



Contents lists available at ScienceDirect



Practice parameters and strength of recommendation data: a variable compass

Taylor Banks, MD; Julia Savitz, MD; and Michael R. Nelson, MD, PhD

Division of Allergy-Immunology, Walter Reed National Military Medical Center, Bethesda, Maryland

ARTICLE INFO

Article history:

Received for publication February 18, 2014.

Received in revised form April 23, 2014.

Accepted for publication April 29, 2014.

ABSTRACT

Background: Practice parameters and guidelines shape and influence the method and manner in which medicine is practiced. With more than 121 scales and methods of assessing and rating evidence, a comparison of practice parameters can appear daunting. An evaluation of the evidence engenders a sense of the evolution of a specialty and a roadmap for the future.

Objective: To assess the level of evidence underlying recommendations in allergy-immunology (AI) practice parameters.

Methods: We analyzed the practice parameters that guide AI (n = 15), otolaryngology (n = 8), pediatrics (n = 13), and internal medicine (n = 10) as they appeared on August 30, 2012. Strength of recommendation data was compared after making adjustments for differences in rating scales.

Results: The strength of recommendation calculated from strong to weak for the AI practice parameters using a standardized format yielded the following grades: A in 195 (13.9%), B in 342 (24.4%), C in 606 (43.2%), D in 231 (16.4%), and E in 29 (2.1%). Controlled trial–based evidence (A and B) demonstrated considerable variability among individual AI practice parameters (range, 1.3%–100%). Evidence from controlled trials was lower in the subspecialty fields (38.3% in AI and 38.2% in otolaryngology) compared with the primary care fields (55.6% in pediatrics and 86.1% in internal medicine).

Conclusion: Considerable variability exists in the strength of recommendations within the AI practice parameters. The guidelines created by the primary care fields rest on a larger base of evidence collected from controlled trials. These findings likely reflect the adopted approach of making recommendations for less well-studied conditions and practices in AI to assist practitioners and patients and at the same time highlight the myriad opportunities for future research.

© 2014 American College of Allergy, Asthma & Immunology. Published by Elsevier Inc. All rights reserved.

He who loves practice without theory is like the sailor who boards ship without rudder or compass and never knows where he may be cast.

—attributed to Leonardo da Vinci

Introduction

As with da Vinci's rudder and compass, practice parameters and guidelines shape and influence the method and manner in which medicine is practiced, helping to plot and maintain a course for the medical practitioner. The earliest commentaries addressing practice parameters note that these documents are developed to refine topics and to facilitate visibility of new evidence and best

practices.^{1,2} Ultimately, practice parameters seek to bring clarity and consistency to areas of confusion and variability, rigorously analyzing and synthesizing available data to provide guidance to practitioners.

Numerous impediments to evidence-based practice have been cited in the literature. Among these is the deluge of published evidence, often of uncertain value, that besets practitioners. Randomized controlled trials have increased from an estimated 5,000 publications annually in 1978–1985 to 25,000 publications annually in 1994–2001.^{3,4} Although practice parameters arose in part to alleviate the burden posed by this marked increase in available information, long-standing criticism has been directed at the variable methods used in crafting assessments and recommendations.^{5–7} This variability has been exacerbated by the proliferation of rating scales, which are estimated to now exceed 120 different iterations.⁸

The bewildering array of evidence and grading systems led to a congressional mandate in 2008 to the Institute of Medicine (IOM) to study standards for developing trustworthy guidelines.^{3,9} This report emphasized the implementation of rigorous standards,

Reprints: Taylor Banks, MD, Division of Allergy-Immunology, Walter Reed National Military Medical Center, 8901 Wisconsin Ave, Bethesda, MD 20889; E-mail: taylor.a.banks@health.mil.

Disclaimer: The views expressed are those of the authors and do not reflect the official policy of the US Department of Army, US Department of Navy, US Department of Defense, or the US government.

Disclosures: Authors have nothing to disclose.

Table 1
Modified grading system^a

Strength	Underlying evidence
A	Well-designed, randomized controlled trials or meta-analyses of controlled trials
B	Controlled trials with minor flaws or based on consistent evidence from observational studies
C	Observational studies (case-control, comparative studies, correlation studies) or extrapolated from other studies
D	Committee reports or based on expert opinion
E	Exceptional situations or not rated

^aThis table is based on the scale of Shekelle et al.¹⁰ and depicts the method used to standardize strength of evidence recommendations and reconcile differences among scales.

highlighted the need for transparency and clarity, and discussed the role practice parameters serve in directing attention to future research needs.

