

SPECIAL ARTICLES

The Evolution of the AASM Clinical Practice Guidelines: Another Step Forward

Timothy I. Morgenthaler, MD¹; Ludmila Deriy, PhD²; Jonathan L. Heald, MA²; Sherene M. Thomas, PhD²¹Mayo Clinic, Rochester, MN; ²American Academy of Sleep Medicine, Darien, IL

One of the goals of the American Academy of Sleep Medicine (AASM) is to provide clear, evidence-based recommendations in our clinical practice guidelines. Periodically, the AASM will assess and update the process by which these guidelines are developed so that it is in line with the standards currently being used for guideline development. The AASM is now taking the next step forward by fully adopting GRADE (Grading of Recommendation Assessment, Development and Evaluation) as the methodology used for evaluating evidence and forming clinical practice guidelines recommendations. Starting this year, AASM recommendations will be based on the following four interdependent domains: 1. quality of evidence; 2. balance of desirable and undesirable consequences; 3. patients' values and preferences; and 4. resource use (when known). AASM strengths of recommendations will be dichotomized into two categories: "Strong" and "Weak," either for or against a patient-care strategy. In an effort to provide clarity and transparency, all AASM recommendations will be actionable statements that include the specific patient population for which the patient-care strategy is recommended, and clearly define the comparator against which the patient-care strategy was evaluated. In some recommendations, the comparator will be an alternative patient-care strategy (e.g., a "gold standard" or previously available alternative), while in other recommendations the comparator will be a placebo or no treatment; this is determined by the availability of evidence, and analyses decisions made by the AASM task force. Implementation of the complete GRADE criteria by the AASM allows us the best path forward towards continuing to provide high quality clinical practice guidelines.

Keywords: guidelines, GRADE, sleep**Citation:** Morgenthaler TI, Deriy L, Heald JL, Thomas SM. The evolution of the AASM clinical practice guidelines: another step forward. *J Clin Sleep Med* 2016;12(1):129–135.

INTRODUCTION

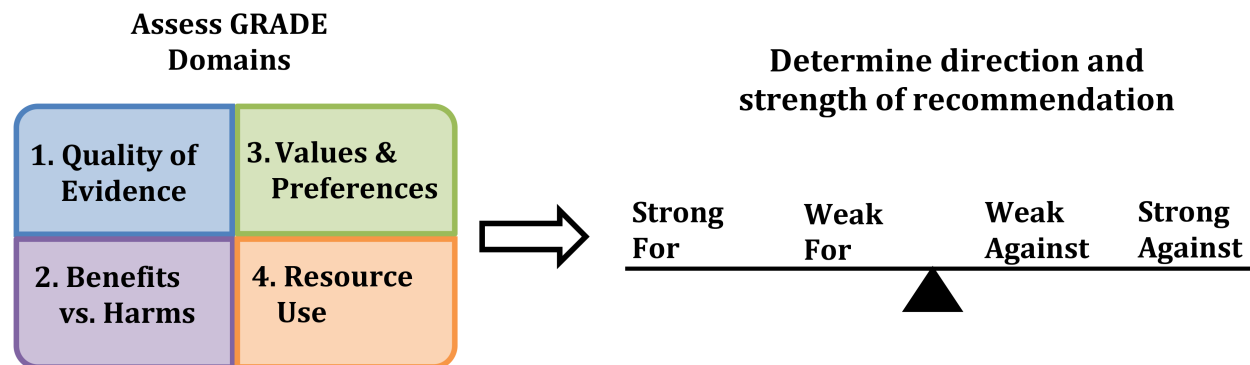
American Academy of Sleep Medicine clinical practice guidelines (and the prior Practice Parameters and Best Practice Guides) are developed to bring the fruits of clinical and basic research into practice, offering guidance to practicing sleep specialists and other health care professionals who work with patients and families who suffer from sleep disorders. These guidelines have influenced clinical practice in the office and at the bedside, on a national and international level. It is therefore imperative that the very best efforts and techniques be applied to produce guidelines that are accurate, clear, and reliable. This document will describe some recent evolutionary changes to how recommendations will be developed by the AASM and how they will appear in AASM clinical practice guidelines, beginning this year.

