

Appraising the Evidence Supporting *Choosing Wisely*[®] Recommendations

Andrew J. Admon, MD, MPH^{1,2,*}, Ashwin Gupta, MD^{1,3}, Margaret Williams, MD⁴, Thomas S. Valley, MD, MSc^{1,2}, Michael W. Sjoding, MD, MSc^{1,2}, Renda Soylemez Wiener, MD, MPH^{5,6}, Colin R. Cooke, MD, MS, MSc^{1,2,7}

¹Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan; ²Division of Pulmonary and Critical Care Medicine, University of Michigan, Ann Arbor, Michigan; ³Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, Michigan; ⁴Department of Internal Medicine, Ohio State University Wexner Medical Center, Columbus, Ohio; ⁵Center for Healthcare Organization and Implementation Research, Edith Nourse Rogers Memorial Veterans Affairs Hospital, Bedford, Massachusetts; ⁶The Pulmonary Center, Boston University School of Medicine, Boston, Massachusetts; ⁷Center for Healthcare Outcomes and Policy, Institute of Healthcare Policy and Innovation, University of Michigan, Ann Arbor, Michigan.

Despite the growing enthusiasm surrounding the *Choosing Wisely*[®] campaign, little is known regarding the evidence underlying these recommendations. We extracted references for all 320 recommendations published through August, 2014, including the 10 adult and pediatric recommendations published by the Society for Hospital Medicine. We then categorized each item by evidence strength, and then assessed a sample of referenced clinical practice guidelines (CPGs) using the validated Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument. Among all recommendations, 70.3% cited CPGs, whereas 22.2% cited primary research as their highest level of evidence. Moreover,

7.8% cited case series, review articles, editorials, or lower quality data as their highest level of evidence. Hospital medicine recommendations were more likely to cite CPGs (90%) as their highest level of evidence. Among the sampled CPGs, the median overall score obtained using AGREE II was 54.2% (interquartile range [IQR] 33.3%-70.8%), whereas among hospital medicine-referenced CPGs, the median overall score was 58.3% (IQR 50.0%-83.3%). These findings suggest that *Choosing Wisely*[®] recommendations vary in terms of evidence strength. *Journal of Hospital Medicine* 2018;13:688-691. Published online first April 25, 2018. © 2018 Society of Hospital Medicine

As healthcare costs rise, physicians and other stakeholders are now seeking innovative and effective ways to reduce the provision of low-value services.^{1,2} The *Choosing Wisely*[®] campaign aims to further this goal by promoting lists of specific procedures, tests, and treatments that providers should avoid in selected clinical settings.³ On February 21, 2013, the Society of Hospital Medicine (SHM) released two *Choosing Wisely*[®] lists consisting of adult and pediatric services that are seen as costly to consumers and to the healthcare system, but which are often nonbeneficial or even harmful.^{4,5} A total of 80 physician and nurse specialty societies have joined in submitting additional lists.

Despite the growing enthusiasm for this effort, questions remain regarding the *Choosing Wisely*[®] campaign's ability to initiate the meaningful de-adoption of low-value services. Specifically, prior efforts to reduce the use of services deemed to be of questionable benefit have met several challenges.^{2,6} Early analyses of the *Choosing Wisely*[®] recommendations reveal similar roadblocks and variable uptakes of several recommendations.⁷⁻¹⁰ While the reasons for difficulties in achieving de-adop-

tion are broad, one important factor in whether clinicians are willing to follow guideline recommendations from such initiatives as *Choosing Wisely*[®] is the extent to which they believe in the underlying evidence.¹¹ The current work seeks to formally evaluate the evidence supporting the *Choosing Wisely*[®] recommendations, and to compare the quality of evidence supporting SHM lists to other published *Choosing Wisely*[®] lists.

METHODS

Data Sources

Using the online listing of published *Choosing Wisely*[®] recommendations, a dataset was generated incorporating all 320 recommendations comprising the 58 lists published through August, 2014; these include both the adult and pediatric hospital medicine lists released by the SHM.^{4,5,12} Although data collection ended at this point, this represents a majority of all 81 lists and 535 recommendations published through December, 2017. The reviewers (A.J.A., A.G., M.W., T.S.V., M.S., and C.R.C.) extracted information about the references cited for each recommendation.

Data Analysis

The reviewers obtained each reference cited by a *Choosing Wisely*[®] recommendation and categorized it by evidence strength along the following hierarchy: clinical practice guideline (CPG), primary research, review article, expert opinion, book, or others/unknown. CPGs were used as the highest level of evidence based on standard expectations for methodological rigor.¹³ Primary research was further rated as follows:

*Address for correspondence: Andrew J. Admon, North Campus Research Complex, Building 16 127W, 2800 Plymouth Road, Ann Arbor, MI, 48109-2800; E-mail: ajadmon@umich.edu

Additional Supporting Information may be found in the online version of this article.