Allergy-immunology (AI) practice parameters have undergone their own changes, although the method of objectively evaluating evidence remains tied to the grading system proposed by Shekelle et al.¹⁰ in 1999. An evaluation of the evidence underlying the AI practice parameters not only engenders a sense of how the specialty has and continues to evolve but also provides a roadmap for the future.

Methods

We hypothesized that AI's unique position, at both the bench and bedside, would be reflected in lower levels of evidence underpinning the recommendations made in its practice parameters. This is in part due to the inclusion of recommendations for complex and rare conditions intended to increase quality outcomes and the use of best practices. We further sought to identify the specific factors associated with variance in such levels compared with other fields. Finally, using the 2008 Medicare Improvements for Patients and Providers Act and a subsequent IOM study on practice parameters as a potential watershed, we strove to evaluate what, if any, effect this has had on AI practice parameters published after this congressional action.

To provide deeper insight into the variables that might inform an analysis of the evidence and methods that frame the AI practice parameters, we included 3 additional specialties in our analysis: otolaryngology (a related subspecialty), internal medicine (a primary care field), and pediatrics (a primary care field). Although the primary objective remained an evaluation of the evidence and variability within the AI practice parameters, the inclusion of these 3 fields offered an opportunity to undertake additional exploration. Contrasts could be drawn not only among fields but also more generally between subspecialties and primary care and adult and pediatric populations.

Our analysis began by querying the Agency for Healthcare Research and Quality National Guideline Clearinghouse website and the websites for the American Academy of Allergy, Asthma, and Immunology, the American Academy of Otolaryngology–Head and Neck Surgery, the American College of Physicians, and the American Academy of Pediatrics. Practice parameters were evaluated as published on these websites as of August 30, 2012. We included only the current or most recent version of the published parameters. Similarly, to avoid overlapping data, we excluded publications jointly sponsored or endorsed by specialties considered in this analysis. Finally, practice parameters that did not indicate the strength of recommendations were excluded from consideration.

As has been noted, myriad scales for assessing evidence exist, complicating a direct comparison of the evidence underpinning the

specialties' practice parameters. Any differences in the scales were reconciled using the system of Shekelle et al.¹⁰ which has long been used in the AI practice parameters (Table 1). Use of this approach ensured continuity in assessing the data within AI and comparisons with the other fields. With the use of this grading system, the strength attached to each recommendation made in each of the field's practice parameters was tabulated and the resulting data set analyzed.

Results

Using the above inclusion and exclusion criteria, we identified 15 practice parameters in AI, 8 in otolaryngology, 10 in internal medicine, and 13 in pediatrics (Table 2). We identified wide variability in the proportion of parameters that were based on controlled trials. The absolute number of recommendations from field to field also varied widely, with AI containing the vast preponderance of recommendations ($n = 1403$) followed by pediatrics ($n = 291$) and more distantly by otolaryngology ($n = 110$) and internal medicine ($n = 38$) (Fig 1).

With the use of the grading scale detailed in Table 1, an analysis of the strength of recommendations within the AI practice parameters yielded the following grades: A in 195 (13.9%), B in 342 (24.4%), C in 606 (43.2%), D in 231 (16.4%), and E in 29 (2.1%). After simplification of the approach to these data, AI recommendations with at least some basis in controlled trials (A and B) represented 38.3% of the total. Marked variability among individual AI practice parameters was also evident, with a range in controlled trial–based evidence of 1.3% to 100% (Fig 2). Five of the practice parameters in AI fall below the field's mean for controlled trial–based recommendations: (1) Practice Parameter for the Diagnosis and Management of Primary Immunodeficiencies (1.3%), (2) Drug Allergy: An Updated Practice Parameter (17.7%), (3) The Diagnosis and Management of Anaphylaxis Practice Parameter (30.6%), (4) The Diagnosis and Management of Rhinitis: An Updated Practice Parameter (36.4%), and (5) Disease Management of Atopic Dermatitis: An Updated Practice Parameter (37.0%).

