off-label use does not have reliable data, but we know it's being used in an off-label fashion in Washington. There's unclear efficacy because of these design issues, and so the value, of course, and cost effectiveness

depends on the efficacy. So, our recommendation is for off-label. We're recommending non-coverage in light of these safety issues. For on-label, as you go through this information and look at the evidence, if you think the safety concerns are valid, we would recommend considering doing a non-coverage decision. As you review it and if you decide that the safety issues are really not valid, then we would encourage you to consider only FDA-approved indications.

Craig Blackmore: Thank you. Any questions from the committee specific to the agency report?

Richard Phillips: I have a question. Can you comment upon whether the use is always accompanied by a cage device (unintelligible)? Or is that even relevant.

Bob Mootz: Well, I think the InFuse product, which is the primary one that's out there, you'll see it in the evidence report, but basically it's infused in a sponge inside a cage.

Richard Phillips: Okay. So that's - you're only talking about…

Bob Mootz: So it is contained, yes.

Craig Blackmore: So if I understand what you're saying in terms of how often this is used, you have numbers but, the 15% that you've given us, but your belief is that it's probably a pretty significant underestimation because of your reporting issues.

Bob Mootz: The 15% is just for billing, from our billing data that we know that it was used.

Craig Blackmore: Whereas, more national averages of 50% or higher.

Bob Mootz: Yeah, and that's just Washington data.

Craig Blackmore: Yeah, thank you.

Bob Mootz: But we think it's in pretty widespread use.

Craig Blackmore: Okay, so it's 10 after 2. Are we ready for a break, or should we press on with the vendor report? Press on. All right, vendor report, please.

Robin Hashimoto: Okay. So, I'm Robin Hashimoto. I'm from Spectrum Research, and I'm just going to present the evidence from our Health Technology Assessment on spinal fusion, BMP use in spinal fusion. Okay, so spinal fusion procedures typically utilize some type of bone graft to facilitate new bone formation and there are 3 properties of bone grafts that contribute to formation of new bone. The first is osteoproduction. Can you hear me?

Craig Blackmore: You can sit over here if it helps and just use that one.

Robin Hashimoto: Okay, oh that's much better. So I'm Robin Hashimoto. Okay, so I'll just start with this again. So, spinal fusion procedures typically utilize some type of bone graft to facilitate new bone formation and there are 3 properties of grafts that contribute to the formation of new bone. The first is osteoproduction, and this refers to the ability of the graft to generate bone and is attributed to the presence of bone-forming cells within the graft. The second is osteoinduction, and this refers to the ability of the graft to induce sole activation, proliferation, and differentiation of bone-producing cells. The last is osteoconduction and refers to the presence of the scaffold on which new bone formation can occur. So, there are, of course, many types of commonly used grafts and autologous bone grafts are considered to be the gold standard, and they possess all 3 characteristics of successful bone grafts. Okay, so the most commonly used form of autologous bone graft in spinal fusion procedures is the iliac crest bone graft or ICBG. Again, this type of bone graft is the gold standard, and although ICBG has the benefit of being osteogenic, osteoconductive, and osteoconductive and is generally accepted to promote fusion, there are limitations to using ICBG that are primarily related to graft harvesting. So, these limitations are listed and can include donor site morbidity, such as pain and infection. The procedure can cause nerve injury or avulsion fractures. There is also, in some patients, a concern that there's limitation in the quantity of bone available. So, of these, donor site pain has really been the primary motivation to pursue alternative bone graft materials for spinal fusion. So one of the more recent bone graft alternatives to be developed are products using bone morphogenetic proteins, or BMPs, and as their name suggests these proteins were identified on their basis to stimulate new bone formation. So, these proteins are members of the transforming growth activator super family of proteins and of the many different forms of these proteins 5 exhibit full osteoinductive potential. BMP-2 and BMP-7 are the 2 forms that

are currently available for clinical use. So, the impetus for developing BMPs as bone graft substitutes was to provide a product that is able to stimulate bone formation without the need for autograft harvesting. Okay, so this is just a very simplified schematic to illustrate how BMPs work. So, BMP is a soluble protein. It binds to specific receptors on host cells, such as missing stem cells or osteoblasts, which are bone-forming cells. binding of the BMP results in receptor activation. This leads to activation of signaling cascades and ultimately the activation of genes that promote cell differentiation, so differentiation of cells and osteoblasts or proliferation of osteoblasts. Okay, so there are 2 products that are currently approved by the FDA. So, the first is InFuse, which you've heard about. This is a Medtronic product and utilizes BMP-2. It received premarket approval in 2002. The second is OP-1, which utilizes BMP-7, and this product received humanitarian device exemption in 2004. So, this is an image of the InFuse product taken from the Medtronic site. So, during surgery, a specific amount of BMP is soaked into an absorbable collagen sponge, or ACS. The collagen sponge is able to function as an osteoconductive scaffold for new bone formation. It's made of type 1 collage, which is naturally present in bone, and the sponge resorbs over time. So, the BMP/ACS product is placed into the LT cage, lumbar tapered fusion device, and 2 devices are inserted side by side into the intervertebral disk space, and together they restore the natural height to the degenerated space. So, after the device is implanted, the BMP begins to dissolve into the disk space. This results in the recruitment of bone-forming cells to the site. BMP induces them to proliferate, and in the case of stem cells to differentiate. At the same time, new vasculature is formed. Once bone has begun to form, it produces it's own signaling cascades, and this further stimulates bone formation. Ultimately, this process leads to the formation of complete trabecular bone and fusion is achieved. Okay, so going back to the FDA products, I'm going to go over what they're indicated for. InFuse is indicated for primary anterior open or laparoscopic fusion at 1 level between L4 and S1 in patients with degenerative disk disease and up to grade 1 spondylolisthesis who have failed at least 6 months of nonoperative care. OP-1 is approved for revision posterolateral lumbar fusion in patients who are compromised by factors such as diabetes or smoking or osteoporosis who aren't expected to have fusion or be able to have autologous bone and bone marrow harvest. So, when BMP is used for these specific indications, BMP-2 or BMP-7, that's considered to be on-label use, and all other uses are considered

to be off-label. So, Ong and Kl reported a large database study on BMP use in 2010, and the authors utilized the nationwide inpatient sample database, and this is a large database that contains representative data from hospital discharges. So, over 340,000 procedures were performed that used BMP between late 2002 and 2007. The authors found that the use of BMP increased over 4-fold between 2003 and 2007, and of the 340,000 procedures done, approximately 92% were for spinal fusions. Interestingly, over 85% of the procedures were for off-label use. Okay, so questions remain about the safety, efficacy, and effectiveness and cost effectiveness of BMP-2 and BMP-7 in spinal fusion procedures. So as such, the aim of the report was to systematically review, critically appraise, and summarize comparative evidence on the clinical efficacy and effectiveness, safety, and cost effectiveness of BMP use in spinal fusion. The specific key questions will be discussed in turn. So, we employed strict inclusion and exclusion criteria. For inclusion studies, must evaluate patients with back and/or leg or neck pain who undergo lumbar or cervical fusion with BMP-2 or BMP-7, and all studies that we identified compared fusion with BMP to fusion with an alternative autograft, allograft, or bone substitutes. So, outcomes of interest were defined by the state. For key question 2 on efficacy and effectiveness, the primary outcomes of interest included pain, function, and fusion. Secondary outcomes included perioperative outcomes, such as length of hospital stay, as well as outcomes such as patient satisfaction and return to work. Safety outcomes of interest, of course, were complications and adverse events, as well as second surgeries. So, we conducted a systematic literature search, and the results of that search are shown here. After systematically reviewing whether each identified reference met our inclusion criteria. We ended up including a total of 36 articles for key question 1, 115 articles for key questions 2, 3, and 4, and 3 articles for key question 5. This slide shows a little bit more detailed information on the quality of literature available. For key question 2, we included 14 randomized control trials, as well as 15 cohort studies. These studies were also included for key question 3 on safety. For safety, we additionally included 12 cohort studies, as well as 33 case series and 16 case reports. Okay, so moving into the results, key question 1 asks about expected treatment outcomes, whether outcome measures have been validated in patients undergoing lumbar or cervical spinal fusion, and whether the minimal clinically important difference has been defined in this patient population. So, to look at expected treatment outcomes, we

surveyed all comparative studies included in key question 2 to identify the most commonly used outcome measures, and as you can see, the Oswestry Disability Index, or ODI, the Neck Disability Index, or NDI, the SF-36 Health Questionnaire, and the Visual Analog Scale for Pain were the most commonly reported outcome measures, and all 4 of these outcome measures are patient reported. The other outcome measures identified were much less frequently used. And fusion, of course, is a very commonly reported outcome measure and should be viewed as a surrogate measure, because fusion doesn't necessarily correlate with clinical outcomes, and the definition of fusion used varied between studies with common requirements including the presence of bridging trabecular bone between the transverse processes, absence of motion, and absence of radiolucent lines over more than half the implant surfaces. So, going back to these 4 most commonly used outcome measures, we determined whether they had been tested for validity, reliability, and responsiveness, and whether the minimally clinically-important difference had been defined in spinal fusion patients. So, as you can see, the SF-36 was the only outcome measure that was validated or tested at all in spinal fusion patients. For the other outcome measures and tests, everything was done in either spine or neck pain patients. So, next looking at the minimal clinically-important difference, we interpreted all of our results in the context of these numbers. So first, looking at the ODI, we didn't find any single agreed-upon value, but we did identify one study that compared lumbar fusion with BMP to autograft, in which the FDA required an ODI change of 15 points for ODI success. So, this is the number that we used to consider that there was a minimally clinically-important difference occurring. For the NDI, again, we didn't find any single agreed upon value. We identified 3 studies in which an MCID value ranged from 7 to 19 points. So, we interpreted results based on a 15 point change, and then similarly for pain, we identified studies that use an MCID value that ranged between 2 and 29 points on a 100 point scale. A recent HTA used an MCID value of 20 points for patients with chronic low back pain, so that's the number that we chose to use.

Seth Schwartz: Can I get a point of clarification, please? So, if 85% of these procedures are done for off-label uses and when I look at slide, well they're not numbered, the slide you have up now, am I interpreting correctly if I say that the first 3 rows are evaluations of pain in patients who had it used in an off-label fashion?

Robin Hashimoto: It's not necessarily where BMP was used at all. These are patients that - these are studies done in patients that had spine or neck pain. So, first we looked for studies in which patients had undergone fusion and when we didn't find very much, then we moved on to the next level and looked for validation in patients that had spine or neck pain.

Chris Standaert: This is just a question on the measures used in studies of (unintelligible), not necessarily BMP.

Robin Hashimoto: And we didn't identify any studies that reported the MCID for the SF-36. Okay, so moving into key question 2 on the efficacy and effectiveness of BMP-2, because there were so many different ways in which BMP was used, I've divided the section up according to whether BMP-2 or BMP-7 was used, whether it was used for on or off-label uses, and whether it was used in the lumbar or cervical spine. Because we do have so much of it, I'm going to focus in on the highest level evidence and provide additional information only when we think it will be helpful to the committee. Okay, so for the efficacy of BMP-2, on-label use in the lumbar spine, 2 RCTs met our inclusion criteria. These patients underwent primary single-level open anterior fusion with InFuse or iliac crest bone graft and BMP was used at the approved dose of 4.2 to 8.4 mg per patient. Patients were followed for 24 months, and these 2 studies served as the pilot and pivotal trials for the FDA SSED on InFuse. So, I've divided the results up by primary and secondary outcomes. The results are shown. The equal sign indicates the outcomes were similar between treatment groups. The plus sign indicates the outcomes were improved with BMP. The minus sign indicates the outcomes were worse with BMP. Outcomes were assigned these values based on whether the differences between groups were clinically meaningful, or if this wasn't available on P values, or just on whether there were large differences in outcomes between groups. So, as you can see, except for perioperative blood loss, which was better for BMP patients, all reported outcomes were similar between treatment groups, and the strength of evidence was low. Okay, so for the efficacy and BMP-2 off-label use in the lumbar spine, we identified 6 RCTs that reported on 27 to 463 patients. There was much more heterogeneity in the interventions used here. The patients underwent primary single or multilevel fusion via a variety of approaches with either BMP-2 or iliac crest bone graft, and BMP was used at a dose that ranged from 4.2, which is where it started for on-label use, but went all

the way up to 40 mg per patient, and that was for the Amplify trials. Patients were followed for up to 24 months. So, as you can see, again, except for a couple of perioperative outcomes, outcomes were similar between treatment groups and the strength of evidence ranged from high to low. Okay, so for the efficacy in BMP-7 off-label use in the lumbar spine, 5 RCTs met our inclusion criteria. These patients underwent primary single-level PLIF or PLF, posterior lumbar interbody fusion or posterolateral fusion with BMP-7 compared with either iliac crest bone graft or autograft, and patients were followed for up to 54 months, and again, except for a couple of perioperative outcomes, outcomes were similar between treatment groups. The strength of evidence was either high or low. For the efficacy of BMP-2, off-label use in the cervical spine, 1 small RCT met our inclusion criteria. So, the study only included 33 patients. The patients underwent primary 1 or 2-level ACDF with InFuse or iliac crest bone graft, and patients were followed for 24 months. Here, there was some benefit associated with BMP use for neck disability index scores and arm pain outcomes. All other outcomes were similar between treatment groups and in all cases the strength of evidence was low. Since the evidence on efficacy was only based on one small RCT, I will also show you the evidence on the effectiveness of BMP-2 use in the cervical spine. So, we identified 5 cohort studies. Each study enrolled between 58 and 775 patients. A variety of interventions and comparators were used, and patients were followed for up to 36 months. As you can see, the strength of evidence for almost all of the outcomes was insufficient and, in general, outcomes were similar between treatment groups. Okay, so moving into key question 3, which asks about the safety of BMP use in the spine. So, there was quite a bit of data for this key question, so I'm going to focus in on some of the outcomes with solid evidence and on outcomes for which BMP may be associated with a higher or lower risk. Okay, so for this key question, we included all adverse events and second surgeries that were reported in 14 RCTs, 27 cohort studies, 33 case series, and 16 case reports. In most cases, I'm going to focus on the evidence from the comparative studies, since that provides the strongest quality of evidence. Okay, so for local safety we didn't find any differences between treatment groups and the risk of deep infections, and the strength of this evidence was low. We did find that there was a greater risk of inflammation or neck swelling with off-label use of BMPs in the cervical spine, and I should point out the data here are interpreted in the same way. So, in this case the minus sign

indicates the outcomes were worse with BMP. So, that's equivalent to the higher risk of complications. Okay, so regarding inflammation and swelling, data from 4 cohort studies and 2 database studies suggested that the rate of inflammation or neck swelling, which can lead to dysphagia or problems swallowing, reintubation, and prolonged hospital stay is higher in patients who received BMP. So, data from 4 cohort studies suggested that the risk of swelling was nearly 35% in BMP patients compared with 9% in control patients, and data from 2 large database studies suggested that the risk was 5% in BMP patients and 3% in control patients. Data from 7 case series give a risk of its complication that falls in the middle of those provided from the comparative studies. Okay, so we didn't find any difference between treatment groups and the risk of dural injury. We did find that there was a greater risk of retrograde ejaculation following BMP use in the lumbar spine, and the strength of this evidence was low. Okay, so this slide provides a bit more detail on the data for retrograde ejaculation. So, the FDA SSED for InFuse reported a 7.9% risk of retrograde ejaculation in BMP patients and a 1.4% risk in the iliac crest bone graft patients. The majority of these events presented within the first 9 weeks following surgery, although 2 of the BMP cases didn't present until between 9 and 19 months. No cases were reported for the last followup, indicating that all cases were transient. Okay, the second study was a 2011 cohort study, and it reported a similarly increased risk of retrograde ejaculation following BMP-2 use for patients undergoing 1 to 2-level anterior lumbar fusion that spanned L5-S1. The authors were able to stratify the data and found that the effect held true for 1-level fusions but not for 2-level fusions, and then for the 5 BMP patients that did develop retrograde ejaculation, 3 of them remained effected at 12-month followups. Okay, so we did not find any difference between treatment groups and the risk of cardiovascular or vascular complications, or in the risk of deep vein thrombosis. There were no differences between treatment groups and the risk of death, except as reported by 1 retrospective cohort study that evaluated BMP-2 use in the cervical spine. So, the study reported that the risk of death in 90 days following surgery was higher in BMP-2 patients compared with control patients, and the causes of death were not reported. So, the significance of this result should be interpreted with some caution, because the study provided no surgical or demographic information. We also found evidence that when used off-label, BMP may be associated with an increased cancer risk. So, a total of 6 studies reported on cancer,

and those are shown here. So as you can see, 1 on-label RCT reported no difference in cancer risks between treatment groups with 24 months followup. Three RCTs reported on off-label use of BMP. So both of the BMP-7 studies, so the second 2 studies, were relatively small. These studies enrolled 34 and 36 patients each, and then the first study, the BMP-2 study from the FDA 2010 report, was the report on Amplify, which uses a very high dose of BMP-2 at 40 mg per patient. So, at 24 months followup, which isn't shown here, the risk of cancer was 3.8% for BMP patients and 0.9% for iliac crest bone graft patients. At 60 months, which is shown the risk was 6.3% for BMP patients and 2.2% for control patients. So, in the study of the types of - 15 cases of cancer that were reported in the BMP group and the 5 in the control group, cancer types varied quite a bit. So, I'll just give you a couple. A few of them included laryngeal cancer, ovarian, prostate, colon, breast, basal cell carcinoma, lymphoma, leukemia, and there are others as well. Finally, 2 cohort studies reported on off-label use of BMPs. One found a higher risk of cancer with BMP use and then the last one found no difference between groups and the risk of specifically pancreatic cancer. Okay, regarding second surgeries, we reported on revisions. Revisions are defined as surgery that modified or adjusted the original implant, hardware removal, and that was defined as removal of one or more components of the original implant, and this included replacement with a different implant. Supplemental fixation, which was surgery to provide additional stabilization to the index site, and reoperation, which was defined as any other procedure at the index level. So, we found that the risk of revision or hardware removal was similar or improved with BMP use. There was moderate evidence that the risk of supplemental fixation was improved with BMP-2 use, and there was low evidence that the risk of supplemental fixation was worse with BMP-7 use. The rates of reoperation were similar between treatment groups for off-label use of BMPs. For graft site morbidity, we found that in patients who underwent iliac crest bone graft harvest, perioperative hip pain, visual analog scale scores range from 5.7 to 8 points on a 10-point scale, as reported by 4 studies. These scores were much lower by the last followup. Nine studies reported that the percentage of patients with hip pain at the last followup ranged widely from 10 to 66% of patients. Other reported complications of the graft site included injury to the lateral femoral cutaneous nerve, ASIS fractures, and deep infection requiring surgery. We also found no evidence of increased back or leg pain in the first 6 months in patients who

underwent fusion with iliac crest bone graft compared with BMP-2. Okay, so moving into key question 4, which asks about the differential efficacy and safety of BMP in subpopulations. We identified 8 cohort studies that provided evidence on this key question, and we found no strong evidence that there was differential safety or effectiveness in any of the subpopulations listed. Key question 5 asks about the cost effectiveness of BMP use in spinal fusion. So, we identified 2 studies that reported on the cost effectiveness of BMP-2 for on-label use in the lumbar spine. The first study is shown here, and this was the cost effectiveness analysis done in the recent AHRQ HTA on BMPs, and the study used data from 1 RCT and costs were taken from the CMS payer perspective, as reported by Medicare. So, the report found that when the BMP cost was bundled and the initial costs were the same in both treatment groups, BMP was more cost effective. When BMP was treated as an additional cost, and the additional cost that they used started at $3000, iliac crest bone graft was the more cost effective strategy. So, the authors concluded the BMP was only cost effective when the initial costs were identical between treatment groups. So, the second study was done by the UK's National Health Service, and the study used data from the same RCT as the AHRQ report. The study found that BMP-2 was unlikely to be cost effective, or BMP-2 was unlikely to be more cost effective than iliac crest bone graft, as the cost per quality adjusted life year was over 120,000 pounds. We identified 1 study that evaluated the cost effectiveness of off-label use of BMP-2 in the lumbar spine. So, this study took data and actual costs from a single RCT that enrolled patients who are at least 60 years old. Patients underwent single or multilevel posterolateral fusion with BMP or iliac crest bone graft, and the study found that although the initial costs were $2300 higher for BMP patients, the final costs were $2300 lower for BMP, as the result of a lower risk of complications. Finally, we didn't find any studies that reported on the cost effectiveness of BMP-7 or on the use of BMP-2 or BMP-7 in the cervical spine. Okay, so summary and implications. For efficacy and effectiveness of BMP use in the lumbar spine in general, outcomes were similar for both treatment groups. There was a bit of evidence that suggested that fusion rates were higher with off-label use of BMP-2; however, fusion is a surrogate outcome measure, and there were no differences between treatment groups in terms of patient reported pain or function. In the cervical spine, the evidence base was much smaller. There was some evidence that BMP use provided some benefit for arm pain, neck disability, and neck

scores and fusions. There was also evidence that suggested that length of stay and neck pain were worse with BMP use, and the strength of evidence was insufficient or low. For safety, we found that there was low evidence of a higher risk of retrograde ejaculation following BMP use in the lumbar spine. It should be noted that retrograde ejaculation is a concern with any anterior approach in the lumbar spine. This complication is generally thought to be transitory; however, one cohort study reported that 3/5 BMP patients who develop this complication postoperatively still had retrograde ejaculation at 12 months. We also found moderate evidence of a higher cancer risk following off-label use of BMP in the lumbar spine. Data varied by study, and data from one RCT, in particular, that used Amplify, which has not been approved by the FDA, reported higher cancer risks with BMP at 60 months. Again, a variety of cancer types were reported, some of which presented relatively soon after surgery, so were likely already present. Given the heterogeneity in cancer type, severity, and time of presentation, it's difficult to make a direct connection between BMP use and the development of cancer. We also found moderate evidence of a higher risk of anterior cervical inflammation or swelling following off-label use of BMP-2 in the cervical spine. This conclusion was based on data from 4 cohort studies and 2 large database studies. This complication occurs perioperatively, and the inflammatory and swelling response may lead to the use of intravenous steroids, dysphagia, and possible PEG placement, reintubation, tracheostomy, surgical exploration, and drainage, and all these things can lead to increased length of hospital stay or readmission. Okay, finally regarding second surgeries, we found moderate evidence that BMP-2 use was associated with a lower risk of supplemental fixation. We found low evidence that BMP-7 use was associated with a higher risk of supplemental fixation. The risk of revision, hardware removal, and reoperation were generally similar between treatment groups with possible benefits being seen with instances listed there. So, we thank you for your attention, and we can take questions at this time.

Craig Blackmore: Thank you. Questions from the committee specifically related to what we've heard on this presentation.

Seth Schwartz: Can I get a point of clarification? I had heard Dr. Shuster say that BMP-7 wasn't being used anymore. Does that? Can that be validated?

Craig Blackmore: Let me put you on hold for one second, and let me introduce our clinical consultant, Dr. Lee. Thank you for coming. If you could just introduce yourself and tell us who you are, that would be very useful.

Michael Lee: Sure. My name is Mike Lee. I'm one of the orthopedic spine surgeons at the University of Washington Medical Center. I believe my CVA and my disclosures are listed in the packet, as such. In regard to BMP-7, the Stryker the product, some of that information is a little bit outdated. I don't believe Stryker actually owns BMP-7 anymore. I think that's actually been sold off to a different company. As for is it in use right now, I actually don't know the answer to that. I don't use it myself, and I can't give you that information. I think most of the folks use BMP-2.

Craig Blackmore: Okay. Just to give you a little background. I don't know if you've been here before. The clinical consultant is a very important role, and we appreciate your being willing to do it. We will - we've heard an evidence report, a group of experts summarizing the evidence, but there's a lot of clinical context, and there's a lot of technical details to these procedures, and you're here to help fill us in on those questions. So, we'll be asking you things periodically, and I appreciate your participation. So, is that? Other questions about the material in the technology assessment?

Richard Phillips: Yeah, I have a question. In the off-label use, where there was an increase in safety problems, with the off - was that related to the amount of the autograft that was - autograft of the graft that was used? In other words, did it occur more frequently when there were multilevel fusions or when there was a greater amount of bone that was used? Or can you discuss that variation?

Robin Hashimoto: Do you have any specific complications in mind? I mean, I can say for cancer you can look at the data and surmise that the high level of BMP-2 used for the amplified data where it was used at 40 mg per patient instead of the on-label use, which is 4 to 8 mg per patient that the risk of cancer might be dose dependent.

Richard Phillips: Just so I understand that. When you were - I saw somewhere like 7 or 8 mg per patient. Is that a single-level fusion, whereas when you're going up does that imply multilevel fusions, or?

Robin Hashimoto: That is my understanding is that.

Michael Lee: Could you restate that again? I'm sorry, could you restate the?

Richard Phillips: The InFuse is defined as a weight, and I was wondering if there is any relationship of the weight to the complications to the safety issue and secondly, does the weight imply a multilevel fusion, so there's increased us of it? So let's say that you would expect twice as much use in a 2-level fusion or 3 times as much in a 3-level fusion, etc.

Michael Lee: To my knowledge, there is no guideline or correlation for that usage. It's largely at the discretion of the user or the surgeon regarding how much InFuse BMP they would use in on or off-label fashion.

Craig Blackmore: So, the surgeon can use 4 mg or 40 mg of the BMP-2…

Michael Lee: Correct.

Craig Blackmore: …at their discretion.

Michael Lee: At their discretion.

Richard Phillips: So, it's not contained within that plastic device itself. It comes out of the plastic device in order to be infused?

Michael Lee: So, are you referring to the on-label?

Richard Phillips: We're trying to understand how it's utilized, because it comes in that plastic wrap, whatever it is, container - the InFuse device - and somehow you have to get it into the spine, and I take it that mechanically what you're doing is you - I thought you were limited by what you could put in it. It's like a packet, but I take it that's not the case.

Michael Lee: So, it comes in a collagen sponge, which kind of resembles a moist paper towel if you will, and that moist paper towel, small strips, which can be rolled up and placed inside of a cage. In an on-label fashion, that cage would be placed in the interbody space from the anterior approach. In an off-label fashion, it can be placed wherever the surgeon sees fit.

Richard Phillips: So, it's not - the question I asked here is, it says I got the wrong impression then that it came in a plastic - in other words, it

comes, you have to put a separate titanium cage or something like that in addition to using the InFuse.

Michael Lee: It depends on where it's placed, whether it's an on-label fashion or off-label fashion. Infusion can be achieved either from an anterior fusion from within the disk space itself, and if that's the case, it's most commonly placed with a cage. It can be placed in a posterolateral fusion as well, in which case sometimes it might be placed there by itself with a bone graft in addition, usually without a carrier in that case. But it can be placed in a variety of different ways.

Richard Phillips: So, the amount - the weight has nothing to do with - you can put as much as you want to in there. I mean within - within guidelines.

Michael Lee: You can, but there are guidelines and recommendations, and we are well aware of complications that can occur with higher doses and if anything, I think it's been a trend to use smaller doses because of that.

Richard Phillips: Okay. The main thing I was trying to do was trying to get an idea of the safety and the amount that's used. That's the crux of my question.

Craig Blackmore: Sorry. I'm going to remind everybody that we're being taped. Speak into the microphone, and please announce your name, at least the first few times, until she gets your voice.

Carson Odegard: Dr. Lee, you mentioned that you don't use BMP-7. Do you have any idea or do the vendors - do we have any data on utilization of BMP - off-label utilization in general?

Michael Lee: Off-label utilization of BMP-7 specifically?

Carson Odegard: Mm-hm.

Michael Lee: I'm not really - I really don't have any data on that off the top of my head. I don't know if Robin does.

Carson Odegard: Do we have any agency dat- we couldn't.

Bob Mootz: No, we couldn't tell from our billing data, other than the 15% number.

Robin Hashimoto: Yeah, and that was the same. I reported a database study, the Ong Study. In that, they just utilized codes, and it was just general BMP use, so you couldn't differentiate.

Carson Odegard: Okay. Thank you.

Michelle Simon: Hi, this is Michelle Simon. I have a question on your page 17, the top slide there on safety and cancer. I tracked this back to the report, it was on page 221 where they actually discussed the research behind this slide. So I'm trying to understand, they had some NS's on there, in the middle section. Are you suggesting those are not significant references?

Robin Hashimoto: Mm-hm.

Michelle Simon: But the level of evidence for those studies is considered moderate, is that right?

Robin Hashimoto: Yes.

Michelle Simon: In the first one, the 0.7 to the 0.7 on the on-label RCT, that is considered low level of evidence, is that right?

Robin Hashimoto: Gosh - is that correct? I don't have that page pulled up.

Michelle Simon: 170.

Craig Blackmore: So if I could ask a question on that same slide.

Robin Hashimoto: Mm-hm.

Craig Blackmore: Can you provide the sample size on these. I'm back to the same issue of telling me it's not significant doesn't help, unless I know if it's powered to detect the clinically meaning difference.

Robin Hashimoto: Yep, just give me one second here.

Craig Blackmore: Yep, sure.

Bob Mootz: While she's looking for the specifics, I would just say that the occurrence of those cancers, of course, are relatively rare, so. They aren't powered to detect significance. The reason we rank that moderate evidence is because the state asked us specifically

to segregate on-label versus off label, and even though there are 2 different kinds of off-label. There's the BMP-7, which is hard to - I mean that mechanism is a little bit different in that concentration of active ingredient. It's difficult to compare with BMP-2. Yet, it was an off-label use in 2 of the studies for the randomized trials, and in one of the randomized trials they used the BMP-2 and they had the higher concentration. So, there's lots of ways to dice and segregate this, and what we were hoping to do is to provide an answer to the state's question of off-label versus on-label, and so we categorized those, and that color is all off-label in those randomized trials. When we're talking about safety, these are not trials that specifically were powered to do that. So, we can't expect them to be statistically significant, but it's bothersome when you have consistency like that, even though the numbers may be small. Consistency is an important ingredient to the strength of evidence, so that's sort of how we got there.

Robin Hashimoto: Okay, to answer the first part of your question, yes. So, the on-label RCT, that's considered to be no increased risk, similar risk, and that was low level of evidence, and that study had about 140 patients in each group, and then the rest of the studies, all the off-label studies were considered to overall have a higher risk of cancer with a moderate level of evidence. The first study there, the BMP-2 FDA-1, that had about 230 patients per treatment group, approximately. The next 2 studies were both pretty small with about 35 patients per treatment group. I'm sorry, 35 patients total. The second to last study used 24 BMP patients and 105 control patients and then the last cohort study, the MINE study, used tens of thousands of patients; 15,000 BMP patients and almost 80,000 control patients.

Richard Phillips: One thing that bothered me about this is the randomized control trials generally don't last more than 3 months and to think that something could cause cancer within 3 months seems a little bit unrealistic. I was wondering if you could comment upon that. What were the length of those randomized control trials that would give that kind of?

Robin Hashimoto: Yeah, and its - they're actually listed there in white at the bottom of the slide. So, the on-label RCT was 24 months followup, and then the FDA trial was 60 months.

Richard Phillips: I didn't see that.

Robin Hashimoto: Mm-hm.

Richard Phillips: I'm sorry, I didn't see that. Thank you.

Robin Hashimoto: Yeah.

Chris Standaert: And these data are from the original studies. I'm sure you're aware of the study by Dr. Carragee where he reanalyzed the data and put out different data with different rates, but I don't see him referenced in there. Did you - this is from the original published studies?

Robin Hashimoto: These are from the original. Yeah.

Chris Standaert: Which differs actually from his reanalysis of those same studies, or you did not look at that.

Robin Hashimoto: The cancer data that Carragee focused on was the second study there, the FDA 2010 report, and that's pretty much what he used. But these are the original data.

Chris Standaert: Did he find the same data.

Bob Mootz: So what he found Chris was that for the Amplify data, the FDA trial reported 1 additional case of cancer. Nine, I believe, was the number, where the randomized trial that was published and peer reviewed was 8, and the data that we report here in our report is the Amplify FDA data, the 9. So, we report the same numbers Carragee does.

Chris Standaert: Okay. Thank you.

Craig Blackmore: Any other questions for the vendor? Okay.

Kevin Walsh: Kevin Walsh. I'm looking in the study. I think it's page 116 that looked at the on-label use of BMP-2 in lumbar spine surgery, and it's the section on, I think it's the section on efficacy, and I think there was a statement that said that in the Burkus trial, it appeared that data from patients classified as failures were not reported after they had failed the treatment. Can you expand on that a little bit?

Robin Hashimoto: That's not an uncommon thing to do in these types of trials, that once a patient fails, then they stop reporting on their data. They'll be included the reoperation rates and whatnot, because that's typically what that implies. So, we considered that when we graded the level of evidence of the study. We consider it, that that's like a failure in the intention to treat. You see intention to treat analysis.

Craig Blackmore: Okay, the next phase is the discussion among the committee members using all our resources to try to work towards a decision. It's now 3:00. I want to give us about a 10-minute break, and then we'll resume and continue with that discussion. So, we'll reconvene at 3:15.

Craig Blackmore: Any other questions for our team?

Richard Philips: I have a question for our clinical expert. My impression, you know I occasionally do chart reviews and things of this. My impression is that most orthopedic surgeons or I should say spinal surgeons, not just orthopedic surgeons, but spinal surgeons are using BMP. Is that false? Or do you think it's divided in the community, and if so why? I don't understand with all the, you know, with the press that came out this last summer, etc. You know, I'm not sure the politics of it and where physicians stand on this.

Michael Lee: Well, in terms of how much BMP is actually being used, I think we actually have data on that, that was presented, right? Was it 10-15% by the agency data there? Was that correct, Dr. Mootz?

Bob Mootz: That's an underestimate. The agencies - the national estimates were 50-80%.

Man: There is a voluntary way that someone submitting a bill can just decide they're going to tell us they use BMP. So, we can tell that 15% of the bills that came in for spinal surgeon, they voluntarily included that information. It is not a billing requirement. The literature suggests that somewhere between 50 and 80% of spinal fusions use BMP.

Michael Lee: I would say that it's definitely fair to say that the prevalence of usage has been increasing.

Richard Phillips: Well, and I guess what I'm getting at is that most of these people aren't dummies, and I'd like to think that they look at some of this

literature, too. But I'm still trying to get a physician's viewpoint of why the utilization continues at a high level in face of this questionable evidence of safety. Can you give me - is there any census - does it really save that much time? Do patients get a good result? Do you have anything that you could throw light on, on that?

Michael Lee: I think it speaks to the potency of the product. I think it's universally agreed upon by all spine surgeons that this is a potent bone former and works very well to promote bone fusion in that it decreases the morbidity of an operation and increases, though the evidence does not suggest yet. I wonder if the evidence is lagging behind. Increasing fusion rates does decrease revision rates for nonunion and other complications. I suspect that is the reason why that the prevalence of usage has been increasing over the past several years.

Richard Phillips: So that issue that one of the physicians presented to us, it was the Burkus paper, whatever. I can't remember the name, where there was about a 5% difference in fusion improvement with BMP. Is that a fair statement, or is that not?

Michael Lee: I don't know if I could apply on that.

Craig Blackmore: Yeah, I'm going to break in. You know, we rely on the evidence vendor for the evidence, and we rely on the clinical expert for the context. I want to avoid having expert opinion interpreting evidence. That's why we have the evidence report. And I apologize, Dr. Lee. We have to sort of keep these boundaries to make sure that everything's evaluated fairly and our process is comprehensive and transparent.

Seth Schwartz: Let me try to make a point, and I want the opinion of my partners here. The way - the direction this discussion is while if using the product decreases frequency of revision surgery and the need for revision surgery as part of a supposed benefit, and I look back at the Burkus trial where the patients that had to undergo disk device removal, revision, or supplemental fixation were not reported after they'd failed the treatment. Am I right in concluding that if we look at the rates in these studies, we really don't know what the rates were, in fact, because, we're just looking at the reported rates?

Craig Blackmore: Which slide are you referring to? I'm not sure I'm with you.

Seth Schwartz: I'm on page 116, again, of the presentation.

Marie Brown: Are you saying because the people drop out at the - who were the treatment failures or the surgical failures are not followed up, and when you calculate the revision rate. Therefore, there is some biased to the revision rate.

Seth Schwartz: I'm just wondering if we're seeing the accurate revision rate, if those people aren't being counted.

Man: It doesn't talk about revision rates on this section.

Marie Brown: I vaguely remember some statement that the people who were noneffective in the first surgery were counted in the…

Craig Blackmore: Dr. Hashimoto, can you help us with the information on revision rates?

Bob Mootz: So, I just want to make sure I understand the question. If you're concerned about whether patients who have failed in what way they weren't counted, it's that they weren't counted for further followups on other outcomes, but they were counted as a failure and it depends on what - how the authors define failure. So, in the context of the Burkus study, part of what's defined as a failure is a revision. So, those revisions are counted.

Seth Schwartz: Okay.

Bob Mootz: Does that make sense? Yeah.

Seth Schwartz: Thank you.

Man: Dr. Lee, can I ask you just one question. Traditionally if you're doing this and you're billing for - or if you're performing a separate autograft from an iliac crest, is that a separate procedure that you would bill for, that if you use BMP you would not bill for?

Michael Lee: That's a question that - to harvest autologous iliac crest bone graft is a separate procedure that we would bill for. Correct. That is the procedure that would be billed for.

Man: So, if you were using BMP and you don't do that, then theoretically the operation would be cheaper from that standpoint?

Michael Lee: From a procedural standpoint, that is correct, from a procedural standpoint, not a total cost standpoint.

Man: I understand, thank you.

Man: Could I ask for some clarification on the complication of the retrograde ejaculation. I'm not sure I understand that as a complication if it's in a patient over the age of 40. Can you comment upon that, or the age groups of these people, and it's a transient thing in some people, too. Is that not correct?

Bob Mootz: Since I'm not over 40, I can't really comment on that.

Marie Brown: Well, they did note a proportion of the people at 12 months who still had it.

Man: Right, but what I'm saying is that I'm not - a lot of people have sterility procedures anyway. So, I'm not really sure I understand whether - if it really is that important. I guess that's what I'm trying to get at.

Bob Mootz: We don't have any data broken down by ages of age-specific information. Just a second.

Robin Hashimoto: I believe in the Carrigee report the average age of the patients, and this, of course, was the males, was 42 years, and I believe the 2 oldest patients were 2 of the 3 patients that had - still had retrograde ejaculation at 12 months.

Man: So it's in the range where people might be effected, then. Okay.

Craig Blackmore: Okay, let's try to get some structure around this. So, I'd like to try to use our tool to do that, and I think we've already heard that we're going to need to break this down. So, let's talk about that for a moment. I think there's one division that we can look at is on use versus off use and our data's been broken out that way, and then the over division is the BMP-2 versus the BMP-7, and it seems reasonable to me to try to split that out, as well. Does that resonate with everyone?

Seth Schwartz: Do we also want to look at separating cervical versus lumbar, or is that?

Chris Standaert: The cervical is an off-label, so it just depends on how we…

Man: So is some lumbar.

Chris Standaert: Yeah, but all the cervicals would be falling in off-label with some lumbars and whether there's a condition for on-label, off-label as a condition or a primary divider.

Craig Blackmore: Okay. I think we may need to do that. We'll discuss that, but let's kind of put that on hold. What I'm going to propose is that we start with the on-label use of BMP-2, as I'm thinking if we're going to approve any or all, we're going to approve that one. So, that'll be kind of our test case and then we'll work backwards from there. So, let's look specifically now at the use of BMP-2 for on-label use. So, this is lumbar and let's work through our tool. So, the first thing is again, delineating the outcomes of interest, the safety outcomes, the efficacy, effectiveness, and the outcomes. So first, Josh and his team have prepopulated the tool with safety outcomes that are of interest. So, we need to sort of run through this and determine if we've got everything that we think is important to cover. If there's anything that we don't think is important, etc. So, if I could get the committee to turn to page 4 on the tool. Obviously, mortality is important. We're going to look at that. Overgrowth and uncontrolled bone formation, which could be promoted. Surgical complications, which could apply either to the lumbar surgery itself or the graft site in the case of the iliac crest graft, reoperations. Now is a reoperation a safety or is that an effectiveness?

Group: Both.

Craig Blackmore: Kinda both?

Chris Standaert: It's even a cost.

Craig Blackmore: All right. So we'll add that to effectiveness and cost. Okay, wound infections, infections from a hematoma, dysphagia, cervical. So, we start with lumbar, and we won't look at that so much. Retrograde ejaculation we've talked about, and bowel obstruction, urinary retention, radiculitis, dural injury, neurologic events, and a body response, a laundry list of things for which we,

well, a lot of these we don't really have a lot of information. They are rare complications. Cancer we have heard something about. Reoperation/revision, which again I think is sort of an effectiveness so we might include that in both, revisions. And then graft site morbidity, obviously, for the iliac crest. Are there any other safety outcomes that the committee feels are important for the deliberation? Okay, then there's outcomes related to efficacy and/or effectiveness and these are the potential benefits of using the BMP. One of them is you might decrease the operative time, particularly around not having to do this additional grafting procedure. Similarly, blood loss, length of stay in the hospital, potentially might be effected. Fusion rates, as well as the need for reoperation or revision might be considered outcomes. Then there are the more sort of functional measures like the Oswestry Index, the SF-36 patient satisfaction measures, measures of neurologic status, and then there's functional in the sort of pure form. Are they back at work? Combined measures of overall success and then medication use or the SF-36 mental health sections listed. Are there other aspects of effectiveness that the committee is including in the deliberation? Okay. So, I'm going to invite any further discussion around safety and effectiveness outcomes before we move towards our nonvoting status, and I'm also going to - I like having somebody on the committee summarize where we are and sort of frame things, as a starting point for a discussion. So, I wonder if I could get one of the committee members to frame where we are in terms of not how they're going to vote but where the trade-offs are, what we know, what we don't know, etc. I'm going to pass out the coffee pot if I don't get some response from my committee here. Come on Michelle, where are we? You're fresh.

Michelle Simon: I think we have quite a bit of evidence that off-label use, both cervical and lumbar, is not the direction we want to go. There seems to be some safety problems there. I think safety is the biggest issue in this particular topic. We don't have much evidence for problems in safety in on-label lumbar use. It seems to occur in off-label uses also. What else? I think there's not a lot of evidence for cost. There is some suggestion that the costs are perhaps better for the BMP procedure, but then studies are small, and one of them is a foreign study, as well. It wasn't even done in the United States, so it's hard to really make that relationship. So, I think the data is not super strong on cost either.

Craig Blackmore: Okay. So drilling down specifically in the on-label use, we're basically balancing these safety concerns, which seem to be a little - they seem to be less in the on-label use, although they certainly exist, and cost concerns are hard to really figure out because of the complexities of whose paying and short-term and long-term and then there's outcomes and do we think these are equivalent, or do we think BMP or non-BMP is superior in terms of outcomes?

Chris Standaert: I'd like to ask our question back to the vendor, because I think what both of you just said is true, and I think in the on-label use safety concerns seem to be less than they are off-label. It certainly seems to fuse people, and you wonder where the benefit is. The benefit would be in avoiding the other surgery for the donor graft and having higher rates of fusion. So, when we get asked these questions, the vote is sort of in any or all populations is there evidence of benefit or whatever the wording is. So, if there are lower revision rates, or if there are higher fusion rates in people who are high risk for nonunion, people who are diabetic, people that are medically compromised, people who smoke, all these sorts of things that mean in those people this might be a better product, because they'd have a more successful outcome from a surgery they need to have. Then, there's some utility to it. You know what I mean? So, is there data on those populations that this actually - or all these only studies on healthy people?

Robin Hashimoto: Well that really gets at key question 4 on the differential effectiveness and safety of subpopulations. So, there was a little bit of data but in general, we didn't find any strong evidence that there was any difference in any subpopulations, including smokers and based on age, number of levels, complexity, whether previous surgeries had been done. So, unfortunately, there's not a good amount of evidence to answer that question right now.

Chris Standaert: It's just insufficient. It wasn't really one way or the other. It sort of.

Robin Hashimoto: Right. I mean, our conclusion, the strength of evidence was insufficient, so.

Chris Standaert: Because theoretically you would think that people with bad bone, people with medical issues, people with prior surgeries, people who had already had a failed fusion, people who smoked.

Robin Hashimoto: Yeah.

Chris Standaert: Okay.

Craig Blackmore: Okay. Any further discussion. Anybody want to build on what Michelle has given us, or disagree or agree? I need vocalizations from the committee so I know we're ready to move on. Or else, I'll assume silence means we're ready to move on.

Seth Schwartz: I think, I don't see much outcome benefit to using this product. I'm reassured that if it's used in an on-label fashion, there's no increased cancer risk. The other risk factors seem to be fairly minor. So, I understand the push of the agency to focus more on safety. So, what I'm seeing is a difference between on-label and off-label use of this product, mostly from a safety point of view.

Carson Odegard: You know, I'm looking at the carrier decisions. It's like they all came out of the same cookbook. They're all pretty close to each other. So, they're either all looking at the same data or they, it's really close.

Craig Blackmore: Did they ask the same experts?

Carson Odegard: I'm just wondering, you know, as a committee, I don't know. I just think we need to talk about BMP-7 and its value, or its effectiveness and its safety, and try to eliminate as much of that as we can.

Seth Schwartz: I understand your point, but what's the distinction in your mind between 7 and 2?

Carson Odegard: The only use that I can see is that for those patients that can't or don't qualify for an allograft, and that they would be - it would be used for those patients that can't tolerate it, either through immunological means or they just. So, that's why I'd like to talk about 7 and see what the committee thinks. And if it's not being used now, and it's.

Chris Standaert: It looks like we don't have as much data about 7.

Carson Odegard: We don't have much about it.

Chris Standaert: And the only indication for 7 was a compassionate use indication, so that probably means it came out before there was much data on it, right? They said well BMP-2 works, and the only indication we have is it is for revision posterolateral lumbar fusion.

Carson Odegard: Right.

Chris Standaert: So, it's not put in a cage. It's not going in the front. They go in the back, and they revise posterior fusion with it, and that's the only indication they have for it, and we heard a couple of issues about uncontained sort of use of this stuff and bone growth and inflammatory issues and controlling dose and all that. That's why I think some of those data don't - the data I saw don't look as good for the BMP-7 as the 2.

Carson Odegard: Mm-hm.

Craig Blackmore: I can see a role for one or both of these products in the sort of failed spine. You know, there's just a subset of individuals who have had their surgery and it hasn't worked, and it's a mess, and it seems to me that that's a pretty desperate situation and it might be appropriate to try something extreme in that circumstance. So, for me, I think I can see sort of like what they have here as a, what do they call it, compassionate use. Having the under extreme circumstances this might be okay, but that is very different from 80% of the time when we do spine surgery, particularly referring to the 7. Again, I'm not speaking to the on-label and some of this other stuff, but I think personally if we're to shift our discussion to 7, I could see making that sort of case for it, but based on our absence of any other data that it's worth it, the risk, I would struggle. I think I'm sort of saying the same you are, right?

Carson Odegard: Right. I mean, I've just got this 7 in the back of my mind, and I start thinking about 2, then maybe I should just forget about it.

Chris Standaert: We don't even have a slide on on-label use of BMP-7.

Craig Blackmore: There is no on-label use.

Carson Odegard: No, there isn't any on-label.

Chris Standaert: Well, there's no study on on-label use.

Craig Blackmore: No. There's no on-label use. It's not approved on label.

Chris Standaert: So, the approved indications - but it's approved.

Marie Brown: 2 is the only one that's approved.

Robin Hashimoto: No, BMP-7 is approved.

Chris Standaert: Yeah, BMP-7 is approved, but there are no studies on the indications.

Robin Hashimoto: That's correct. The studies that were used in the FDA report are actually considered off-label uses because they were primary fusions. They were not revision fusions. They were not compromised patients.

Chris Standaert: So even after that, it only got approved for the primary revision - for the revision fusion even though it was not studied under those circumstances.

Robin Hashimoto: And the rational for that wasn't very clear in the report.

Chris Standaert: I think that would follow, yes.

Robin Hashimoto: Yes.

Chris Standaert: Because we have no (unintelligible) only indication.

Craig Blackmore: So back to procedure. Is it the committee's preference to proceed from BMP-2 approved to sort of one extreme to BMP-7 or is it our preference to start with the BMP-7 extreme and proceed to the BMP-2 approved extreme? Carson wants us to address the 7.

Chris Standaert: I say you start with 2. That's my vote.

Craig Blackmore: All right. Well we'll go back to starting with 2 for approved indications. So, we talked about outcomes and whether they're equivalent and whether there's a real outcome benefit here, and we've talked about cost, and we've talked about safety. Does anybody have any further comment to make on BMP-2 lumbar-approved indications? Because if not, I want to go to a nonbinding vote and get a sense of where we are.

Marie Brown: I just, in terms of the agency directors, one of the things that could really drive cost is if the company keeps escalating the cost of this material and then that drives up the cost of the bundled package. Does that - is that a possibility or am I just worrying unnecessarily?

Bob Mootz: Again, I think there's tradeoffs in costs, because theoretically if you have reduced OR time, that saves some money, and I'm not sure that there's actually reduced hospital time or hospital stay. Our data actually has a slightly higher hospital stay in the 14% that had the BMP, and so the cost issues, I think, are - they could be a wash. It could be slightly higher. It could be an upward cost driver over the longterm, but direct right now, the way our system is set up, cost isn't a primary concern for us.

Craig Blackmore: Any other comments, questions? I don't want to cut us off prematurely. Okay, then let's go look at things, the nonbinding vote, and this will give us a feel for where we are. So, this is going to be a decision onto whether the committee members think that the use of rhBMP-2 in the lumbar spine for approved indications is more, less, equivalent, or unproven when compared to fusion performed without the use of this agent.

Man: Compared to autograft?

Craig Blackmore: Right, so compared to autograft, which is usually going to be iliac crest, although it wouldn't necessarily have to be. Right, so the first vote will be whether we think BMP is more, less, unproven effectiveness for strictly on-label lumbar indications.

Josh Morse: 7 equivalent, 1 more.

Craig Blackmore: And then the second component of this, do we think it is as safe as the use of crest, or as the use of autograft?

Josh Morse: 3 equivalent, 4 less, 1 more.

Craig Blackmore: And then finally, we want to know if it's cost effective.

Josh Morse: 8 unproven.

Craig Blackmore: So, it sounds like we're pretty much, there's a lot of agreement around equivalent. There's a lot of agreement that we don't know anything about cost effectiveness, and on safety, we're in

the kind of either equivalent or it's less safe camp. So, it sounds like the decision for the committee members on this topic is going to be around whether it's worth it. If we do believe, those of us who do believe that it's a little less safe, is it worth it? Anybody want to comment, or should we proceed? Okay, so we're going to conduct this as separate votes, so this is semantics. Instead of saying we're voting for the use of BMP under conditions and one of the conditions would be approved indications of the lumbar spine, what we're going to do is have a specific vote for approved, FDA-approved indications, in the lumbar spine. Is that going to work for you, Josh? I can reformulate it later if it doesn't.

Josh Morse: I think that would work. Say it again.

Craig Blackmore: So, instead of having a decision document that says we approve, if we go this way, but we approve BMP under certain conditions, and one of the conditions is that it has to be an FDA-approved indication of the lumbar spine. This decision would be for FDA-approved indications in the lumbar spine, we vote to cover or not cover. Does that make sense?

Josh Morse: I think that makes sense.

Chris Standaert: My only problem with that is the FDA indication says patient's DDD plus or minus grade 1 spondy and failed greater than 6 months of nonoperative care, which don't necessarily match the indications for fusion, that this committee already decided awhile ago. So I think we could say, just simply, are we voting on the first sentence of their indication, which is primary anterior open or laparoscopic fusion at 1 level between L4 and S1 and leave it there and leave out the second part of the FDA indication, because that will conflict with what this committee already decided in terms of the agencies approval of fusion in general. So, we're just using the first sentence of that, and we're saying under - do we cover as a condition primary anterior open or laparoscopic fusion at one level between L4 and S1? Just to be clear.

Carson Odegard: I would agree with that, too.

Craig Blackmore: I think maybe another way to look at this is to take a straw vote and then lump it all together in the end and define it that way so that in the end we would have our formal vote be on everything once, and the conditions will be number 1, approved indication

that we decided on a couple years ago, and then depending on what we decide. So, in that context, I want to entertain straw, a show of hands, nonbinding votes on where the committee is in terms of whether we should cover BMP-2 for FDA-approved indications of the lumbar spine. So, if I can nonbinding just show your hands if you think this should be something we include. Okay. So, I'm seeing - I'm not counting, this is unofficial - but I'm seeing that downstream we're going to need that. We're going to write that down. It looks like the committee is yes on FDA L-spine. Okay, so then the next question is, how about non-approved indications. We're still talking BMP-2, and this could be, we can break it down into lumbar, or we can lump lumbar and cervical together. Where do we stand on this, and I'm welcoming discussion? We're not voting at this point. So we're not using the cage, we're just using the product. Seth do you want to take a stab?

Seth Schwartz: Yeah, I'll take a little stab at this. I think it's a little difficult to tease out the efficacy data here, but I think what we're seeing is that there's not a dramatic difference. Essentially, if you're looking at non-inferiority, I think you could say that BMP is not non-inferior in these settings. So, the efficacy is going to be fairly similar in these settings, and the question really becomes a safety issue. I think when we look in the cervical spine, I think the safety data is a lot more compelling that it's a bad idea, and I think we've kind of heard that from our expert - or from the people that testified for us that that's pretty much generally viewed upon as correct. I think when you look at the lumbar spine, the nonapproved indications, I think, probably have a slightly higher safety concern than an approved one. So, I think that's where we get into the situation, and that may be where we can think about conditions, such as the more complex patients revision spines. People where there might be a more compelling reason to use it.

Craig Blackmore: Okay. Any other comments? So what I'm taking from Seth, first of all, is that we should talk about off-label lumbar. We shouldn't try to lump lumbar and cervical, and starting with that, I need people's input on the use of BMP for off-label lumbar indications. Can I get a straw from (unintelligible) even?

Chris Standaert: One of the problems with the off-label stuff is the cancer data, and I know it's not significant, but it almost looks like there's a dose relationship here, and so if you go off-label lumbar and you don't restrict dosing, which means multilevel fusions, which

means lots of stuff in the back of the spine for a posterolateral fusion, which means a multilevel posterior fusions or whatever. You get out of the range where the FDA data is supported, and the numbers start to go up, and it looks more troublesome. I agree, the cervical stuff is a different story. Then, the problem becomes where in the off-label use do you find benefit when we don't - again, the 1 FDA indication for BMP-7 they didn't study, which is sort of odd. Those are the people you'd think about doing it, people with a failed fusion of some sort, but the data they had was really insufficient for all these special subpopulations, the smokers, the failed fusions, and so it's hard to draw the line. It just sort of -we'd be doing what probably the FDA did, saying these sound like people who might need help. They might need something else to make them fuse, because they are at such high risk so try this, even though the risks are there without the data.

Craig Blackmore: So, we could approve it for all lumbar indications. We could not approve it for indications that are off-label, or we could try to define conditions around what our decision might be.

Chris Standaert: We can start putting up conditions and just start with that one from the FDA and see if we want to add another condition of somewhere in the lumbar spine, and if people are in agreement that we're not going to go for cervical spine. We just leave that off as not one of our conditions. Then see whether, as Seth said, is there some other place we think this might be applicable beyond the pure FDA indication for an anterior interbody fusion?

Craig Blackmore: That sounds useful. Margaret, can we get you to pull up a blank sheet? So, one potential area, which has been identified, is paralleling the BMP-7 compassionate use, whatever the terminology is, guidelines for cervical, and that is circumstances in which there's been failure of prior fusion. I don't have the other ones in front of me. Do you have the other ones? What are the other BMP-7 cervical indications as a starting point?

Marie Brown: It's injury. Isn't it back injury?

Craig Blackmore: So from the BMP-7 lumbar indications on here, revision posterolateral lumbar fusion, compromised patients for whom autologous bone and bone marrow harvest are not feasible or not expected to result in fusion.

Chris Standaert: Do you want that up as an indication, or do you want some discussion?

Craig Blackmore: So, I don't want the FDA indications up there. I don't want the…

Chris Standaert: But that's the first - that's one of the indications we have been talking about, yeah? Are we on BMP-2 still?

Craig Blackmore: Yeah, but we were - we thought we were okay on the FDA-approved indications.

Chris Standaert: Right, but these are going to be…

Craig Blackmore: All right, so we can add…

Chris Standaert: …and list all of our indications…

Craig Blackmore: That's fine.

Chris Standaert: And then we have one vote on BMP-2.

Craig Blackmore: Okay, so we'll start with those and then on page 5, the other ones we're talking about are the ones that have been applied previously to the BMP-7 and that's for revision posterolateral lumbar fusion or patients who are compromised for whom autologous bone and bone marrow harvest are not feasible or not expected to result in fusion. So I need input from the committee.

Chris Standaert: I guess I can see why they said that, because that sounds reasonable. It's just lacking in data. You know what I mean?

Group: Yep.

Chris Standaert: There's no data.

Seth Schwartz: Can we ask our clinical expert, what do you - what else would you do in that situation? What other options are there for these patients?

Michael Lee: So, for patients who are at risk for nonfusion?

Chris Standaert: I feel like the people for whom BMP-7 is indicated, revision posterolateral fusion, or people where you can't harvest bone or you just don't think autologous bone is going to work.

Michael Lee: Other options could include allograft, which is not as effective as autologous iliac crest bone graft, bone graft extenders, which again are not as effective as autologous iliac crest bone graft, harvesting bone graft other places outside the iliac crest, though that's not really done typically. There really is a challenging situation to achieve fusion in patients who are at risk for nonfusion for whatever reason.

Seth Schwartz: Would you say that this is a circumstance where clinically BMP is standard of care at this point?

Michael Lee: I don't know if I'd use the word standard of care. I think this is a situation for, if I can apply, or many surgeons would consider this a useful situation for Infuse BMP to augment the likelihood for challenge and reduce the biological challenge of the patient to mount that fusion.

Seth Schwartz: Thank you.

Craig Blackmore: So, I think what the FDA is trying to get at here, for the rhBMP-7 indications and I think what we're trying to get at is to somehow capture the challenging patient who is at higher risk of nonunion, right?

Seth Schwartz: For whom there's no evidence.

Craig Blackmore: There's not going to be any evidence. So, we sort of, we're there already. We know there's no evidence, but I think that's what - so maybe we take in the big picture at this point in saying if we can define a set of indications that seems to capture that group, would the committee view that favorably for coverage? So, if we can craft a set of indications that put into words the high risk for nonunion group, is that a group that we would view favorably for coverage?

Seth Schwartz: In the lumbar spine only.

Craig Blackmore: In the lumbar spine off-label. Carson?

Carson Odegard: I'm not sure, yet. I'd - you're lumping both products together?

Craig Blackmore: No, I'm doing 2, because I don't want to go to 7 just yet.

Carson Odegard: Okay, because 7's obviously got a high- a higher cancer risk; 2 is pretty equivalent. Off-label cancer risks about 3 times higher.

Craig Blackmore: So, if we define patients of high risk lumbar spine for nonunion, and we're talking about BMP-2, would you - nonbinding?

Carson Odegard: I'd go along with that.

Craig Blackmore: Michelle, where are you on this?

Michelle Simon: I'd be closer.

Craig Blackmore: You'd be closer. So, maybe?

Michelle Simon: Maybe.

Craig Blackmore: Yeah. Richard, nonbinding, where's your mind on it at this point?

Richard Phillips: Well, I'm not sure. I mean, I sort of made my mind up where I really have my concerns. I only have one concern, and that's about the cancer. Other than that, I have no problems with the use of it. Even at that, there's no statistical significance about the cancer. It's just a trend in multiple studies. So, and I don't know how to deal with that. I'm really having a lot of trouble with that, and it only occurs in the lumbar. So, that's when you're talking about this, the lumbar, I sort of have reticence about moving ahead and approving that, because from my perspective that's the only thing that really bothers me about the utilization of this at this point.

Craig Blackmore: Okay. So we're up in the air on that one. So, let's move onto the cervical. Where is the committee in terms of the use of BMP-2 in the cervical spine? This is off-label, obviously. I've heard some concerns over here. Are there committee members who would offer the opposite viewpoint that we should be supporting this in the cervical spine, or not?

Marie Brown: No.

Carson Odegard: I don't think there's any evidence.

Michelle Simon: The evidence is pretty insufficient. The evidence that we do have on that.

Chris Standaert: There are some significant problems, too. Yeah.

Craig Blackmore: So, at this point, is the committee on board with the limitation that would read lumbar only?

Group: Mm-hm. Yes.

Craig Blackmore: So, that's easy. And then in terms of the BMP-7, does the committee believe that BMP-7 should be used in the equivalent manner as the off-label BMP-2, or not at all, or with less restriction?

Marie Brown: Where were you seeing the 3 times risk for 7 versus 2 carcinoma?

Richard Phillips: For off-label, I don't know what study it is, but comparing it at 24 months, 3.8% versus 0.9% and 6.3% versus 2.2% at 60 months. So, there's about a 3 times.

Craig Blackmore: So, you're looking at the slide on page 17.

Richard Phillips: Yeah. Actually, it's in the report on page 221. And then when you get to 7, it goes up to about 5, almost 6 times in a 12-month period risk of cancer.

Robin Hashimoto: Can I just clarify this, on that table, is that okay? So, for the BMP-2 use at 24 and 60 months, so that's the Amplify report, so that's one study. BMP-7, there's 2 different studies, and actually the first data - the first study there where it says 18 and 16, that's actually the number of patients. That was labeled incorrectly, so the percentages are actually in the last column, so 5.6 and 0%, and then for the next study, it was 12.5% and 8.3%. So, I think it's important to keep in mind that both of those studies were small with 34 and 36 patients total for the BMP-7.

Richard Phillips: Okay. Mm-hm.

Seth Schwartz: This is for our vendor. Are there any studies that compare BMP-7 to BMP-2, or has that not been done?

Robin Hashimoto: That's not something that we looked at.

Chris Standaert: I think my worry with BMP-7 is that it's not actually FDA-approved for anything that it was studied on, and the numbers are a lot smaller in the BMP-7 studies, and it hasn't really met a bar where

the FDA has decided to approve it for what they - it's not approved interbody fusion. It's approved on a compassionate use basis for things that might go badly otherwise.

Michelle Simon: For which it may not be used, either.

Chris Standaert: So, I have trouble with that one. It's totally unconstrained. It's a putty. It's not a sponge or in a cage or whatever. It doesn't have a delivery device.

Craig Blackmore: So, I think one perspective I'm hearing is that there is insufficient evidence to support BMP-7 at all. The kind of leading argument would be that there's potential value, and we might use it in a sort of a compassionate use basis, as the FDA has defined it.

Group: Mm-hm.

Craig Blackmore: Nonbinding vote to include this as a condition or are we mostly on board with limiting our potential approval to BMP-2.

Marie Brown: Yes.

Chris Standaert: Can I ask Dr. Lee a question.

Craig Blackmore: Sure.

Chris Standaert: So, you said you don't use BMP-7, but if you're doing a posterolateral fusion where you can use sort of an unconstrained delivery of BMP. Say you had a posterolateral fusion in a revision complex patient, and you thought you should use BMP to improve your odds, is there a difference between using BMP-2, which is sort of a sponge-like delivery system, as opposed to a BMP-7, which is really a putty or a different sort of whole texture. I assume you intermingle with bone, morcellized bone and all that sort of stuff. Is there any advantage to having the BMP-7 as a tool in that context? Or would you just sort of apply the BMP-2 sponge in the midst of your bone graft?

Michael Lee: I don't know that there's an advantage or disadvantage of the putty, the putty form versus the sponge carrier form.

Chris Standaert: Just from the technical performance from the procedure aspect, yeah.

Michael Lee: I don't know that there's any advantage or disadvantage. With the BMP-2, the sponge form, it's usually - sometimes you can make it a little egg roll with the bone graft inside the sponge and put it out in the posterolateral gutter, and it essentially forms a similar shape grafting construct. So, I don't know if there's much disadvantage or advantage.

Chris Standaert: Okay. Thank you.

Craig Blackmore: So, in a formal show of hands, one condition that's being proposed here is that the condition be that approval only be the BMP-2. Is that something that the committee favors? You can nod. You don't have to show hands. I'm seeing nods. Okay. So, let's put that in here that the condition is BMP-2 only.

Seth Schwartz: Can we just separate it out and just have one vote on BMP-7 and one vote on BMP-2 and the BMP-7 we'll just say not cover. Is that an option?

Craig Blackmore: Well, we could. We could do that. Is that cleaner? Okay.

Chris Standaert: We have a dangling closet there.

Craig Blackmore: Actually, you know, we're doing well. So, okay, I'm going to take Seth's advice. I'm going to go through the process for BMP-7, and we're going to return to our BMP-2 discussion, which I think we're well on our way towards. So, first we have nonbinding.

Man: What was our last official nonbinding? That was lumbar?

Man: It was BMP-2 for (unintelligible) indications.

Craig Blackmore: Right. Okay. So, this is nonbinding. This is for BMP-7 for any indication. So, the way, this is nonbinding, and the way it's written here is, is there sufficient evidence under some or any situations that the technology is effective, safe, and cost effective. So, committee members, you should vote more effective if you believe that under any circumstance the BMP-7 is more effective than fusion with graft. So, we're not comparing to BMP-2. We're comparing to graft and ditto with safety and effectiveness. Is that clear? Okay. So, first is the effectiveness.

Josh Morse: 6 unproven, 1 less, and 1 more.

Craig Blackmore: Okay, and then the safety piece.

Josh Morse: 8 less.

Craig Blackmore: Okay, and then cost effectiveness.

Josh Morse: 8 unproven.

Craig Blackmore: Okay, further discussion, or shall we proceed to binding vote?

Marie Brown: Vote.

Man: So, I just need to be clear. So that was BMP-7 for any application?

Craig Blackmore: For any application. So the following vote is, will BMP-7 be approved, and our decisions are that we will either choose to cover, which means it's covered unconditionally, to not cover, meaning it's not covered for any indication, or to cover with conditions, which we would then define.

Josh Morse: 8 no cover.

Craig Blackmore: Okay, and having made a binding vote, it is our responsibility to see how that corresponds to Medicare National Coverage Decision and reconcile if we disagree, and there's no National Coverage Decision. Correct, Josh?

Josh Morse: Yes.

Craig Blackmore: So, there's no National Coverage Decision, so we are set. Okay. Now we return to the BMP-2 discussion and our list of potential conditions, and we have the FDA indications, and we have lumbar only as a condition, and we are now reconciling the lumbar off-label indications and whether or not we will cover the use of BMP-2 under those conditions.

Chris Standaert: I think we removed the words off-label because we have - we're just doing one vote, right, for just lumbar spine BMP-2?

Craig Blackmore: Right. So what I want to address now is lumbar spine BMP-2 with the exception of those 2 indications, which correspond to the on-label but, yes. We're not going to be using that, but that's…

Marie Brown: So, take that off.

Chris Standaert: Take those 2 words out.

Craig Blackmore: Actually, take off lumbar spine, leave it as BMP-2, and then list as a condition lumbar spine only. Then, just for point of clarity, we've only looked at adults, right? So, this will be, this decision is only addressing individuals of ages 18 and over. Okay. So…

Man: Is that the FDA or is that (unintelligible).

Craig Blackmore: I think that's the committee's decision.

Marie Brown: It just says adults.

Man: (Unintelligible).

Craig Blackmore: That's with the FDA, but we're going with adults. We're going with 18 and older. Okay. So, we have to make a decision on lumbar indications other than those 2 that are listed. So, one choice is that we would cover this for all lumbar indications, basically all lumbar indications. Again, if I can get a sense of the committee with head nods or hand raises, who would be more comfortable covering all lumbar use of this procedure? Is anybody in that place? Okay. Then I need somebody to take a shot at - well, I guess I'll ask the opposite question. With nods or hand waves, is there anybody who believes we should have additional indications to the 2 that are listed up here, which are the ones that are derived from the FDA? Do we believe this is sufficient, or do we believe there are other appropriate indications that should be covered. So, I guess nod or give me your hand if you think we should be adding to the list of approved indication.

Marie Brown: Are we still keeping nonfusion?

Chris Standaert: We have that last - that's not a sentence. I'm not sure where that. So are you referring to that or no?

Craig Blackmore: We need to finish that first. Pull up the FDA. Okay. Can you add the words, Margaret. Do you know what we're talking about?

Margaret Dennis: No.

Craig Blackmore: I can't understand how we could have confused anybody (unintelligible).

Michelle Simon: We're giving up on conservative care…

Craig Blackmore: I'm sorry.

Michelle Simon: We're giving up on conservative care lingo, right? Did we already decide that (unintelligible) earlier?

Group: Yeah, right.

Man: We aren't making the decision on whether they have the (unintelligible) or not.

Craig Blackmore: We can add - this is the point where we would add that in the lumbar.

Marie Brown: You mean add having failed conservative care?

Group: No.

Man: Add the compromised patients from autologous (unintelligible).

Craig Blackmore: You were talking having failed conservative care as an indication for surgery or having failed surgery already as an indication for BMP?

Michelle Simon: No, an indication for BMP.

Chris Standaert: So, a revision surgery.

Craig Blackmore: So, that would be a proposed additional approved indication, which that's where we are is deciding that.

Chris Standaert: Question for the vendor.

Craig Blackmore: Oh that's in there already. Yes, it is.

Chris Standaert: Quick question for the vendor? I had some, again I had a sense when looking at the cancer data that it was almost like a dose response curve? Is that just a way that the data was presented that you see that, or does there seem to be like an effective increased risk with increased dose? Because that gets into, you

know, nonfusion - if it's not just a revision but it's a primary fusion, you just think it's a low-risk to fuse. I mean, again, you get into five-level fusions with lots of BMP, and if there's a problem with high doses of BMP, and that's what the problem is, maybe that shouldn't happen. Maybe it really should be a revision thing or for one-level fusions that are challenging. I don't know. That's why I'm curious about that. Is there a dose - does it seem that way or no?

Robin Hashimoto: If you're asking my opinion, you could definitely interpret the data that way. I don't think that you could say - I don't think there's enough evidence to say conclusively that it is dose-responsive, a dose-response effect, but that's one possibility.

Craig Blackmore: Okay. Now we have our list completed. So, these are the FDA indications.

Richard Phillips: Didn't you ask the question did anybody else think anything that should be added?

Craig Blackmore: Yeah, so that's where we are. So, we didn't have the list complete before. Now we do, and now we're asking the question is there anything else that should be added.

Richard Phillips: I personally don't think there's a problem with the cervicals, within - I shouldn't say that totally. I don't have a problem with some of the cervical indications, personally, but I suspect I'm in the minority here.

Craig Blackmore: Okay.

Marie Brown: The cancer risk is higher.

Richard Phillips: I'm sorry.

Marie Brown: The cancer risk is higher.

Craig Blackmore: I mean, we took a straw poll, and it seemed to be that most of the committee were not on board with the cervical.

Richard Phillips: Yeah, that's what I said. I didn't hear that specifically, but I.

Craig Blackmore: And we haven't officially voted.

Richard Phillips: I sense that from what's been said.

Craig Blackmore: So, that's why it's up there, because that's the. So, in terms of the lumbar, have we covered what we believe are the appropriate indications or conditions on this list, or do we believe?

Chris Standaert: I just want to clarify the wording. So this says patients at risk for nonfusion due to revision posterolateral lumbar fusion or comprised patients. This looks like 2 different groups of patients. As I read the FDA thing, what it really seems to say are revision posterolateral lumbar fusion in compromised patients for whom. It's revision patients who have these risk factors. It's not - so do we want all revision patients can have it? And all compromised patients can have it? Or do we want revision patients who are compromised.

Craig Blackmore: I'm happy with both, personally.

Chris Standaert: I'm just throwing that out, because I'm just trying to - we have to be reading this the same way. Is that - right now it's worded as 2 different groups.

Craig Blackmore: But they're both high-risk groups.

Chris Standaert: They are high-risk groups.

Man: For whom there's no evidence.

Craig Blackmore: For whom there's no evidence.

Marie Brown: Mm-hm.

Chris Standaert: But in the second, you're allowing it in any application, any method, PLIFs, XLIFs, ALIFs, posterolateral, all that stuff, because you're not defining the delivery mechanism.