Received: October 24, 2017; Revised: January 17, 2018;

Accepted: February 4, 2018

© 2018 Society of Hospital Medicine DOI 10.12788/jhm.2964

TABLE 1. Highest Quality of Evidence per Recommendation

Reference Category	All Recommendations (n, %)	Hospital Medicine Recommendations (n, %)
Clinical Practice Guideline	225 (70.3%)	9 (90%)
Primary Research Article	71 (22.2%)	1 (10%)
Systematic review and meta-analysis	29 (9.1%)	–
Randomized controlled trial	13 (4.1%)	–
Observational study	28 (8.8%)	1 (10%)
Case series	1 (0.3%)	–
Review Article	7 (2.2%)	–
Editorial/Opinion	7 (2.2%)	–
Others	10 (3.1%)	–

systematic reviews and meta-analyses, randomized controlled trials (RCTs), observational studies, and case series. Each recommendation was graded using only the strongest piece of evidence cited.

Guideline Appraisal

We further sought to evaluate the strength of referenced CPGs. To accomplish this, a 10% random sample of the *Choosing Wisely*[®] recommendations citing CPGs was selected, and the referenced CPGs were obtained. Separately, CPGs referenced by the SHM-published adult and pediatric lists were also obtained. For both groups, one CPG was randomly selected when a recommendation cited more than one CPG. These guidelines were assessed using the Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument, a widely used instrument designed to assess CPG quality.^{14,15} AGREE II consists of 25 questions categorized into six domains: scope and purpose, stakeholder involvement, rigor of development, clarity of presentation, applicability, and editorial independence. Guidelines are also assigned an overall score. Two trained reviewers (A.J.A. and A.G.) assessed each of the sampled CPGs using a standardized form. Scores were then standardized using the method recommended by the instrument and reported as a percentage of available points. Although a standard interpretation of scores is not provided by the instrument, prior applications deemed scores below 50% as deficient.^{16,17} When a recommendation item cited multiple CPGs, one was randomly selected. We also abstracted data on the year of publication, the evidence grade assigned to specific items recommended by *Choosing Wisely*[®], and whether the CPG addressed the referring recommendation. All data management and analysis were conducted using Stata (V14.2, StataCorp, College Station, Texas).

RESULTS

A total of 320 recommendations were considered in our analysis, including 10 published across the two hospital medicine lists. When limited to the highest quality citation for each of

the recommendations, 225 (70.3%) cited CPGs, whereas 71 (22.2%) cited primary research articles (Table 1). Specifically, 29 (9.1%) cited systematic reviews and meta-analyses, 28 (8.8%) cited observational studies, and 13 (4.1%) cited RCTs. One recommendation (0.3%) cited a case series as its highest level of evidence, seven (2.2%) cited review articles, seven (2.2%) cited editorials or opinion pieces, and 10 (3.1%) cited other types of documents, such as websites or books. Among hospital medicine recommendations, nine (90%) referenced CPGs and one (10%) cited an observational study.

For the AGREE II assessment, we included 23 CPGs from the 225 referenced across all recommendations, after which we separately selected six CPGs from the hospital medicine recommendations. There was no overlap. Notably, four hospital medicine recommendations referenced a common CPG. Among the random sample of referenced CPGs, the median overall score obtained by using AGREE II was 54.2% (interquartile range [IQR] 33.3%-70.8%, Table 2). This was similar to the median overall among hospital medicine guidelines (58.2%, IQR 50.0%-83.3%). Both hospital medicine and other sampled guidelines tended to score poorly in stakeholder involvement (48.6%, IQR 44.1%-61.1% and 47.2%, IQR 38.9%-61.1%, respectively). There were no significant differences between hospital medicine-referenced CPGs and the larger sample of CPGs in any AGREE II subdomains. The median age from the CPG publication to the list publication was seven years (IQR 4-7) for hospital medicine recommendations and three years (IQR 2-6) for the nonhospital medicine recommendations. Substantial agreement was found between raters on the overall guideline assessment (ICC 0.80, 95% CI 0.58-0.91; Supplementary Table 1).

In terms of recommendation strengths and evidence grades, several recommendations were backed by Grades II-III (on a scale of I-III) evidence and level C (on a scale of A-C) recommendations in the reviewed CPG (Society of Maternal-Fetal Medicine, Recommendation 4, and Heart Rhythm Society, Recommendation 1). In one other case, the cited CPG did not directly address the *Choosing Wisely*[®] item (Society of Vascular Medicine, Recommendation 2).