In 2009, the AASM began adopting GRADE (Grading of Recommendations Assessment, Development, and Evaluation) as the methodology for evaluating evidence and forming clinical practice guideline recommendations.¹ The major impetus for this change was that, compared to the prior employed methodologies which heavily emphasized study design, GRADE placed weight on the systematic evaluation of a wider spectrum of characteristics of the evidence (e.g., not only study design, but also the directness with which the evidence addresses the clinical question, appropriate use of blinding, the degree of unexplained heterogeneous results across studies). The GRADE system also made it easier for

users to assess the judgments and decisions that are integral to forming recommendations, by making those judgments explicit and transparent. Finally, GRADE made it possible to place merit on well-conducted observational studies, the significance of which may be underestimated by other evidence-grading systems. This factor was thought to be especially important in an evolving and unique field such as sleep medicine, where large randomized controlled studies are not always available. This systematic approach, combined with the explicit transparency and flexibility of the GRADE system, made it a good choice for guideline development. Not surprisingly, the GRADE method has also evolved since its adoption by the AASM. The well-reasoned changes offer an opportunity and challenge to further improve the development and presentation of the AASM clinical practice guidelines.

When GRADE was first adopted by the AASM, the AASM had produced 31 practice parameter/best practices papers. The recommendations within carried strengths of "Standard" (a generally accepted patient-care strategy, which reflects a high degree of clinical certainty), "Guideline" (reflects a moderate degree of clinical certainty), or Option (reflects uncertain clinical use). During the transition to GRADE in 2009, it was felt that the best approach would be to adapt GRADE methods for assessing the strength of the evidence while maintaining the nosology of Standard, Guideline, Option for the levels of recommendations.¹ The first AASM Practice Parameter using this modification of GRADE was published in 2010.²

Figure 1—An overview of using GRADE to make recommendations.



THE IMPETUS FOR ANOTHER UPDATE

At the time of this adoption, the GRADE Working Group was still developing a series of publications that explained the specifics of implementing GRADE; fifteen papers of the 20-paper series have since been published in the *Journal of Clinical Epidemiology*.³⁻¹⁷ The most recent papers, published in 2013, detail the process of using GRADE to determine the strength of recommendations.^{16,17}

In late 2010 the GRADE Working Group also published a list of criteria that must be met for full GRADE compliance,¹⁸ summarized here:

1. “Quality of evidence” should be defined consistently with the definition used by the GRADE Working Group.
2. Explicit consideration should be given to each of the GRADE criteria for assessing the quality of evidence, although different terminology may be used.
3. The overall quality of evidence should be assessed for each critical and important outcome and expressed using four (e.g., high, moderate, low, very low) categories that are consistent with the definitions used by the GRADE Working Group.
4. Evidence summaries should be used as the basis for judgments about the quality of evidence and the strength of recommendations. Reasons for upgrading and downgrading should be described transparently.
5. Explicit consideration should be given to each of the GRADE criteria for assessing the strength of a recommendation (the balance of desirable and undesirable consequences, quality of evidence, values and preferences, and resource use) and a general approach should be reported.
6. The strength of recommendations should be expressed using two categories (weak and strong) for or against a patient-care strategy and the definitions for each category should be consistent with those used by the GRADE Working Group.
7. Decisions about the strength of the recommendations should ideally be transparently reported.

Based on the details contained within these publications, and correspondence with the developers of GRADE, it was evident that the AASM modification of GRADE required

further updating. Specifically, maintaining the nosology of “Standard,” “Guideline,” and “Option” did not meet Criterion 6. Additionally, determining the strengths of recommendations based only on the quality of evidence and the balance of benefits and harms, and not explicitly considering patient values and preferences and resource use as separate domains, did not meet Criterion 5. Therefore, as of 2016, the AASM has updated its practices and will now be using GRADE according to the modernized criteria.

THE FOUR DOMAINS OF GRADE: HOW FULL ADOPTION OF GRADE WILL INFLUENCE AASM PRACTICE RECOMMENDATIONS

AASM recommendations developed using GRADE will be based on the following four inter-dependent domains: (1) quality of evidence; (2) the balance of desirable and undesirable consequences; (3) patients’ values and preferences; and (4) resource use (when available) (**Figure 1**). These domains and their role in the development of AASM recommendations are summarized below.

1. Quality of Evidence

AASM practice recommendations are based on a systematic review of the literature, focusing on patient-centered outcomes as much as possible. For these recommendations, “quality of evidence” reflects the certainty that an estimated effect is sufficient to support the recommendation (GRADE Criterion 1), and should not be confused with the strength of a recommendation. The quality of evidence includes effect estimates for both beneficial and harmful outcomes and is assigned one of four categories (high, moderate, low, or very low), which reflect the certainty in the evidence (**Box 1**) (GRADE Criterion 3). For each clinical question addressed in a guideline, the selected outcomes for the patient-care strategy (i.e., intervention or diagnostic test of interest) are rated as “critical,” “important,” or “not important” for decision making. The quality of evidence for each “critical” and “important” outcome is determined based on the literature review, and then the overall quality of evidence is determined for each recommendation.

Box 1—Definitions of quality of evidence categories.

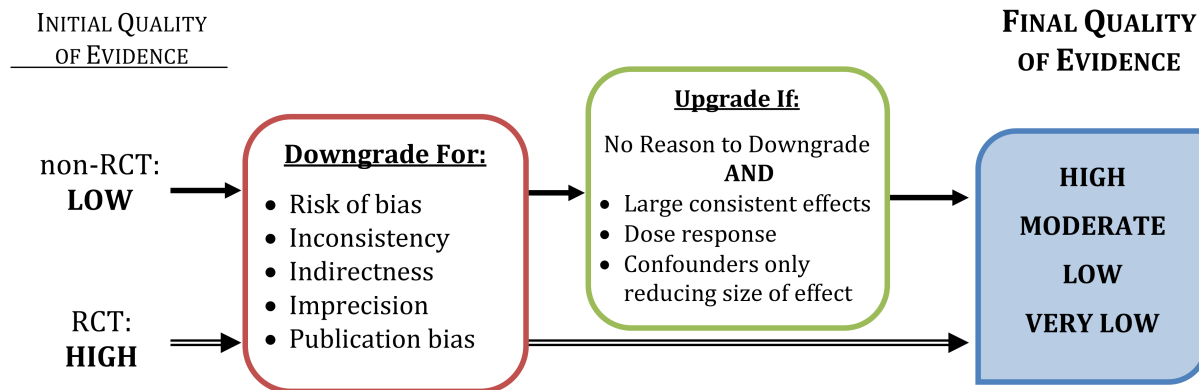
High: corresponds to a high level of certainty that the estimate of the effect lies close to that of the true effect.

Moderate: corresponds to a moderate level of certainty in the effect estimate; the estimate of the effect is likely to be close to the true effect, but there is a possibility that it is substantially different.

Low: corresponds to a low level of certainty in the effect estimate; the estimate of the effect may be substantially different from the true effect.

Very low: corresponds to very little certainty in the effect estimate; the estimate of the effect is likely to be substantially different from the true effect.

Figure 2—How GRADE is used to determine quality of evidence for individual outcomes.



Box 2—Factors evaluated to determine quality of evidence.

Risk of Bias: Risk of bias refers to the potential for bias by the authors of the publication(s), which might affect the reported outcomes. Examples of risk of bias include lack of allocation concealment, lack of blinding, incomplete accounting of patients and outcome events, and selective outcome reporting.

Publication Bias: Publication bias is the most difficult problem to detect and refers to selective publication of results that favor the treatment or diagnostic tool. Potential sources of publication bias include preliminary or pilot studies, incomplete reporting of “negative outcomes”, selective reporting by journals (e.g., editorial considerations, peer review), and author revision and resubmission.

Imprecision: Imprecision refers to the confidence of the task force that the estimate of effect supports a recommendation, as measured by the relationship of the mean difference and 95% confidence interval in relation to a clinical significance threshold. The clinical significance threshold is the threshold at which a treatment effect or diagnostic outcome is considered to be clinically significant and provide a true benefit or true harm.

Inconsistency: Inconsistency, also known as heterogeneity, occurs when the estimate of effect for an outcome differs greatly across studies and reduces the confidence that the estimate of effect accurately reflects the true effect for patients.

Indirectness: Indirectness occurs when the studies used to determine the estimate of effect for a patient-care strategy do not represent the patient population for which the recommendation will be used. Indirectness can result from differences in patient populations, interventions, outcome measures, or from indirect comparisons.