In analyzing those AI practice parameters published before and after the 2008 congressional mandate to the IOM, similar proportions of parameters fell above and below the specialty mean for controlled trial–based recommendations. Three of 5 parameters (60.0%) that fell below the specialty mean for controlled trial–based recommendations were published concurrently with or after the IOM study. Similarly, 6 of 10 parameters (60.0%) above the specialty mean were published at the same time as or after the IOM study.

In comparing AI to other fields, the proportion of controlled trial–based recommendations varies significantly, with otolaryngology having a similar percentage (38.2%) and the primary care fields presenting much higher figures (55.6% in pediatrics and 86.1% in internal medicine; Fig 3). A deeper analysis reveals more nuanced findings. The most rigorous level of evidence (A) is similar among the parameters for AI (13.9%), internal medicine (13.1%), and pediatrics (18.9%), whereas otolaryngology's parameters have a much lower percentage (5.4%). Also of note, expert evidence and opinion (D) was highest in pediatrics (25.1%) and AI (16.4%).

Discussion

Four recommendations contained in the IOM report are of particular interest in evaluating the data presented and looking toward future directions for practice parameter development (Table 3). These critical ideas emphasize rigor in reviewing and summarizing evidence coupled with clarity and transparency in presenting that evidence. These concepts have garnered substantial interest and commentary within the medical literature and are of significant import in addressing the results of our study.^{7,9,11}

Table 2Data from evaluated practice parameters^a

Practice parameter	Controlled trial–based recommendations, %	Year published	No. of recommendations
Allergy-Immunology			
Primary immunodeficiencies	1.3	2005	224
Drug allergy	17.7	2010	198
Diagnosis and management of anaphylaxis	30.6	2010	98
Diagnosis and management of rhinitis	36.4	2008	110
Disease management of atopic dermatitis	37.0	2004	54
Food allergy	41.7	2006	108
Allergen immunotherapy	42.8	2011	98
Influenza vaccine and egg allergic recipients	42.8	2010	7
Stinging insect hypersensitivity	48.1	2011	27
Diagnosis and management of sinusitis	58.4	2005	77
Allergy diagnostic testing	60.0	2008	250
Exercise-induced bronchoconstriction	62.3	2010	53
Contact dermatitis	76.0	2006	75
Adverse reactions to vaccines	76.9	2012	13
Attaining optimal asthma control	100	2005	11
Otolaryngology			
Polysomnography for sleep-disordered breathing before tonsillectomy in children	0	2011	6
Benign paroxysmal positional vertigo	21.4	2008	14
Cerumen impaction	25.0	2008	16
Acute otitis externa	30.0	2006	10
Hoarseness (dysphonia)	43.7	2009	16
Tonsillectomy in children	43.7	2010	16
Adult sinusitis	50.0	2007	18
Sudden hearing loss	64.3	2012	14
Pediatrics			
Hyperbilirubinemia in the newborn infant of ≥ 35 weeks of gestation	14.7	2004	34
Management of sinusitis	40.0	2001	5
Guidelines for adolescent depression in primary care	42.8	2007	28
Otitis media with effusion	46.7	2004	15
Childhood obstructive sleep apnea syndrome	57.1	2002	7
Diagnosis and management of acute otitis media	57.1	2004	7
Diagnosis and management of the initial urinary tract infection in febrile infants and children	60.0	2011	10
Pharmacologic treatment of migraine headache in children and adolescents	60.0	2005	10
Prevention of infective endocarditis	60.0	2007	15
Prevention of intravascular catheter-related infections	64.5	2002	121
Diagnosis and management of bronchiolitis	66.7	2006	21
Neurodiagnostic evaluation of the child with a simple febrile seizure	71.4	2011	7

(continued)

Table 2 (continued)

Practice parameter	Controlled trial–based recommendations, %	Year published	No. of recommendations
Attention-deficit/hyperactivity disorder in children and adolescents	90.9	2011	11
Internal Medicine			
Current pharmacologic treatment of dementia	33.3	2008	3
Hormonal testing and pharmacologic treatment of erectile dysfunction	33.3	2009	3
Screening for osteoporosis in men	66.7	2008	3
Improve palliative care of pain, dyspnea, and depression at the end of life	80.0	2008	5
Diagnosis and management of stable chronic obstructive pulmonary disease	85.7	2011	7
Diagnosis and treatment of low back pain	100	2007	4
Intensive insulin therapy for the management of glycemic control in hospitalized patients	100	2011	3
Oral pharmacologic treatment of type 2 diabetes mellitus	100	2012	3
Using second-generation antidepressants to treat depressive disorders	100	2008	4
Venous thromboembolism prophylaxis in hospitalized patients	100	2011	3

^aAllergy-immunology parameters below the controlled trials–based mean (38.3%) are in bold.