TABLE 2. AGREE II Assessment by Domain for Random Sample of Referenced Clinical Practice Guidelines (CPGs)

Domain	Sample of All CPGs (Median Scaled Score %, IQR) (n = 6)	Hospital Medicine CPGs (Median Scaled Score %, IQR) (n = 24)
1. Scope and Purpose	75.0 (72.2-83.3)	84.7 (75.0-91.7)
2. Stakeholder Involvement	47.2 (38.9-61.1)	48.6 (44.4-61.1)
3. Rigor of Development	54.2 (34.4-74.0)	66.1 (59.4-74.0)
4. Clarity of Presentation	80.6 (75.0-86.1)	80.6 (80.6-83.3)
5. Applicability	50.0 (37.5-64.6)	52.1 (47.9-62.5)
6. Editorial Independence	50.0 (37.5-64.6)	60.4 (25.0-70.8)
Overall Guideline Assessment	54.2 (33.3-70.8)	58.3 (50.0-83.3)

Abbreviations: CPGs, clinical practice guidelines; IQR, interquartile range.

DISCUSSION

Given the rising costs and the potential for iatrogenic harm, curbing ineffective practices has become an urgent concern. To achieve this, the *Choosing Wisely*® campaign has taken an important step by targeting certain low-value practices for de-adoption. However, the evidence supporting recommendations is variable. Specifically, 25 recommendations cited case series, review articles, or lower quality evidence as their highest level of support; moreover, among recommendations citing CPGs, quality, timeliness, and support for the recommendation item were variable. Although the hospital medicine lists tended to cite higher-quality evidence in the form of CPGs, these CPGs were often less recent than the guidelines referenced by other lists.

Our findings parallel those of other works that evaluate evidence among *Choosing Wisely*® recommendations and, more broadly, among CPGs.¹⁸⁻²¹ Lin and Yancey evaluated the quality of primary care-focused *Choosing Wisely*® recommendations using the Strength of Recommendation Taxonomy, a ranking system that evaluates evidence quality, consistency, and patient-centeredness.¹⁸ In their analysis, the authors found that many recommendations were based on lower quality evidence or relied on nonpatient-centered intermediate outcomes. Several groups, meanwhile, have evaluated the quality of evidence supporting CPG recommendations, finding them to be highly variable as well.¹⁹⁻²¹ These findings likely reflect inherent difficulties in the process, by which guideline development groups distill a broad evidence base into useful clinical recommendations, a reality that may have influenced the *Choosing Wisely*® list development groups seeking to make similar recommendations on low-value services.

These data should be taken in context due to several limitations. First, our sample of referenced CPGs includes only a small sample of all CPGs cited; thus, it may not be representative of all referenced guidelines. Second, the AGREE II assessment is inherently subjective, despite the availability of training materials. Third, data collection ended in April, 2014. Although this represents a majority of published lists to date, it is possible that more recent *Choosing Wisely*® lists include a stronger focus on

evidence quality. Finally, references cited by *Choosing Wisely*® may not be representative of the entirety of the dataset that was considered when formulating the recommendations.

Despite these limitations, our findings suggest that *Choosing Wisely*® recommendations vary in terms of evidence strength. Although our results reveal that the majority of recommendations cite guidelines or high-quality original research, evidence gaps remain, with a small number citing low-quality evidence or low-quality CPGs as their highest form of support. Given the barriers to the successful de-implementation of low-value services, such campaigns as *Choosing Wisely*® face an uphill battle in their attempt to prompt behavior changes among providers and consumers.⁶⁻⁹ As a result, it is incumbent on funding agencies and medical journals to promote studies evaluating the harms and overall value of the care we deliver.

CONCLUSIONS

Although a majority of *Choosing Wisely*® recommendations cite high-quality evidence, some reference low-quality evidence or low-quality CPGs as their highest form of support. To overcome clinical inertia and other barriers to the successful de-implementation of low-value services, a clear rationale for the impetus to eradicate entrenched practices is critical.²² *Choosing Wisely*® has provided visionary leadership and a powerful platform to question low-value care. To expand the campaign's efforts, the medical field must be able to generate the high-quality evidence necessary to support these efforts; further, list development groups must consider the availability of strong evidence when targeting services for de-implementation.

Acknowledgment

This work was supported, in part, by a grant from the Agency for Healthcare Research and Quality (No. K08HS020672, Dr. Cooke).

Disclosures: The authors have nothing to disclose.

References

1. Institute of Medicine Roundtable on Evidence-Based Medicine. The Healthcare Imperative: Lowering Costs and Improving Outcomes: Workshop Series

- Summary. Yong P, Saudners R, Olsen L, editors. Washington, D.C.: National Academies Press; 2010.
2. Weinberger SE. Providing high-value, cost-conscious care: a critical seventh general competency for physicians. *Ann Intern Med.* 2011;155(6):386-388.
 3. