

REVIEW ARTICLE

Transparency Matters: Kaiser Permanente's National Guideline Program Methodological Processes

Carrie Davino-Ramaya, MD; L Kendall Krause, MD, MPH; Craig W Robbins, MD, MPH; Jeffrey S Harris, MD, MPH; Marguerite Koster, MA, MFT; Wiley Chan, MD; Gladys I Tom, MS

Abstract

Introduction: The practice-guideline process of collecting, critically appraising, and synthesizing available evidence, then developing expert panel recommendations based on appraised evidence, makes it possible to provide high-quality care for patients. Unwanted variability in the quality and rigor of evidence summaries and Clinical Practice Guidelines has been a long-standing challenge for clinicians seeking evidence-based guidance to support patient care decisions.

Methods: A multidisciplinary group of stakeholders, with representation from all eight Kaiser Permanente Regions, is responsible for creating National Guidelines. Conducting high-quality systematic reviews and creating clinical guidelines are time-, labor-, and resource-intensive processes, which raises challenges for an organization striving to balance rigor with efficiency. For these reasons, the National Guideline Program elected to allow for the identification, assessment, and possible adoption of existing evidence-based guidelines and systematic reviews using the ADAPTE; Appraisal of Guidelines Research and Evaluation; Assessment of Multiple Systematic Reviews (AMSTAR); and Grading of Recommendations Assessment, Development, and Evaluation (GRADE) frameworks. If no acceptable external guidelines are identified, the Guideline Development Team then systematically searches for relevant high-quality systematic reviews, meta-analyses, and original studies. Existing systematic reviews are assessed for quality using a measurement tool to assess systematic reviews (the AMSTAR systematic review checklist).

Study Appraisal: Following the screening and selection process, the included studies (the “body of evidence”) are critically appraised for quality, using the GRADE methodology, which focuses on four key factors that must be considered when assigning strength to a recommendation: balance between desirable and undesirable effects, quality of evidence, values and preferences, and cost. The evidence is then used to create preliminary clinical recommendations. The strength of these recommendations is graded to reflect the extent to which a guideline panel is confident that the desirable effects of an intervention outweigh undesirable effects (or vice versa) across the range of patients for whom the recommendation is intended.

Dissemination: The Care Management Institute disseminates all KP national guidelines to its eight Regions via postings on its Clinical Library Intranet site, a Web-based internal information resource.

Introduction

As the US health care system continues to undergo significant structural and financial change, evidence-based medicine—which, through clinical recommendations and practice guidelines, brings to the bedside the best available evidence of effective testing and treat-

ment—is used increasingly to reduce unwarranted variation in care and to form the basis for efficient, high-quality care. Physicians want to provide the best-quality care; the practice guideline process of collecting, critically appraising, and synthesizing available evidence, then developing expert panel

recommendations based on appraised evidence, makes this possible.

At times, national and international health care organizations and professional societies have issued conflicting recommendations about various clinical processes of care. Unfortunately, this inconsistency has made the identification and selection of high-quality clinical guidance a daunting task for any frontline clinician. It also has made cooperation in guideline development difficult, creating some confusion and contributing to the resource intensive nature of guideline development. Differences in the criteria and processes used to appraise and interpret the same body of evidence are a part of the problem, as are inconsistencies in the processes used to translate evidence into recommendations.

How can we, as consumers or clinicians, be certain of the quality and rigor of clinical practice guidelines (CPGs)? One would assume that guideline developers follow a standard, transparent protocol when searching for, evaluating, analyzing, synthesizing, and summarizing relevant data. One might also assume that expert guideline panels formulate recommendations in the same way. However, unwanted variability in the quality and rigor of evidence summaries and CPGs represents a long-standing challenge for clinicians seeking evidence-based guidance to support patient care decisions. Until recently, there have been no universally accepted standards for evidence summaries and guidelines, which has led to significant variability in the way guidelines are developed. Similarly, because there have not been common standards for documentation, there has been a lack of transparency in materials available for

Carrie Davino-Ramaya, MD, is a Practice Leader and Methodologist of Guidelines and Evidence-Based Medicine in the Quality Management and Systems Department for Northwest Permanente in Portland, OR. E-mail: carrie.m.davino@kp.org. **L Kendall Krause, MD, MPH**, is a Project Manager at the Institute for Health Research and an Evaluation Consultant for the Population and Prevention Services Department of Kaiser Permanente Colorado in Denver. E-mail: lauren.k.krause@kp.org. **Craig W Robbins, MD, MPH**, is the Medical Director of Evidence-Based Practice for the Care Management Institute and Guideline Director for the Colorado Permanente Medical Group in Denver. E-mail: craig.w.robbins@kp.org. **Jeffrey S Harris, MD, MPH**, is a Senior Physician with The Permanente Medical Group at the San Rafael Medical Center in CA. E-mail: jeffreys.harris@kp.org. **Marguerite Koster, MA, MFT**, is a Practice Leader for the Southern California Permanente Medical Group Technology Assessment and Guidelines Unit in Pasadena, CA. E-mail: marguerite.a.koster@kp.org. **Wiley Chan, MD**, is the Director of Guidelines and Evidence-Based Medicine for Northwest Permanente in Portland, OR. E-mail: wiley.chan@kp.org. **Gladys I Tom, MS**, is a Senior Project Manager for the Care Management Institute in Oakland, CA. E-mail: gladys.i.tom@kp.org.

review, making informed guideline comparison and adoption difficult.

Kaiser Permanente's (KP's) methodological approach to guideline development, together with the recently published standards of the Institute of Medicine (IOM),¹ may provide some assurances. The IOM standards call for the development of unbiased, scientifically valid, and trustworthy evidence reviews and CPGs. As KP and other guideline development organizations analyze and adopt these recommendations, it is not unreasonable to suggest that the new IOM standards might well form the basis against which transparency, objectivity, and consistency of guidelines will be measured (see Sidebar: Institute of Medicine Standards).

Kaiser Permanente's Guideline Development Infrastructure

KP was an early adopter of evidence-based methods. More than a decade of efforts by a network of KP clinicians, pharmacists, methodologists, evidence analysts, project managers, and other experts in evidence-based medicine has positioned KP as a national leader among health care organizations in developing evidence-based CPGs (see Sidebar: Definition of Clinical Practice Guidelines).^{2,3} To fully appreciate the link between KP and the IOM standards, it is worth briefly tracing the history of KP's guideline development program.

KP, the largest nonprofit integrated health care delivery system in the US, includes 8 geographic Regions in 9 states and the District of Columbia, covering more than 8.6 million members. KP's National Guideline Program (NGP) was established, first as an initiative, in 2005

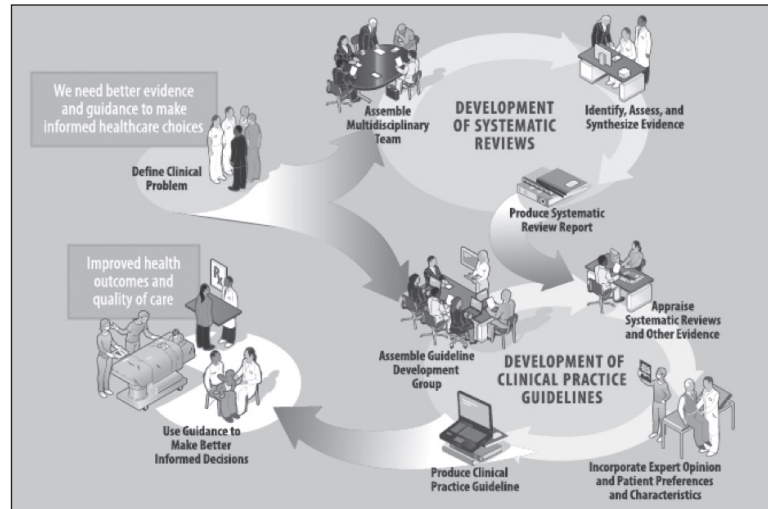


Figure 1. The Institute of Medicine approach to systematic reviews and clinical practice guidelines.¹

¹ Reprinted with permission: Interactive figure available from: www.iom.edu/Reports/2011/Clinical-Practice-Guidelines-We-Can-Trust.aspx; click on "Launch graphic."

by KP's Care Management Institute (CMI). The NGP receives direction and oversight from its National Guideline Directors (NGD), who represent the 8 Regions and Medical Groups of KP.

The CPGs, together with a variety of evidence-based clinician and patient tools linked to KP's electronic health record (EHR), provide evidence and reminders to ensure consistent, effective, and up-to-date evidence-based care. The NGP provides the organization with evidence-based recommendations to support care delivery, to help reduce unwarranted variation in care, and to improve clinical outcomes. To ensure that recommendations are framed consistently and accurately reflect the body of scientific evidence, the Guideline Quality (GQ) committee—a subcommittee of the National Guideline Directors—ensures that guideline development follows a set of rigorous, evidence-based, systematic, and transparent processes. Additionally, through active involvement in national and international organizations, members of the GQ committee remain current on evidence-based medicine and the guideline-development processes, and contribute to the ongoing development of the science of evidence-based practice.

In general, the need for evidence and/or guidance regarding a specific clinical problem prompts a search for existing evidence—most commonly in the form of a formal systematic review. Evidence is

then appraised and synthesized to create CPGs that provide guidance for clinical decisions. Guidelines based on well-conducted systematic reviews of the evidence provide an explicit linkage between the best evidence and clinical practice.

A Closer Look at the Institute of Medicine Standards

The IOM standards were developed in part in response to the *Medicare Improvements for Patients and Providers Act of 2008*⁴ when the US Congress asked the IOM to study and to report on best methods used to develop CPGs. The IOM then developed eight standards for the development of rigorous, trustworthy systematic reviews and CPGs (Figure 1).^{1,5} These two recent IOM reports—*Finding What Works in Health Care: Standards for Systematic Reviews*⁶ and *Clinical Practice Guidelines We Can Trust*¹—further describe detailed standards to increase the rigor of evidence reviews and guideline development, documentation, and reporting. Having received immediate attention on an international level, both reports raise the bar for conducting reviews of scientific evidence and developing evidence-based CPGs. It is expected that adherence to these standards will reduce bias, conflicts of interest, and variability in guideline development. Proponents also hope these standards present a unique opportunity to increase national and

Institute of Medicine Standards

The eight Institute of Medicine standards for the development of unbiased, scientifically valid, and trustworthy clinical practice guidelines (CPGs) include:

1. Establishing transparency
2. Management of conflict of interest
3. Guideline development group composition
4. CPG-systematic review intersection
5. Establishing evidence foundations for and rating strength of recommendations
6. Articulation of recommendations
7. External review
8. Updating

international collaboration for guideline development and dissemination, and will make the guideline development process less resource intensive.

Historical Perspective on Kaiser Permanente's National Guideline Development Program

KP's development of evidence-based CPGs began in 1991, when the Southern California Region hired David Eddy, MD, a pioneer in evidence-based medicine, to consult on the development of the Region's clinical guideline and technology assessment programs. At that time, most of KP's Regions had been developing practice guidelines based on expert consensus and review of selected studies. Dr Eddy's methodology—detailed in a manual developed in collaboration with the Council of Medical Specialty Societies Task Force on Practice Policies⁷—emphasized a rigorous and explicit approach based on systematic searching and selection of all available evidence for the topic of interest. He also emphasized the importance of critical appraisal of relevant studies, detailed estimation of an intervention's effect on important health outcomes, and an explicit descrip-

tion of the link between the evidence and eventual guideline recommendations.

By the mid-1990s, a growing interest in evidence-based guideline methodology, coupled with a desire to share guideline development resources across KP's Regions, led to the founding of the KP Interregional Guidelines Steering Group (IRGSG). Influenced by Dr Eddy's explicit approach, the IRGSG developed a position paper for interregional collaboration and a common methodology outlining principles and processes for evidence-based guideline development. By the late 1990s, under the sponsorship of CMI, the IRGSG evolved into the National Guideline Directors with representation from all Regions and agreement on a core set of nationally endorsed guidelines to be developed using rigorous evidence-based methods.

In 2010, KP's NGP became a member of the Guidelines International Network (G-I-N), a collaborative, international, not-for-profit association of organizations and individuals involved in development and use of CPGs. G-I-N seeks to improve the quality of health care by promoting systematic development of clinical guidelines and application of these guidelines in practice. Through G-I-N, the NGP was exposed to a number of international collaborative groups that had been developing frameworks and tools to improve the guideline development process, including:

- ADAPTE collaboration: provides a structured framework and systematic approach to adopt or adapt preexisting CPGs as an alternative to *de novo* guideline development.⁸
- Appraisal of Guidelines Research and Evaluation (AGREE) tool: used to assess the methodologic quality of existing CPGs.⁹
- Assessment of Multiple Systematic Reviews (AMSTAR) tool: used to evaluate the methodologic quality of systematic reviews.¹⁰
- Grading of Recommendations Assessment, Development, and Evaluation (GRADE): used to grade the quality of a body of evidence and the strength of recommendations.¹¹

Although KP's NGP had invested significant resources to create a unique internal guideline development process, the ADAPTE, AGREE, AMSTAR and GRADE frameworks offered opportunities and suggestions to make KP's guideline develop-

ment process more systematic, transparent, and explicit. Furthermore, the appropriate application of these tools allows KP to extend its rigorous processes to adopt or to adapt preexisting high-quality guidelines and evidence summaries, and to tailor recommendations for KP's specific cultural and organizational context.

In early 2010, backed by KP's CMI, guideline quality methodologists, and KP Regions, the National Guideline Directors agreed that the ADAPTE, AGREE, AMSTAR, and GRADE frameworks should be incorporated into KP's guideline methodology. To reflect these changes, the KP National Guideline GQ Committee revised the NGP methodology. *The Kaiser Permanente National Guideline Program Process and Methodology for Systematic Development of Clinical Practice Recommendations*¹² defines and describes the methods the NGP employs when creating CPGs. The following is a high-level overview of the KP National Guideline development processes.

Guideline Development and Methodology at Kaiser Permanente

KP employs an integrated, evidence-based, systematic, and transparent approach to the development of clinical guideline recommendations. This iterative process involves collection of data to create evidence-based resources, including CPGs and point-of-care decision-support tools within the EHR.

Following guideline implementation, care processes and outcome measures are compared with internal targets and

Guidelines based on well-conducted systematic reviews of the evidence provide an explicit linkage between the best evidence and clinical practice.

Definition of Clinical Practice Guidelines

Clinical practice guidelines (CPGs) have been part of the US health care system for over 30 years and should contain recommendations that are evidence based, defined as the "conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients."¹ Creating and applying CPGs uses a method commonly employed to codify best practices in a centralized and consistent manner. High-quality CPGs "articulate goals of care, enumerate potentially beneficial clinical approaches, and may reduce undesirable variation in care while supporting rational management of health conditions."²

References

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996 Jan 13;312(7032):71-2.
2. O'Connor PJ. Adding value to evidence-based clinical guidelines. *JAMA* 2005 Aug 10;294(6):741-3.

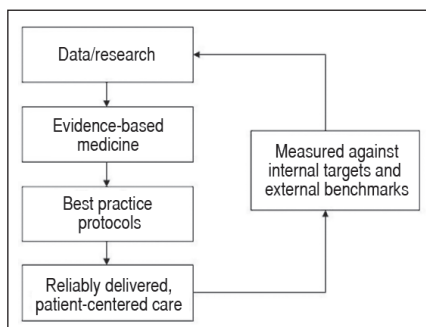


Figure 2. The Kaiser Permanente approach to evidence-based practice.

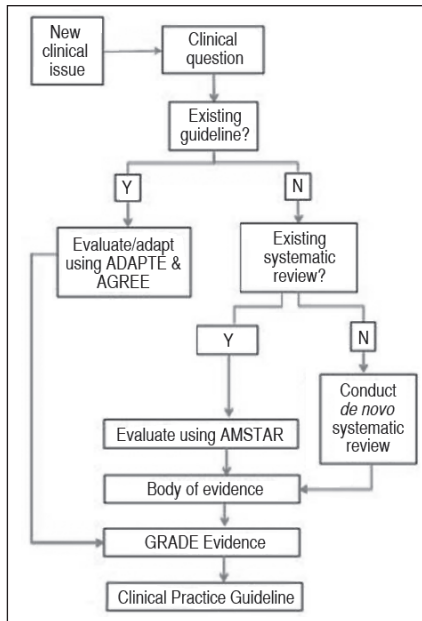


Figure 3. Schematic representation of Kaiser Permanente's process for guideline creation.

AGREE = Appraisal of Guidelines Research and Evaluation; AMSTAR = Assessment of Multiple Systematic Reviews; GRADE = Grading of Recommendations Assessment, Development, and Evaluation

external benchmarks; data are fed back into the system to further improve clinical practice (Figure 2).

A multidisciplinary group of stakeholders, with representation from all eight KP Regions, is responsible for creating National Guidelines. Each Guideline Development Team (GDT) includes physicians and other clinical experts (such as psychologists, pharmacists, clinical nurse experts, social workers, etc), evidence analysts, and a methodologist. The GDT serves as the expert panel that refines and approves the clinical recommendations that compose each guideline. To ensure that recommendations accurately reflect the body of scientific evidence, and are relevant to clinician and patient needs, the guideline development process follows the process described in Table 1 and schematically in Figure 3.

Choosing Topics for Clinical Guidelines: Challenges and Implications

Each year, the NGP evaluates and selects priority topics to be included in its guideline portfolio. Clinical questions to

be addressed are prioritized on the basis of an assessment of several internal and external factors. These factors may include:

- Quality-of-care concerns
 - Unwarranted variation in clinical and/or operational practice
 - Multiple treatment options
- Evolving evidence base
- High cost or resource use
- High prevalence of condition or risk factor
- Regulatory and/or accreditation requirements and metrics
- Strategic priorities
- Public interests
- Payer or employer group interests.