Seth Schwartz: So, just to be clear, you're taking the indication from putty with BMP-7, for which we've decided we're not going to cover, and you're applying it to BMP-2.

Craig Blackmore: As a high-risk group.

Seth Schwartz: For whom there's no evidence in the material we've looked at.

Craig Blackmore: For whom the evidence is very limited, yeah.

Seth Schwartz: Okay. I just want to be clear on what we're voting on.

Craig Blackmore: We're not - we're not, yes. It hasn't been studied as a group. The data is slim, but it's a way of saying these are the patients in whom the potential benefit of having more bone growth is greater, because they are at higher risk of nonunion or other complications, and that's what the attempt is to capture.

Seth Schwartz: I understand it. I just want to…

Craig Blackmore: Can you reflect on it.

Seth Schwartz: I tend to look at the data about the amount of BMP that's used in the cancer risk as being strongly suggestive, and so without any evidence that this group would benefit, all I see is risk. So, that's my interpretation. I understand the supposition, but that's not - suppositions haven't floated us very far down the river for most of the other decisions we've made.

Marie Brown: Right.

Craig Blackmore: It's, you know, we're in that space of where do we draw a line and you don't have randomized clinical trials for every single subset of something, and I think the committee has to make a decision where we're going to draw the line, and this is one proposal and we don't have to - we can go back on it. I'm just trying to get our thoughts down so we have something to make a decision about. Maybe the best way is to have, we'll draft a couple of different potential conditions, because that's the direction we're headed, although we haven't voted on that, and committee members will vote on which group of conditions they are most comfortable with.

Seth Schwartz: Okay.

Craig Blackmore: Does that?

Seth Schwartz: Well, I defer to your procedural decision.

Craig Blackmore: Okay.

Chris Standaert: Question for Dr. Lee. Microvision, so in a revision, so if - I'm in my head trying to think about is this 1 group or 2 groups. So, we're talking about allowing it in people who you need a revision fusion on because the primary fusion failed, and either you can't go get an iliac crest for some reason or you just really don't think this is going to work, because it failed once already and you wanted to do something else, so a revision fusion. My question, so, if it's that group, so are all revision fusions posterolateral or do you have some other way of doing a revision fusion that would be reasonable where you would use this product that wouldn't be posterolateral? So, do we need that word in there or not? And then the other - so I'm trying to sort that one group, anyway. Answer that - we'll start with that question.

Michael Lee: So, fusion can be done either posterolaterally, it can be done as an interbody fusion from a posterior or transverse approach. It can be done as an interbody fusion from a lateral approach, as utilized with the XLIF or can be done from an anterior approach interbody, which appears that's already been decided upon. So, to answer your question, revision fusion can occur posterolaterally or interbody from a posterior approach, as well. So, there are more options of revision fusion than just posterolateral.

Chris Standaert: So you would do - you can do a revision interbody fusion?

Michael Lee: Yes.

Chris Standaert: Okay. So then, the question would be do we allow this for every approach? So, again now then we have revision patients. Do you let any approach you need to do in a revision patient to fuse them who is compromised? Or again, is this any sort of revision patient, you can use it, or plus any patient where you just can't get bone or think it's not going to work, which is sort of ill-defined? Leaving this as 2 groups is sort of ill-defined, that last one.

Marie Brown: It's more broad.

Craig Blackmore: So, I need the members of the committee to give me input here. Do we want to include all revisions or the subset of revisions?

Michelle Simon: I think if we're wandering in the field of no data, I guess I would be more inclined to stick at least with the humanitarian exemption, as worded. That's at least something, some guideline,

and that would be compromised patients in whom revision posterolateral lumbar fusion is happening, or needs to happen.

Craig Blackmore: Okay, are there other opinions? Anybody? I like it better separate, personally.

Chris Standaert: So, if you like it better separate, how would you go about (unintelligible) that last…?

Craig Blackmore: My feeling is, the reason to do this is because you want to grow bone, and if it's going to be - there is obviously risk and we're saying, I think, that we think that in some cases the risk might exceed the benefits, or sorry, the benefits might exceed the risks, but it's really hard to tell where that group is. So, my take on it would be to try to define that group but be fairly generous recognizing that even being generous you're going to be causing a much more of a restriction on the use of this than currently exists. That's a take on it. I'm not - I'm one voice. That's my…

Chris Standaert:: I guess I'm - the last words, are not expected to result in fusion, again very vague. So, you say well I'm going to do 5 levels, and I don't think it will fuse it. I'm going to put this stuff in, because I can't get that much off their hip, you know.

Craig Blackmore: I feel like our ability to micromanage is limited. If we think it has enough value that we're going to approve it under some conditions, I feel like we should not try to second guess the decisions. I mean, if we say we don't think it works, then that's great, but my bias is to not try to micromanage if we think it's useful in these high-risk people. Again, that's one opinion. I'm not.

Seth Schwartz: I have one other thought. I'm just kind of going back to the local decision on the sleep apnea situation, which is that they require that the surgeon have a discussion about the risks with the patient. I'm just wondering if you could put something on that, you know, that you have to discuss the cancer risk involved in potentially using it in this situation. It's just kind of another thing, because I think you're right in saying that basically this is a situation where there's some data for efficacy, we're in a little bit of a grey area, there is safety data, but these may be situations where the benefit of this outweighs the harms, but there are harms, and this might be a situation where kind of shared decision making with the patient becomes a sensible thing to include.

Chris Standaert: I mean you're sort of mandating informed consent, which is already mandated already. That should be part of informed consent, one would think, in theory.

Seth Schwartz: Well, yeah.

Chris Standaert: That should be what people are discussing already. So, putting that, I don't know.

((Crosstalk))

Seth Schwartz: I guess what I think, and maybe I'm wrong in saying this, but my sense from the committee is that the cancer risk is the one thing that's holding us up here, in this situation. We're looking at these exceptions as the cancer risk that's uncertain to us, that we're not sure about, and I don't know when these surgeons, or spine surgeons are having this discussion with their patients, are they saying, we might use this product that's going to put you at increased risk of cancer. I don't know that that's part of the informed consent process for this normally, because this is a rare outcome that's sort of - that I think one of our presenters even said that spine surgeons don't believe that's true anymore. So, I don't know if that's true, but I mean it's - if this is the only thing that's important, which is what we're basically basing this decision on, or perhaps we're making this decision that may be an important piece.

Craig Blackmore: I think it's a part of our decision, but I think there's a lot more. We're considering a large body of data.

Seth Schwartz: Well, I think that's true, but if - so, you could take a hand vote, but if cancer was not a concern would anybody on the committee have a problem with this statement?

Craig Blackmore: Yeah, I'm not saying…

Chris Standaert: There's no data.

Seth Schwartz: All right.

Michelle Simon: Do we have a definition of compromised patients? We might want to. Is it smokers? That was one. There was a list. There are a couple of different lists, I think, defined differently.

Craig Blackmore: No matter how we do this, we're going to have to leave a lot to the judgment of the surgeon and the agency, and in my mind it's just a matter of how much we want to manage that decision. It may be that we want to manage it a lot and we want to have a lot very specific words in here, and we've certainly done that, 2 mm versus 3 mm, or whatever. Or, it may be that we don't feel comfortable doing that. I mean, I don't feel comfortable doing that, but we can.

Michelle Simon: It may not leave the decision to the surgeon, it may leave it to the agency medical directors, though, because they're the ones that would be approving it, right?

Craig Blackmore: Yeah.

Michelle Simon: So, if we're not clear with them, then.

Man: But it's going to have be approved anyway, and they're going to have to get - the surgeon is going to have to prove that it's (unintelligible).

Marie Brown: Well, regardless of the origin of the compromised patients, I'm okay with a broad definition of compromised patient.

Craig Blackmore: Yeah.

Marie Brown: I mean, I don't want to micromanage that.

Craig Blackmore: I think we need to make a decision about how tightly we want to control the restrictions, and I think we'll have to go with maybe go with 2 proposals. One is a very general proposal, which I might put up, which said any revision or any compromised patient, based on this definition, which is going to leave a lot for discretion, and the alternate proposal would be perhaps, Chris is saying that it has to be revision posterolateral lumbar fusion on a compromised patient, which is a very restrictive definition. We're going to have to make a decision as a group which one of those. We can fine tune it, but which one of those is more comfortable.

Michelle Simon: Can we ask Gary how he interprets?

Gary Franklin: I guess I'm a little confused now. So, are you talking about - maybe Dr. Lee can help here. If the fusion fails at a single level

and you're coming in for revision at that level, that might be reasonable. Just the terms that are there now could be any number of levels. You know, maybe the first one failed, but now you're coming in to do 3 levels with no dose kind of limit. So, then you have a whole another situation that - you know. The main reason for bony failure, I think, is in smokers. Smoking causes - is associated with a 2 or 3-fold increased risk of bony failure, and Dr. Lee can please chime in here, but I'm not sure that if a graft fails from smoking, I don't know whether BMP also wouldn't fail related to smoking. Maybe you could comment on that.

Michael Lee: Smoking is a recognized risk factor for pseudoarthrosis, or nonunion. In revising a nonunion, multiple levels may be addressed the second time around, only because one of the strategies for revising a failed fusion is to obtain more fixation points for stronger rigidity of that constrict. So, it is feasible that in revising a single-level fusion, multiple levels may be instrumented and arthrodesed at that time, and the last point Dr. Franklin?

Gary Franklin: Smoking. If a bony graft fails from smoking, wouldn't BMP also fail from smoking? What is the evidence that BMP could overcome the smoking effect?

Michael Lee: I don't know that evidence - I don't know that to be the case. Anecdotally, many people will use InFuse BMP in smokers because of the higher risk of nonunion, anecdotally.

Craig Blackmore: So, again, we are, as we often are, trying to define conditions in the context of incomplete evidence, and that's our reality. So, is the committee without specific words but to help us go forward, is the committee comfortable with a very restrictive set of conditions or a relatively nonrestrictive set of conditions to try and basically define high-risk, and I need to know so we can drill down one way or the other. So, who thinks we should try to be highly specific in how we define the indication? So, without the specific words, so who thinks we should give a general high-risk patients a you can use it, and who thinks we should have specific conditions, condition A, condition B, condition C defined very precisely. That's my question, because I think that's where we are. So, I'll start and say I think we should be general and try to leave it at high-risk and leave it open, and then who thinks we should try to be much more precise and limiting.

Michelle Simon: I don't really like either of them. I don't think there's necessarily evidence for us to say that there is a benefit in high-risk patients. We don't have that evidence. So, why are we saying that that's okay?

Craig Blackmore: Okay. So, what do you think the alternative is?

Michelle Simon: I don't think we should do this at all, this clause. I mean, there is FDA-approved cases, and we should stick with those. That's how I feel.

Craig Blackmore: So, only FDA indications. Okay. So, I need a show of hands or nods of heads. Is that where the committee is, that we should be limiting this to the FDA indications that we've seen, and give me a show of hands if you think that's where we should be limiting.

Chris Standaert: I'm counting that last paragraph as one of the FDA indications, but I'm - the problem - the safety thing seems to just get higher when you get off-label. The safety stuff gets higher.

Craig Blackmore: Okay. I need.

Chris Standaert: For BMP-7 it is. That last paragraph shows (unintelligible) BMP-7.

Man: That's for BMP…

Seth Schwartz: Reword your question. Indications for BMP-2 FDA-approved indications?

Craig Blackmore: So, I think if I understand what Michelle is suggesting is that we limit BMP-2 to the BMP-2 approved FDA indications, which are…

Chris Standaert: Just the 1-level fusion between L4 and S1.

Craig Blackmore: Scroll there up, scroll higher on there. Primary anterior open or laparoscopic fusion at L4-5, and that's it.

Group: That's it.

Marie Brown: And that would restrict coverage from what we have now fairly - it would restrict it more.

Craig Blackmore: Anything we do restricts coverage from what it is now.

Chris Standaert: 85% of the use now is not that. That's only 15% of the use, is what I saw.

Marie Brown: Okay.

Craig Blackmore: So that would not allow use of this for anybody else. So, I'm seeing 3 people who are in that camp. So, am I assuming then that the others believe that there should be other additional indications? So, then it's a matter of defining what those are.

Michelle Simon: Can I say one more thing.

Craig Blackmore: Yes.

Michelle Simon: There is purportedly a study coming out this fall, which might provide a lot more data and information, and when we re-view this topic at whatever timeframe we rereview it, we may have a lot more granularity to make a better decision on it.

Craig Blackmore: We might, and we might not. I don't know. We may have a lot more data and still have the same question, because I - I mean, yeah. Okay. It is a challenge. Okay, I think what we need to do is vote on each line and take it from there. We've already voted on lumbar. We've already voted on adults. We've already voted on primary L4-5.

Man: Straw poll vote.

Craig Blackmore: We haven't officially voted. Sorry, this is unofficial. This is just a straw to set up our final vote. So, I'm going to take votes now on first on the top 2 lines and then second on the bottom paragraph. So, the provision and compromised sentences together, as separate indications, and then I'll vote separately on revision in a compromised patient only, as a more restrictive. No I won't. I'm going to take votes on the first 2. Okay.

Carson Odegard: Before we do that, because we still have that question on whether separating these things out, would it help to put the word or in there at all? That…

Craig Blackmore: So, we're voting, we'll put the word or in, and we'll have a straw vote.

Carson Odegard: Okay.

Craig Blackmore: Thank you. Right there. So, that's what we're voting on. We're not voting on the bottom 3 lines. We're voting on - we're taking an unofficial poll on whether we want to include this as a coverage indication, and the alternative to this is the 3 lines below. So, you get to either vote on the top of the page or the bottom of the page. So who wants the top of the page? Who wants the bottom of the page? Are there 8 of us? Did we get 8 votes, or did we get 7 votes.

Josh Morse: I thought I got 7 votes. Did I miss one?

Man: I didn't vote.

Craig Blackmore: You didn't vote? Why?

Seth Schwartz: Because I don't support either of them.

Craig Blackmore: Because you don't support either of them.

Chris Standaert: I think if you did the top one, you actually don't even need the first sentence. You don't need the revision postero - you're opening it up for anybody who can't have a bone graft or needs something else to help their - you think will not fuse.

Craig Blackmore: I would take out the word posterolateral. Say either revision or you don't have bone to graft, to harvest.

Seth Schwartz: That's even more expanding.

Craig Blackmore: Yes.

Man: Take out the word posterolateral in that one.

Craig Blackmore: All right. Okay. Questions brought up by Dr. Franklin, should this say single level? So, I want a show of hands whether this - the indication should be restricted to single-level? All right. I want discussion.

Chris Standaert: Several of us aren't in agreement with that first section, so saying a single-level or double-level, I'm not going to respond, because I don't agree with it. So, the bottom part - it's a posterolateral, it can only get so big, so it's a 1 or 2-level deal, because it's posterolateral fusion, and that's sort of what the FDA says is a

compassionate use thing, and that's why I'm going with it, because it gives some out if things go bad.

Craig Blackmore: This is your opportunity to influence what those conditions are. So, if you're not voting you're not.

((Crosstalk))

Marie Brown: So, would you do just the last paragraph?

Chris Standaert: I voted for just the last paragraph.

Marie Brown: Okay.

Craig Blackmore: Okay. Committee, do we think single-level should be included as a restriction on coverage or not? I'm not hearing yeses or no's. Is there - I'm a little…

Seth Schwartz: I guess I'm just trying to still wrap my head around how all this looks, because I'm just sort of thinking about, you know, where things are now to where we're going, and I guess I think what we're doing here is we're excluding - we're saying, okay, there's this specific indication for this one approach for single-level fusion in a standard patient, we're saying for primary cases anything other than that, we shouldn't use it, because the risks are not - the risks outweigh the benefits. We're then saying that of all the patients to get fusion, what is the failure rate? The failure rate's 3 to 5%. I can't remember what it was, but just in general primary without using BMP is 5% say, give or take. So, then we're saying that the cancer risk in patients who use BMP we think is somewhere on the order of, I don't know, what are the numbers? What was the actual number for cancer risk? 1 or 2%? So, if you're talking about only using this on the revision patients, then you're talking about 1% of 5%, so you're talking about 1 in 5000 patients is at potential risk with that restriction. That sounds like a pretty good reduction compared to just using it on everybody. I don't know what this patient population looks like for just these compromised patients. I don't know how many people who get this kind of operation are optimized patients, because if 50% of the patients who are going to have this surgery are, you could make an argument on compromised, I'm a little less comfortable with saying we should say go ahead and use it. If we're talking about a fairly small percentage of patients, like the revision patients, which we know is going to be limited to 5% of people, I

feel much more comfortable saying we should use it. We can have that situation be at the surgeon's discretion.

Chris Standaert: You're not defining compromised, and revision, like Dr. Lee said, if you're advising a 1-level fusion, sometimes the best way to revise a 1-level fusion is to turn it into a 2-level fusion. So, to restrict somebody to a revision lumbar fusion that's only 1 level, you're actually tying, frankly you're doing the opposite of what your intention is. You're tying a surgeon's hands more. So if you make this 1-level, you really confine this a lot more.

Seth Schwartz: I'm not arguing when you say 1-level. That's not what I'm looking at here. I'm saying, I'm just thinking in terms of defining - I'm just thinking in terms of what are the populations? What are we opening ourselves up to? What are we opening this up to with these potential exceptions.

Craig Blackmore: So, we're struggling. Are we struggling around the compromised patient or are we struggling around the revision? Would the group be more comfortable with, having already heard that we don't like both, would the group be more comfortable with limiting this to revision and dropping the compromise piece? Is that where the hangup is, or not?

Chris Standaert: Revision in compromised patients is a much smaller patient population than compromised patients, because by definition, you're in the revision population, so you're down to the 5% or less. Now you're saying that compromised revision patients, and then you're giving the surgeon leeway. I mean, that's why the wording is that way, I assume. They have to give - you're giving somebody leeway to say I need to redo this fusion, and I don't think this person is going to do very well if I just use iliac crest or I can't get it. So, I need something else. BMP's around and maybe it will help me. That's a fairly restrictive population, because you're implying that maybe it's 50%, maybe it's 20, maybe it's 10, and of people who fail, it's probably a higher rate. The numbers are a little small, but when you say just compromised patients who are not expected to get a fusion, I mean that's - that could, defining it could (unintelligible) population.

Seth Schwartz: I think that's what I was really getting at, which is that - so we're talking about, we've already kind of agreed that on-label use, well we haven't agreed, but I guess the leaning is that for on-label use we're pretty much okay with that. So, we're talking about for off-

label use, who are we talking about? So, I think that if you're going to say, if compromised patients could be defined as half of the population that's having this, because they're smokers or they're old, or they're osteopenic or whatever else, then that's a pretty big window of the exception, but if we're talking about revision patients, or revision patients who are also compromised, then again we get - when looking at this risk-benefit ratio, it starts to look a little better. I'm not sure that the posterolateral is still necessary, if you restrict this to revision patients who are compromised, that gets you a pretty tight group where the risk/benefit ratio might be acceptable to us. I don't know. That's kind of - looking at it that way, I think that makes a little bit more sense to me, and I guess I would rescind my previous nonbinding vote and go with the second one, if we get rid of the posterolateral portion.

Craig Blackmore: So, the upper half of the page now, is that a condition that the group is more comfortable with?

Group: No.

Seth Schwartz: I'm talking about the lower half. I'm talking about get rid of posterolateral in the lower half.

Craig Blackmore: Right. If I change the or to an and on the top half? Use the bottom, get rid of the word posterolateral.

Chris Standaert: Just use the bottom.

Marie Brown: Just use, yeah, and take out posterolateral.

Chris Standaert: Bottom paragraph. Bottom sentence there.

Craig Blackmore: Don't erase (unintelligible).

Chris Standaert: See, you're done. See, that was fate. Fate intervened, what are you doing? So bottom paragraph, remove that word posterolateral. That means you can revise it.

Craig Blackmore: So, nods or hands or some indication as to whether the bottom condition, as we've got here. Is that more comfortable?

Group: Mm-hm. Yes.

Craig Blackmore: I'm not seeing enthusiastic nods, but I'm seeing little nods. That's good. Okay, Margaret. Now, if we could eliminate from patients at risk at the top through the fusion half. Yes, thank you. And then, do we still have the rest of it on there? Okay, stop. Are we at a point where - is there anything else that needs to be on this list? Dr. Franklin, is this implementable as a list of conditions?

Gary Franklin: I believe that we can implement anything that you ask us to, because we do prospective review on these. If we weren't doing prospective review, this would be pretty hard to do, but because we do prospective review with a pretty detailed checklist of what is actually approved, I believe that we could do anything you ask us to do.

Craig Blackmore: Okay.

Man: Do you need the lumbar spine only like you have (unintelligible).

Craig Blackmore: Yes. No, but yes. We'll leave it on. Okay. I think we're ready to move on to voting. I think we've worked the conditions down to a place where we're as comfortable as we're going to be. We still need to decide if we're going to cover or if we're going to cover with conditions, or not cover. I don't believe we've done the official nonbinding vote on the use of BMP-2 aside from the FDA indications.

Marie Brown: And we need to.

Chris Standaert: We haven't done any binding vote on BMP-2.

Craig Blackmore: We haven't done a nonbinding vote.

Chris Standaert: A nonbinding vote.

Craig Blackmore: We did a nonbinding vote on the FDA indications.

Chris Standaert: Oh, that's - okay. Yeah, I got you.

Craig Blackmore: The nonbinding, but in terms of process I want to stay on track. So, the next thing we do is a nonbinding vote on the effectiveness, safety, and cost effectiveness of BMP-2 and this is specifically excluding the FDA indication of primary anterior open or laparoscopic fusion at 1 level between L4 and S1, because we've already voted on that. So, the committee will vote that it is more

effective if they believe there is any circumstance outside of 1-level between L4 and S1 where BMP-2 is more effective, and I'll just point out that the bottom category that we've defined there, if you believe it's more effective in that category, you would vote more. Is the question clear? Okay. So, the nonbinding vote - actually I'm going to rethink this.

Josh Morse: 6 unproven, 2 more, and this is for effectiveness.

Craig Blackmore: Next is the safety.

Josh Morse: 8 less.

Craig Blackmore: And the cost effectiveness?

Josh Morse: 8 unproven.

Craig Blackmore: Now, we make a binding vote. This is the coverage decision for BMP-2 and your choices are to vote for cover, which means cover under all conditions, to not cover, which means never cover, or to cover with conditions, and the conditions are predefined, as use in the lumbar spine only, use for primary anterior open or laparoscopic fusion at 1 level between L4 and S1, or use in revision lumbar fusion on a compromised patient for whom autologous bone and bone marrow harvest are not feasible or not expected to result in fusion, and I will point out that the decision only applies to adults defined as age 18 and over.

Seth Schwartz: Can I ask a procedural question?

Craig Blackmore: Yes. So, there's really only 2 possible votes here, correct? Cover or not cover, because we've already gone through conditions.

Chris Standaert: No, it's cover with conditions (unintelligible).

Craig Blackmore: No, we haven't - you can say no cover, you can say…

Seth Schwartz: We're saying no cover for all of this or no cover for BMP-2.

Craig Blackmore: All right. I'll start over. You have 3 choices. Choice number 1 is no cover, and that means it is never covered.

Seth Schwartz: BMP-2.

Craig Blackmore: This list is completely irrelevant, we're not paying for it.

Seth Schwartz: Right.

Craig Blackmore: Cover means, we're covering everything, this list is completely irrelevant. We're covering everything. The third choice is cover with conditions, and the conditions are these.

Carson Odegard: And then the third one is abstain.

Craig Blackmore: Abstain, but I hope you won't.

Seth Schwartz: You're not giving. I mean, it doesn't matter. It's a procedural issue, and I only have one vote, but.

Craig Blackmore: I'm not giving you a choice for other conditions, because the group has already decided these are the conditions. So, we can either accept these conditions or you can vote for nothing or everything. Those are your 3 choices. Is that clear to everybody? Okay.

Josh Morse: We have 8 cover with conditions.

Craig Blackmore: So, we are, procedurally we are to determine if our decision is consonant with the National Coverage Decision, and if it is not, we are to justify our reasons for not agreeing, and there is no National Coverage Decision. Thank you, all, for your hard work. Thank you, Josh and the team.

Josh Morse: Thank you.

Marie Brown: Do we give you all this paper?

Craig Blackmore: Further comments?

Marie Brown: And notebooks and stuff?

Craig Blackmore: We are adjourned.