The initial quality of evidence for each outcome is assigned according to the design of the studies included in the body of evidence. For treatment guidelines, the evidence from randomized controlled trials (RCTs) is assigned high quality, while evidence from non-RCT observational studies is assigned low quality. For this reason, RCTs and non-RCTs are often evaluated separately. For diagnostic guidelines, RCTs are still ideal and are assigned high quality, however due to the nature of diagnostic tests, RCTs are often unavailable. Therefore, appropriately designed observational studies can be assigned an initial quality rating of high; however such assignments must be considered carefully.¹⁹

After the initial quality assignment, high quality evidence can be graded down, while low quality evidence can be graded up or further down (Figure 2), based primarily on evaluation of the following categories: risk of bias within and across

studies⁶; publication bias⁷; imprecision of the estimate of effect⁸; inconsistency across studies⁹; and indirectness of the population, intervention, or comparator¹⁰ (Box 2) (GRADE Criterion 2). In some rare instances the quality of evidence can be upgraded (Figure 2), increasing confidence in the estimated effect. Additional information about grading up the quality of evidence can be found in paper 9 of the GRADE series.¹¹

Rather than assigning a quality rating to individual studies (e.g. as done under the Oxford system), GRADE determines the quality of evidence for each “critical” and “important” outcome across studies. Thus, “quality” is not a reflection of the quality of individual studies, as well-conducted studies can produce evidence that does not provide clear support for clinical decision-making. For example, when well-conducted randomized trials provide contradictory, or widely inconsistent,

results for a therapeutic outcome, the quality of evidence for that outcome would be downgraded for inconsistency across studies, resulting in lower confidence that the estimated effect is representative of the true effect that patients will see.

The overall quality of evidence for a recommendation for or against a given care strategy usually rests upon the lowest quality of evidence across all “critical” outcomes for the patient-care strategy of interest (GRADE Criterion 3).¹³ As discussed previously, the quality of evidence refers to the confidence that the estimate of a given patient-care strategy’s effect is representative of the true effect. Suppose, for example, that we evaluate the quality of evidence for a patient-care strategy based on three outcomes that are each determined to be “critical” for clinical decision making: sleep latency, the number of awakenings, and total sleep time. Based on a systematic review of the evidence, the patient-care strategy possesses moderate quality evidence that it decreases sleep latency, high quality evidence that it decreases the number of awakenings, but only low quality evidence that it increases total sleep time. A chain is only as strong as its weakest link; therefore the overall quality of evidence that the patient-care strategy will accomplish all the “critical” outcomes would be low.

Transparency is a key element in the GRADE approach to guideline development; therefore all reasons to upgrade or downgrade the quality of evidence are presented as footnotes in a Summary of Findings table for each patient-care strategy. These tables serve as the basis for the determination of the final quality of evidence, one of the four domains upon which the recommendations will be based (GRADE Criteria 4),^{14,15} and will be used when determining the balance of benefits and harms of the patient-care strategy.

2. Balance of Benefits and Harms

The balance of benefits and harms (i.e., desirable and undesirable consequences of a patient-care strategy) is assessed for each recommendation. This evaluation is based on the evidence considered in the Quality of Evidence assessment⁴ (i.e., the estimate of effect for both beneficial and harmful outcomes of a patient-care strategy), the results of which are presented in the Summary of Findings tables. The balance of benefits and harms can be categorized as follows: benefits outweigh harms, benefits equal harms, harms outweigh benefits, or the balance between benefits and harms is unclear. The direction of a recommendation (for or against) will be partially determined by the balance of benefits and harms, while the strength of a recommendation (strong or weak) will be influenced by the magnitude of the balance (**Figure 3**).

In addition to an objective review of the evidence, determining the balance of benefits and harms also includes consideration of patient values and preferences (further detailed below). For example, a patient-care strategy might demonstrate clinically significant improvement in total sleep time and sleep latency, while the evidence also suggests that headache and nausea are potential side effects. Based on this evidence, a recommendation is made in favor of the patient-care strategy. However the side effects lower the confidence of the task force that all patients would choose this patient-care strategy over an alternative, resulting in a weak recommendation for the patient-care

strategy. Alternatively, if the side effects were considered to be minimal or unimportant to most patients when compared to the beneficial outcomes of the patient-care strategy, then the recommendation might be strong. In this way, patients’ values and preferences form an axis upon which the beneficial and harmful outcomes of a patient-care strategy are balanced (**Figure 3**).