Each of the eight KP Regions has the opportunity to present topics of interest, which are then reviewed and voted on by the National Guideline Directors, taking into consideration the criteria above. If any KP primary or specialty care group presents topics that are not selected for inclusion in NGP priorities (eg, the topic doesn't meet prioritization criteria or resource availability), the group has the option to pursue the development of other practice support tools independent of the NGP.

Guideline Scope

Once a topic has been identified, a GDT is assembled; this team helps specify

Table 1. The Kaiser Permanente process for clinical practice guideline creation

Determine scope of the clinical content to be addressed in the guideline
Develop key clinical questions, including specification of patient populations, comparative interventions, and outcomes
Identify and evaluate existing recommendations and guidelines
Conduct a comprehensive search of relevant databases and other sources to identify relevant evidence
Screen, select, and extract data from studies
Perform a critical appraisal of the strengths and limitations of the identified studies
Assess, synthesize, and grade the body of evidence
Develop recommendations and rationales that are consistent with the evidence
Review recommendations
Approve guideline
Disseminate and implement guideline
Update underlying evidence periodically

Table 2. Adaptation of ADAPTE process used by the Kaiser Permanente Guideline Program

Preparation for the ADAPTE process <ul style="list-style-type: none"> • Identify Guideline Development Team (GDT) and staff • Solicit suggestions for external guidelines from the GDT or other subject matter experts
Define health questions (clinical questions)
Search and screen guidelines that address the clinical questions
Assess the identified guidelines, by clinical question, using the AGREE II tool, with an emphasis on Domain 3, <i>Rigor of Development</i> ⁹
Decide whether to adopt certain recommendations or the entire guideline
Draft modifications of recommendations or guidelines for specific needs and circumstances of Kaiser Permanente (KP) and its members
Draft a guideline report (this is a report to the GDT from the Lead Team consisting of the guideline assessment and recommendations for GDT consideration)
Other ADAPTE steps that may be used within the KP National Guidelines Program process include: <ul style="list-style-type: none"> • External review (for example, by Chiefs of relevant specialties) • Plan for updates and implementation • Guideline production and dissemination

AGREE = Appraisal of Guidelines Research and Evaluation

Table 3. AGREE domains of guideline quality and usability⁹

Scope and purpose
Stakeholder involvement
Rigor of development
Clarity of presentation
Applicability
Editorial independence

AGREE = Appraisal of Guidelines Research and Evaluation

the scope of the guideline, including target populations, comparative interventions, important outcomes, and other clinical issues. This process provides direction for framing specific clinical questions, which commonly address issues of risk, diagnosis, prognosis, therapy, and harm.

Traditionally, KP's CPGs have been created in-house; that is, the GDT maintained control of the entire guideline development process—from defining the clinical question through conducting systematic reviews, evaluating the evidence, and creating clinical recommendations within the framework of a clinical guideline—regardless of other preexisting external evidence reviews or guidelines.

External Guidelines

Conducting high-quality systematic reviews and creating clinical guidelines are time-, labor-, and resource-intensive processes, which raises challenges for an organization striving to balance rigor with efficiency. For these reasons, the NGP elected to allow for the identification, assessment, and possible adoption of existing evidence-based guidelines and systematic reviews using the ADAPTE, AGREE, AMSTAR and GRADE frameworks.

Using these frameworks, the NGP is able to evaluate the quality and applicability of preexisting external guidelines. For example, if KP were interested in creating a guideline related to human immunodeficiency virus (HIV) prevention and screening, the GDT might investigate whether any high-quality relevant guidelines or systematic reviews had previously been created or conducted by groups such as the Centers for Disease Control and Prevention, the US Preventive Services Task Force, the Cochrane Collaboration, etc.

If such guidelines are found, the guideline team uses the ADAPTE process to

assess congruence with identified clinical questions, analyze guidelines for quality, and make adaptations as needed to fit KP's context and needs (Table 2). Guidelines are further evaluated using the AGREE tool (part of the ADAPTE process), which evaluates guideline quality standards in six domains of guideline quality and usability (Table 3).

No External Guidelines Available

If no acceptable external guidelines are identified, the GDT then systematically searches for relevant high-quality systematic reviews, meta-analyses, and original

studies. Existing systematic reviews are assessed for quality using a measurement tool to assess systematic reviews (the AMSTAR systematic review checklist),¹³ another tool new to KP's methodology (Table 4). If systematic reviews are identified and deemed high quality by the AMSTAR checklist, the GDT may opt to use a preexisting review, rather than complete a systematic review *de novo*. In these cases, evidence analysts perform a supplementary search to identify any new studies that may have been published following the date of the preexisting review.

Table 4. AMSTAR review checklist¹⁰

1. Was an 'a priori' design provided? The research question and inclusion criteria should be established before the conduct of the review.
2. Was there duplicate study selection and data extraction? There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.
3. Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (eg, Central, EMBASE, and MEDLINE). Keywords and/or MESH terms must be stated and where feasible, the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.
4. Was the status of publication (ie, grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language, etc.
5. Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided.
6. Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions, and outcomes. The ranges of characteristics in all the studies analyzed (eg, age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported).
7. Was the scientific quality of the included studies assessed and documented? 'A priori' methods of assessment should be provided (eg, for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo-controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.
8. Was the scientific quality of the included studies used appropriately in formulating conclusions? The rigor and scientific quality of the methods used should be considered in the analysis and in the conclusions of the review, and explicitly stated in formulating recommendations.
9. Were the methods used to combine the findings of studies appropriate? For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (ie, χ^2 or I^2 tests for homogeneity). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (ie, is it sensible to combine?).
10. Was the likelihood of publication bias assessed? An assessment of publication bias should include a combination of graphical aids (eg, funnel plot, other available tests) and/or statistical tests (eg, Egger regression test).
11. Was the conflict of interest stated? Potential sources of support should be clearly acknowledged in both the systematic review and in the included studies.

Reprinted with permission from: Open access article: Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007 Feb 15;7:10. Available from: www.biomedcentral.com/1471-2288/7/10. AMSTAR = Assessment of Multiple Systematic Reviews

If no existing guidelines or systematic reviews are identified, the team conducts a full formal systematic review of the literature. This process includes the creation of a search strategy, database and hand searches for relevant studies, and evaluation of studies for relevance to the clinical question and the area of inquiry (eg, screening, diagnosis, prognosis, testing, treatment, or harms). Meta-analysis, or the statistical synthesis of results from independent studies, is often, though not always, employed. By combining the results from all relevant studies, this approach allows guideline developers to understand the combined weight and quality of the evidence, including direction and strength of association, and possible underlying heterogeneity (or differences in the estimates of effects between studies).

Grading Evidence Quality

Following the screening and selection process, the included studies (the “body of evidence”) are critically appraised for quality, using the GRADE¹¹ methodology. This approach is widely used internationally, and provides a systematic, transparent, and explicit framework to grade the qual-

ity of bodies of evidence. The higher the quality of the available evidence, the more confident one can be that an estimate of effect or association is close to the actual outcome of interest.

Steps in the GRADE approach include:¹⁴

- Identifying critical or important clinical outcomes
- Assessing the quality of evidence for each important outcome and across outcomes
- Formulating recommendations based on quality of evidence, balance of benefits and harms, patient values and preferences, and resource and cost implications.

The quality of evidence for each important outcome includes consideration of six elements: study design, study quality, consistency, directness, precision, and possible reporting bias. The GRADE system classifies the quality of evidence into the following four levels with corresponding implications:

- High quality—further research is very unlikely to change confidence in the estimate of effect
- Moderate quality—further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate

- Low quality—further research is very likely to have an important impact on confidence in the estimate of effect and is likely to change the estimate
- Very low quality—any estimate of effect is very uncertain (Table 5).¹⁵

When based on randomized control trials, the body of evidence is initially categorized as high quality. However, the evidence grade may be reduced for reasons including study limitations, inconsistency of results, indirectness, imprecision, and publication bias. When based on observational studies (eg, cohort or case-control studies), the body of evidence is initially categorized as low quality. However, the evidence grade may be increased if, for example, the magnitude of treatment effect is very large or if there is evidence of a dose-response relationship.

Translating Evidence into Recommendations

The evidence is then used to create preliminary clinical recommendations. The strength of these recommendations is graded to reflect the extent to which a guideline panel is confident that the desirable effects of an intervention outweigh undesirable effects (or vice versa) across the range of patients for whom the recommendation is intended. GRADE focuses on four key factors that must be considered when assigning strength to a recommendation: balance between desirable and undesirable effects, quality of evidence, values and preferences, and cost (Table 6).¹⁵

The final outcome of this process is the creation of evidence-based clinical recommendations that address the clinical questions posed at the outset of the process. A recommendation may contain one or more parts. Rationale statements—explicitly addressing the four GRADE strength of recommendation domains—are written to explain the logic that links the evidence synthesis and GRADE analysis to the appropriate recommendation (Table 7).

Guideline Dissemination and Implementation

Following the completion of a CPG, next steps in KP’s process include dissemination and implementation activities.

Grade	Definition
High	Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low	Any estimate of effect is very uncertain.

GRADE = Grading of Recommendations Assessment, Development, and Evaluation

Factor	Comment
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted.
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted.
Values and preferences	The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted.
Costs (resource allocation)	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted.