Confirming our initial hypothesis, AI's practice parameters had a lower number of recommendations based on controlled trials compared with the practice parameters of the primary care fields. Otolaryngology's practice parameters reflected a similar, lower level of controlled trial–based recommendations. AI and otolaryngology's smaller base of controlled trial evidence is likely reflective of these fields' specialized nature. As subspecialties, both treat more complex and less commonly encountered disease processes compared with those addressed in the primary care fields. Despite this, it is notable that the proportion of AI recommendations with the strongest (A) evidence is similar to that found in pediatrics and actually exceeds the percentage observed in the internal medicine practice parameters.

Additional review of our study's results highlights several areas that merit further consideration. The first is the cause of the variability in the proportion of controlled trial–based recommendations noted within AI practice parameters and compared with other fields' parameters. Two likely contributing factors that are somewhat related are that many of the parameters address broadly or poorly defined conditions and that many of these conditions lack well-established, accurate, and reliable means of testing or evaluation. Similarly, particularly among the subspecialty practice parameters, rare conditions with small numbers of patients will undoubtedly have less available evidence in the medical literature with which to develop recommendations. Finally, those conditions with significant pediatric populations may affect the proportion of controlled trial–based recommendations in practice parameters that address this population or conditions that predominantly affect children, given the impediments and difficulties involved in pediatric research.

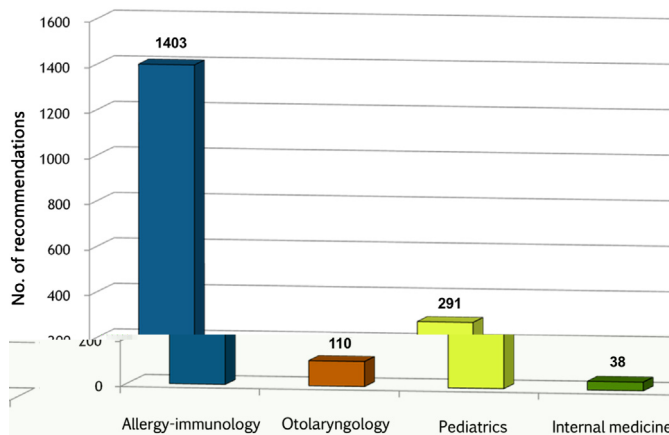


Figure 1. Absolute numbers of recommendations contained within each field's practice parameters.

The second area is the marked differences observed in the number of recommendations made within each specialty's practice parameters. AI far outstrips the other fields in that its practice parameters contain more than 4 times ($n = 1403$) the number found in the nearest specialty ($n = 291$ in pediatrics). Furthermore, a deeper analysis reveals that the 5 AI parameters that fell below the specialty mean for controlled trial–based recommendations contain 48.8% of the total number of recommendations while representing only one-third of the total practice parameters. With the exception of the atopic dermatitis practice parameter, the remaining 4 parameters fall in the top 6 for the field in terms of absolute number of recommendations (Table 2). One potential explanation is that these larger numbers of recommendations reflect a roadmap for practitioners, attempting to bring clarity in less frequently practiced specialty areas and to provide clear boundaries while preserving greater individual discretion among practitioners. As discussed earlier, this may find some basis in those conditions with poorly understood mechanisms of disease. Another explanation of the observed difference in number of recommendations is the possible philosophical differences among specialty guideline task forces. Examples of possible differences in approach to developing recommendations include entire vs partial scope of practice, recommendations for any vs only those practices with controlled trial evidence, and combined vs discrete aspects of clinical practice.

A third area that warrants additional scrutiny is the role of expert opinion in practice parameters. A higher portion is observed in both AI and pediatrics. This may in part stem from difficulties in

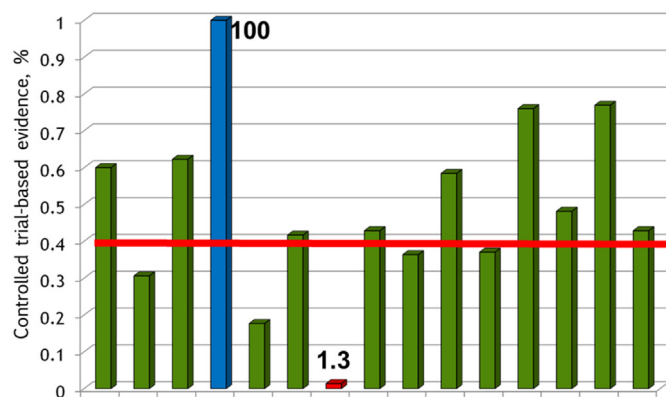


Figure 2. Percentage of controlled trial–based evidence by individual allergy-immunology practice parameters. The horizontal line represents the specialty's mean for controlled trial–based evidence (A and B evidence, 537/1,403 [38.3%]).

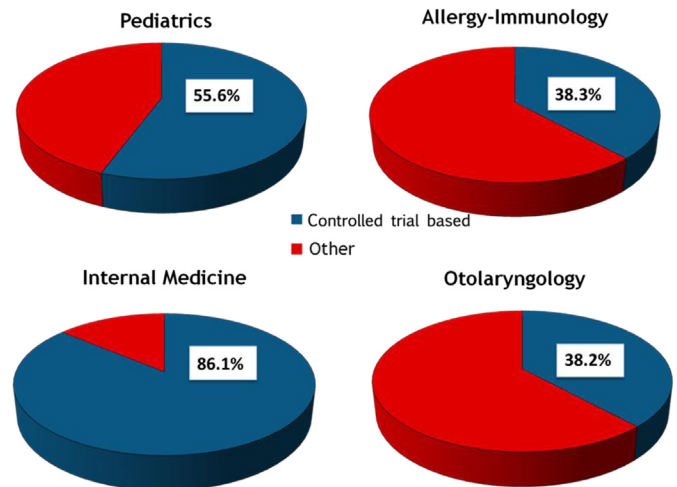


Figure 3. Comparison of the percentage of controlled trial–based (strength of recommendation A and B) recommendations contained in each field's practice parameters.

study design or approval in the pediatric population or with therapeutic modalities. It may also indicate deficits in the understanding of underlying mechanisms or rarity of the conditions addressed in the practice parameters. Although expert opinion and committee reports are more generally regarded as embodying lower levels of evidence, newer evaluation systems, such as the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE), offer an emerging alternative in which qualitative appraisals of studies allow more balanced assessments, particularly in areas where randomized controlled trials are not always a feasible option.^{11–14}

Several limitations of our study should be noted. The first is that the study does not represent a comprehensive review of practice parameters because a wide range of specialties are not included. Instead, we present a sampling of AI, a related subspecialty, and 2 primary care fields. The second is that our data were not normalized for statistical analysis. This was a purposeful decision because we hoped to present the recommendations in situ as they are available to practitioners. Although there are several iterations of a number of the existing practice parameters, we only present data from the most current form, again examining the parameters' recommendations because they are generally available to practitioners, rather than presenting an evolutionary or developmental view. Third, as has been noted, proposed changes in the guidance that shapes practice parameters are being implemented. Dedication to continuous improvement is a recognized goal of the AI Practice Parameters Task Force, and these changes continue an aggressive approach to address the well-studied and understudied clinical conditions that practitioners face.

This study represents an initial foray into this topic because it is a rich area for future research and an ongoing dialogue within our

Table 3

Institute of Medicine Key standards for guidelines: a selection of critical standards from the institute of medicine report on trustworthy guidelines

Key standards for a trustworthy guideline

Rigorous systemic evidence review considering quality, quantity, and consistency of available evidence

Summary of evidence relevant to each recommendation

Transparency regarding the role of opinion, theory, and clinical experience

Provides a rating of the strength of evidence underlying each recommendation

Modified from Laine et al.⁹

specialty and more generally as a medical community. As suggested by one of our study's limitations, we hope to further evaluate how the recommendations within the practice parameters have and continue to evolve. This includes evaluation of more discrete new parameters, such as those that address hypersensitivity to furry animals and rodents, and ongoing updates, such as the atopic dermatitis practice parameter.^{15–17} Another avenue for future consideration is a comparison of systems such as that used by the AI practice parameters for evaluating evidence and those that use newer frameworks, such as GRADE.