Thus, the balance of benefits and harms includes both an objective review of the evidence, and a subjective value-judgment of the consequences of a patient-care strategy. Such decisions will be clearly explained in the guideline in order to help clinicians understand the judgments that went into determining the strength of a recommendation, and allow them to better apply the recommendation to their patients.

3. Patients’ Values and Preferences

Patients’ values and preferences encompass the perspectives, beliefs, and expectations for health and life of the patient population, and refer to the process that individuals use when considering the consequences of using a patient-care strategy.¹⁷ Such information can be gathered from the literature (when available), stakeholders (such as patient advocacy groups), and the experience of practicing clinicians. For the purposes of making practice recommendations, patient values and preferences focus on the patient population under consideration, rather than on individual patients, and should therefore include consideration of their uniformity across the patient population. If values and preferences vary greatly across the patient population under consideration, the strength of a recommendation should be lower than if all patients’ values and preferences are similar for a particular patient-care strategy. For example, if the inconvenience created by using a patient-care strategy is considered by some patients to be too great, while other patients are willing to accept the inconvenience in return for improved sleep, then the strength of that recommendation would be weaker than the recommendation for a patient-care strategy that creates little or no inconvenience.

Rather than being a post hoc exercise, considered only after the evidence review is completed, inclusion of patients’ values and preferences spans the entire GRADE process (**Figure 3**). Patients’ values and preferences guide the evaluation of benefits and harms (discussed above), make determinations about resource use (further detailed below), help determine thresholds for setting clinical significance, and influence the selection of outcomes (e.g., determining patient-important harmful outcomes of a patient-care strategy) used as the basis for the evidence review. For example, taste distortion is a harmful outcome of some patient-care strategies, but patients may not consider it a significant factor when deciding to use a patient-care strategy; therefore, taste distortion would not be considered an outcome of interest when reviewing the evidence. However next-day drowsiness is a harmful outcome that may influence many patients’ decisions regarding the use of a patient-care strategy, therefore next-day drowsiness would be considered a “critical” or “important” outcome. Patients’ values and preferences can also help with setting clinical significance thresholds for the outcomes of interest by taking into account the magnitude of change that patients feel is significant enough to warrant using a patient-care strategy. This includes changes in both beneficial and harmful outcomes (e.g., how large the

improvement in quality of sleep should be, and how much daytime dizziness would be tolerable). These assessments might differ from the estimates of clinicians, providing valuable guidance when making recommendations that are patient-centered. In these ways, patients' values and preferences are incorporated into the process of evaluating evidence and determining the direction and strength of a recommendation.

4. Resource Use

When available, information about resource use should be considered when determining the strength of a recommendation. Resource use refers not only to the monetary cost of a patient-care strategy, but also the availability and potential health-disparity associated with recommending a patient-care strategy.¹⁷ For example, a patient-care strategy that is high cost, or is not widely available to the entire patient population might receive a weaker recommendation than a patient-care strategy that is low-cost or more widely available.

Patients' values and preferences can also influence resource use evaluations by helping to determine the financial cost that patients would be willing to incur to receive the benefits of a patient-care strategy, or the variability of that decision. For example, if a patient-care strategy provides significant improvement in sleep parameters, but the financial burden would be untenable for many patients, the strength of the recommendation would be weak. Another example would be if a patient-care strategy shows moderate improvement in sleep parameters but is low-cost, it might receive a strong recommendation.

guidance by clinical sleep experts, while facilitating clinical decision-making on an individual basis. Therefore, the AASM recommendations will be based on the four domains discussed above¹⁶ (GRADE Criterion 5), and detail the judgments and decisions supporting the strength and direction of the recommendation. The consideration of all of these components, rather than relying solely on the quality of evidence, is intended to encourage clinicians to determine how best to implement AASM recommendations for each patient they encounter.

AASM recommendations will be dichotomized into two strengths, "Strong" and "Weak", directed either for or against a patient-care strategy (GRADE Criterion 6). **Table 1** provides example characteristics of each direction and strength. It should be noted that these are not prescriptive characteristics, but serve as examples of possible reasons that a task force might assign a particular direction and strength to a recommendation.