CMI disseminates all KP national guidelines to its eight Regions via postings on its Clinical Library Intranet site,¹⁶ a Web-based internal information resource. From the Clinical Library, physicians and health care professionals throughout KP can then access the guidelines.

Ideally, the goal of guideline implementation is to enhance the likelihood of successful changes in practice, specifically the increased application of evidence-based practices. KP’s multifaceted approach to guideline implementation (Figure 4) includes, as the central feature of its process, its integrated EHR, KP HealthConnect. This allows KP to provide point-of-care support and guidance for clinicians.

Guidelines are embedded in the process of care delivery through KP HealthConnect standard orders, alerts and clinical tools, program protocols, clinician reminders, member education and outreach, and other tools. Generally, these tools are developed by each Region to meet local needs. Prompts within the EHR or “SmartRx” prompts provide explicit guidance regarding pharmacologic choices; best practice alerts provide active reminders to clinicians and patients; “SmartSets” are standard order sets that provide comprehensive recommendations (eg, lab tests, medications, etc) for a variety of illnesses and disease states.

A variety of additional decision support tools and summary documents,

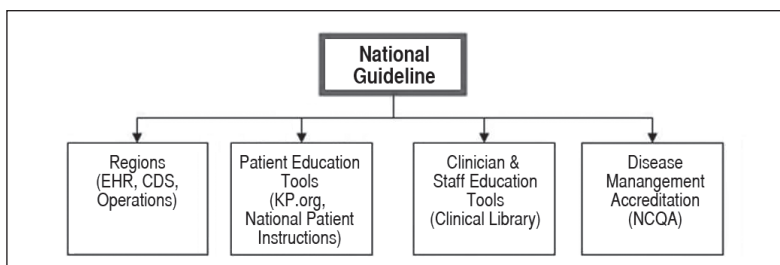


Figure 4. Guideline implementation at Kaiser Permanente.

CDS = clinical decision support; EHR = electronic health record; NCQA = National Committee for Quality Assurance

including guideline change documents, diagnosis-specific quick reference guides, clinician and staff educational materials, and patient education materials may be developed to support recommendations. Guideline change documents outline any changes as compared with previous versions of a CPG.

Monitoring and Auditing Criteria

KP has the capacity to monitor and evaluate clinician adherence to clinical guidelines as well as to monitor effects on health care outcomes. The “measurable outcomes” factor—defined as the degree to which the guideline identifies markers or endpoints to track the effects of implementation of a recommendation—is especially salient in arenas with established monitoring criteria (eg, Healthcare Effectiveness Data and Information Set and Joint Commission metrics).

Support of Quality-Improvement Initiatives

KP’s CPGs are intended not only to enhance patient and clinical decision making but to improve health care outcomes and quality of care, to meet state and federal regulatory requirements, and to support voluntary organization accreditation and internal quality improvement and patient safety initiatives. Through the use of KP patient data, clinical care gaps can be identified, performance and quality goals developed, and specific initiatives implemented through clinical guideline recommendations that address best practices.

Conclusion

If you are a user or developer of guidelines, a consumer or clinician, or are an administrator or operations specialist, transparency regarding CPGs and how they are developed should matter to you.

Table 7. Rationale/decision table example: aspirin therapy in the general primary prevention population					
Subgroup	Quality of evidence	Balance of benefits vs harms and burdens	Values and preferences	Resource implications	Recommendation
General primary prevention population	High—no serious limitations	These analyses show a modest benefit overall. However when investigated independently, evidence fails to show the benefit of aspirin in healthy patients without a history of diabetes, stroke/TIA or CKD. There may be minimal benefit of aspirin in patients with low risk for CVD, although a small risk of major bleeding would also exist. At <10% 10-year CVD risk, the GDT believes risk would outweigh benefits. As risk for CVD increases, the benefit of aspirin outweighs the risk of major bleeding.	The GDT believes, as risk for CVD increases, patients would generally value the potential mortality benefit of aspirin higher than the potential risk of serious bleeding. Because of variation in CVD risk, there would likely be significant variability in patient acceptance of aspirin therapy.	Low	For patients with no established CVD, atherosclerotic cerebrovascular disease, noncoronary atherosclerosis, type 2 diabetes, or CKD; clinicians: 1. Should prescribe aspirin to those with >20% 10-year risk for CVD. (Strong Recommendation) 2. May prescribe aspirin to those with 10% to 20% 10-year risk for CVD. (Weak Recommendation) 3. Should probably not prescribe aspirin to those with <10% 10-year risk for CVD. (Weak Recommendation)

CAD = coronary artery disease; CKD = chronic kidney disease; CVD = cardiovascular disease; GDT = guidelines development team; TIA = transient ischemic attack

The recently published IOM standards¹ are intended to reduce variability and to standardize the level of rigor and quality that goes into developing CPGs and in performing systematic reviews of the evidence. KP's new methodology meets the majority of these standards.