These findings should not be viewed as detrimental. The significant expertise and work embodied in the development of existing practice parameters represent a guiding force and an important contribution to medical knowledge and practice. As was part of the original intent of practice parameters, our study suggests that these publications continue to serve as a means to highlight areas for future research. This can aid all fields in identifying and targeting opportunities to further enlarge our understanding of the diseases and processes that we treat and will in turn strengthen future practice parameters. Although variation within parameters and a smaller proportion of controlled trial–based recommendations exist in AI practice parameters, this largely reflects the field's specialized and still evolving nature. These important documents continue to strive to ensure consistency of practice within defined boundaries and simultaneously protect individual discretion among health care professionals to offer personalized health care. This is bolstered by the commensurate portion of A-level evidence compared with the primary care fields, suggesting a strong commitment to rigorous, evidence-based practice. New and training allergists along with seasoned investigators should be heartened in that they can play a significant and meaningful role in the way in which AI is practiced and in unlocking its underlying biological mechanisms through research.

References

- [1] Eddy DM. Practice policies: where do they come from? *JAMA*. 1990;263:1265.
- [2] Rothschild RD. Practice parameters benefit all: patient management strategies improve outcomes for patients, lower costs for payers. *Health Progress*. 1993;74:24–29.
- [3] Graham R, Mancher M, Wolman DM, et al. *Clinical Practice Guidelines We Can Trust*. 2011. Washington, DC: National Academies Press; 2011 http://www.nap.edu/catalog.php?record_id=13058. Accessed October 30, 2012.
- [4] Davidoff F, Miglus J. Delivering clinical evidence where it's needed: building an information system worthy of the profession. *JAMA*. 2011;305:1906–1907.
- [5] Hasenfeld R, Shekelle PG. Is the methodological quality of guidelines declining in the US? *Qual Saf Health Care*. 2003;12:428–434.
- [6] Rosenfeld RM, Shiffman RN. Clinical practice guideline development manual: a quality-driven approach for translating evidence into action. *Otolaryngology*. 2009;140:S1–S43.
- [7] Ransohoff DF, Pignone M, Sox HC. How to decide whether a clinical practice guideline is trustworthy. *JAMA*. 2013;309:139–140.
- [8] Steering Committee on Quality Improvement and Management. Classifying recommendations for clinical practice guidelines. *Pediatrics*. 2004;114:874–877.
- [9] Laine C, Taichman DB, Mulrow C. Trustworthy clinical guidelines. *Ann Intern Med*. 2011;154:774–775.
- [10] Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Developing guidelines. *BMJ*. 1999;318:593–596.
- [11] Mustafa RA, Santesso N, Brozek J, et al. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *J Clin Epidemiol*. 2013;66:736–742.
- [12] Aurora RN, Morgenthaler TI. On the goodness of recommendations: the changing face of practice parameters. *Sleep*. 2010;33:1273–1276.
- [13] Canfield SE, Dahm P. Rating the quality of evidence and strength of recommendations using GRADE. *World J Urol*. 2011;29:311–317.
- [14] Guyatt GH, Oxman AD, Schunemann HJ, et al. GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. *J Clin Epidemiol*. 2011;64:380–382.
- [15] Portnoy J, Kennedy K, Sublett J, et al. Environmental assessment and exposure control: a practice parameter–furry animals. *Ann Allergy Asthma Immunol*. 2012;108:223.e1–223.e15.
- [16] Phipatanakul W, Matsui E, Portnoy J, et al. Environmental assessment and exposure reduction of rodents: a practice parameter. *Ann Allergy Asthma Immunol*. 2012;109:375–387.
- [17] Schneider L, Tilles S, Lio P, et al. Atopic dermatitis: a practice parameter update 2012. *J Allergy Clin Immunol*. 2013;131:295–299.