A strong recommendation for a patient-care strategy is a recommendation that clinicians should, under most circumstances, always be doing (i.e. something that might qualify as a Quality Measure, further discussed below). Accordingly, a strong recommendation against a patient-care strategy would be something that clinicians should, under most circumstances, NOT be doing. A strong recommendation against a patient-care strategy may be the result of harmful outcomes that outweigh the beneficial outcomes, a patient-care strategy that has been found clinically ineffective and therefore a poor use of resources or an alternative patient-care strategy that is more effective and better tolerated.

Weak recommendations reflect a lower degree of certainty in the appropriateness of the patient-care strategy and require that the clinician use their clinical knowledge and experience to refer to the individual patient's values and preferences and determine the best course of action. A weak recommendation for a patient-care strategy may suggest that a majority of well-informed patients would choose the strategy, however

DETERMINING STRENGTHS OF RECOMMENDATIONS

The goal of AASM clinical practice guidelines is to summarize the evidence for patient-care strategies and provide

Figure 3—The complexity of determining the direction and strength of a recommendation.

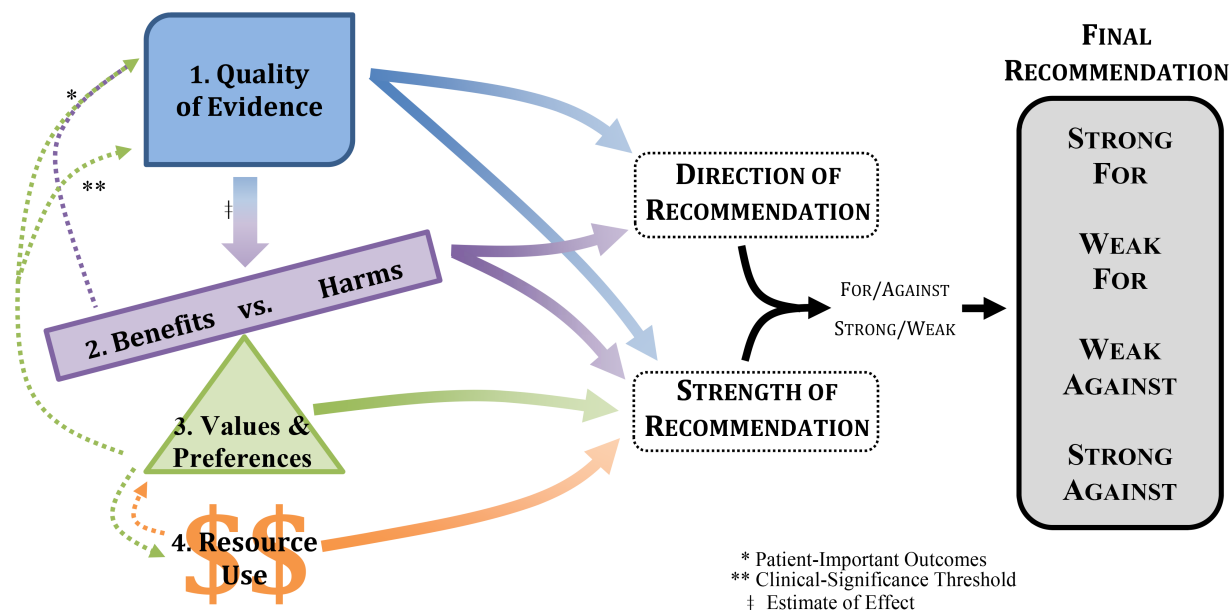


Table 1—Example characteristics of AASM strengths of recommendations.

AASM Strength of Recommendation		Example Characteristics Guiding Recommendation
FOR	STRONG	<ul style="list-style-type: none"> • There is a high degree of clinical certainty that the balance between benefits vs. harms (i.e., <u>net benefits</u>) favors benefit for this patient-care strategy. • The vast majority of well-informed patients would most likely choose this patient-care strategy, compared to alternative patient-care strategies or no treatment.
	WEAK	<ul style="list-style-type: none"> • There is a lower degree of clinical certainty that the balance between benefits vs. harms (i.e., <u>net benefits</u>) favors benefit for this patient-care strategy. • The majority of well-informed patients would most likely choose this patient-care strategy, compared to alternative patient-care strategies or no treatment.
AGAINST	WEAK	<ul style="list-style-type: none"> • There is a lower degree of clinical certainty in the balance between benefits vs. harms (i.e., <u>net harms</u>) of this patient-care strategy. • The majority of well-informed patients would most likely not choose this patient-care strategy, compared to alternative patient-care strategies or no treatment.
	STRONG	<ul style="list-style-type: none"> • There is a high degree of clinical certainty in the balance between benefits vs. harms (i.e., <u>net harms</u>) of this patient-care strategy. • The vast majority of well-informed patients would most likely not choose this patient-care strategy, compared to alternative patient-care strategies or no treatment.

a percentage of patients may not. A weak recommendation against may suggest that a majority of well-informed patients would not choose this patient-care strategy, however a percentage of patients may choose this strategy. Thus weak recommendations are conditional, based on the individual circumstances of the patient and clinician.

The decisions and considerations that are made when determining the direction and strength of a recommendation will be clearly detailed in the clinical practice guideline (GRADE Criteria 7). This will allow the individual clinician to understand why a patient-care strategy received its recommendation, and facilitate clinical decision-making in day-to-day practice.

FORMAT OF THE AASM RECOMMENDATION STATEMENTS

In an effort to provide clarity and transparency, all AASM recommendations will be actionable statements that include the specific patient population for which the patient-care strategy is being recommended, and clearly define the comparator against which the patient-care strategy was evaluated. For consistency across guidelines, recommendation statements will use the language “We recommend” for all strong recommendations (either for or against a patient care strategy), while all weak recommendations will use the language “We suggest.” To facilitate the implementation of the recommendation, each statement will specify the population or sub-population that the recommendation applies to, and any subgroups that the recommendation excludes. In some recommendations the comparator will be an alternative patient-care strategy (e.g., a “gold standard” or previously available alternative), while in other recommendations the comparator will be a placebo or no treatment; this is determined by the availability of evidence and analyses decisions made by the AASM task force.

“NO RECOMMENDATION”: THE LIMITS OF MAKING RECOMMENDATIONS

The objective of clinical practice guidelines is to provide a broad range of evidence-based recommendations for treating a specific disease or disorder, facilitating evidence-based treatments or diagnostics for the majority of patients, in the majority of clinical care settings. However, in an emerging field such as sleep medicine, where there may be very few or no large quantitative studies on the topic of interest, “No recommendation” can be a frequent result. For rare disorders, such as Non-24 hour Circadian Rhythm Sleep-Wake Disorder, the availability of quantitative data can be extremely limited, making it difficult to confidently recommend for or against a patient-care strategy.

It is possible to make recommendations, even strong recommendations, based on a small number of studies, particularly when the balance of benefits and harms and patient values and preferences are clear. However, the task force may not feel confident making a recommendation for or against a patient-care strategy when only a few small studies are available, the quality of evidence is low, the evidence is contradictory, or the benefits/harms ratio is unclear. In this case, “No recommendation” may be indicated, to reflect the lack of certainty in the body of available evidence. Therefore, it will be left to the discretion of the clinician to determine if that particular patient-care strategy is appropriate to use, on a case-by-case basis. It will also serve as a strong indicator that research in this direction is warranted. Discussion of the available evidence will be provided in such cases to assist the clinician to make a more educated decision.

FROM RECOMMENDATIONS TO QUALITY MEASURES

Strong recommendations for patient-care strategies are often made when high quality evidence is available, there is high

certainty in the balance of benefits and harms, and patients' values and preferences are generally consistent. When a patient-care strategy receives a strong recommendation, it is something that a practitioner should, under most circumstances, always do (or not do, in the case of "strong against"). With the expanding body of evidence for patient-case strategies that improve outcomes for patients with sleep disorders, we need to ensure that patients receive recommended therapies (and do not receive therapies that are strongly recommended against). In order to improve the quality of care given to patients with sleep disorders, clinicians must have valid, reliable, and practical tools to evaluate the effects of implementing recommended interventions and tests. The AASM, in consultation with numerous stakeholders, has developed Quality Measures for the diagnosis and management of several sleep disorders^{20–25}; many of these measures are based upon previously established clinical practice recommendations. Going forward, strong recommendations for or against a patient-care strategy will form the foundation of new quality measures, or be used to update existing quality measures. In this way, AASM clinical practice guidelines will be used to provide guidance to clinicians treating patient with sleep disorders, and serve as a metric from which the quality of care can be evaluated and improved.

FINAL THOUGHTS

Developing clinical practice guidelines requires time and careful consideration. Developing best practices for writing clinical practice guidelines is an ongoing journey, as evidenced by the plethora of grading systems and document styles that are currently used by guideline developers across the world.²⁶ With this update to the AASM guideline development practices, the adoption of the GRADE system by the AASM is now up to date. It is the goal of the AASM to promote high quality patient-centered care by providing clear, evidence-based recommendations; our implementation of the complete GRADE criteria allows us the best path forward towards continuing to meet this goal.

REFERENCES

1. Aurora RN, Morgenthaler TI. On the goodness of recommendations: the changing face of practice parameters. *Sleep* 2010;33:1273–6.
2. Aurora RN, Casey KR, Kristo D, et al. Practice parameters for the surgical modifications of the upper airway for obstructive sleep apnea in adults. *Sleep* 2010;33:1408–13.
3. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol* 2011;64:380–2.
4. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 2011;64:395–400.
5. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64:401–6.
6. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol* 2011;64:407–15.
7. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol* 2011;64:1277–82.
8. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol* 2011;64:1283–93.
9. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *J Clin Epidemiol* 2011;64:1294–302.
10. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol* 2011;64:1303–10.
11. Guyatt GH, Oxman AD, Sultan S, et al. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 2011;64:1311–6.
12. Brunetti M, Shemilt I, Pregno S, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol* 2013;66:140–50.
13. Guyatt G, Oxman AD, Sultan S, et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J Clin Epidemiol* 2013;66:151–7.
14. Guyatt GH, Oxman AD, Santesso N, et al. GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. *J Clin Epidemiol* 2013;66:158–72.
15. Guyatt GH, Thorlund K, Oxman AD, et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes. *J Clin Epidemiol* 2013;66:173–83.
16. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol* 2013;66:719–25.
17. Andrews JC, Schunemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation--determinants of a recommendation's direction and strength. *J Clin Epidemiol* 2013;66:726–35.
18. GRADE Working Group. Criteria for applying or using GRADE. [Web] 2010 [cited 2015 1 July]; Adopted from: <http://www.gradeworkinggroup.org/intro.htm>.
19. Schunemann HJ, Oxman AD, Brozek J et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;336:1106–10.
20. Morgenthaler TI, Aronsky AJ, Carden KA, Chervin RD, Thomas SM, Watson NF. Measurement of quality to improve care in sleep medicine. *J Clin Sleep Med* 2015;11:279–91.
21. Aurora RN, Collop NA, Jacobowitz O, Thomas SM, Quan SF, Aronsky AJ. Quality measures for the care of adult patients with obstructive sleep apnea. *J Clin Sleep Med* 2015;11:357–83.
22. Edinger JD, Buysse DJ, Deriy L, et al. Quality measures for the care of patients with insomnia. *J Clin Sleep Med* 2015;11:311–34.
23. Krahn LE, Hershner S, Loeding LD, et al. Quality measures for the care of patients with narcolepsy. *J Clin Sleep Med* 2015;11:335–55.
24. Trotti LM, Goldstein CA, Harrod CG, et al. Quality measures for the care of adult patients with restless legs syndrome. *J Clin Sleep Med* 2015;11:293–310.
25. Kothare SV, Rosen CL, Lloyd RM, et al. Quality measures for the care of pediatric patients with obstructive sleep apnea. *J Clin Sleep Med* 2015;11:385–404.
26. West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence. *Evid Rep Technol Assess (Summ)* 2002:1–11.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication October, 2015

Accepted for publication October, 2015

Address correspondence to: Timothy I. Morgenthaler, MD, Mayo Clinic, Rochester, MN; Email: research@aasmnet.org

DISCLOSURE STATEMENT

Dr. Morgenthaler is member of the American Academy of Sleep Medicine Board of Directors. Dr. Deriy, Mr. Heald and Dr. Thomas are employees of the American Academy of Sleep Medicine.