In reviewing KP's CPG methodology, it is apparent that the NGP is well-aligned with the following IOM standards: establishing transparency, management of conflict of interest, multidisciplinary and balanced guideline development group composition, effective management of the CPG and systematic review intersection, assessing evidence foundations and assigning strength of recommendations, and articulation and updating of recommendations.

Because the NGP has been developed primarily to create CPGs for use within its integrated delivery system, KP has not previously pursued patient and consumer involvement in guideline development and review. However, the National Guideline Directors are currently exploring options for meeting these standards and standards related to the development and public posting of systematic reviews.

This article demonstrates that the goals of KP's NGP are reflected in the new methodology as well as in the CPGs that are produced at the national level. This new and rigorous methodology provides the framework for developing CPGs that end users can trust in their medical care decisions.

Although challenging, the new IOM standards present guideline developers

and clinicians with several opportunities. If other organizations continue to improve the rigor and transparency of their processes, it is likely that sharing high-quality systematic reviews and guidelines will increase. Furthermore, increased collaboration between developers of systematic reviews and CPGs could influence public funding and allocation of resources to address health topics that are the most important to patients, clinicians, and health care delivery systems. ♦

Disclosure Statement

The author(s) have no conflicts of interest to disclose.

References

- Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E, editors; Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. *Clinical practice guidelines we can trust*. Washington, DC: The National Academies Press; 2011.
- Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996 Jan 13;312(7032):71-2.
- O'Connor PJ. Adding value to evidence-based clinical guidelines. *JAMA* 2005 Aug 10;294(6):741-3.
- Medicare Improvements for Patients and Providers Act of 2008, Pub L No. 110-275 (July 15, 2008).
- Eden J, Levit L, Berg A, Morton S, editors; Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. *Finding what works in health care: standards for systematic reviews*. Washington, DC: The National Academies Press; 2011.
- Finding what works in health care: Standards for systematic reviews. Washington, DC: The National Academies Press; 2011.
- Eddy DM. *A manual for assessing health practices and designing practice policies: the explicit approach*. Philadelphia, PA: American College of Physicians; 1992.
- Adapte.org [homepage on the Internet]. The Adapte Collaboration; 2007 [cited 2012 Jan 31]. Available from: www.adapte.org.
- Brouwers M, Kho ME, Browman GP, et al. AGREE II: Advancing guideline development, reporting and evaluation in healthcare. *CMAJ*. 2010 Dec 14;182(18):E839-42.
- Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007 Feb 15;7:10.
- Gradeworkinggroup.org [homepage on the Internet]. GRADE Working Group; 2005-2011 [cited 2012 Jan 31]. Available from: www.gradeworkinggroup.org/index.htm.
- Kaiser Permanente National Guideline Program Process and Methodology for Systematic Development of Clinical Practice Recommendations [monograph on the Intranet]. Oakland, CA: Kaiser Permanente; 2011 Mar 15 [cited 2012 Feb 8]. Available from: <https://wiki.kp.org/wiki/download/attachments/102107054/NGP+Methodology+3-15-11.pdf?version=1&modificationDate=1305579496000;passwordprotected>.
- Stuck AE, Kharicha K, Dapp U, et al. Development, feasibility and performance of a health risk appraisal questionnaire for older persons. *BMC Med Res Methodol* 2007 Jan 11;7:1.
- Atkins D, Best D, Briss PA, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004 Jun 19;328(7454):1490.
- Guvatt GH, Oxman AD, Kunz R, et al; GRADE Working Group. Going from evidence to recommendations. *BMJ* 2008 May 10;336(7652):1049-51. Erratum in *BMJ* 2008 Jun 21;336(7658):doi:10.1136/bmj.a402.
- Clinical library [Web page on the Intranet]. Oakland, CA: Care Management Institute; 2012 [2012 Jan 19]. Available from: <http://cl.kp.org;passwordprotected>.

Anecdotalism

The danger as always is anecdotalism, the inherent deception of drawing a broad generalization from very few cases. A pejorative cloud hangs over the anecdote in medicine. Exceptions are allowed: there seems to be little danger in the illustrative use of a similar case during the discussion of a differential diagnosis or in the citation of a notable exception as a cautionary example that marks the limits of a topic under discussion.

— *Doctor's Stories*, Kathryn Montgomery Hunter, PhD,
Professor of Medicine and Medical Ethics and Humanities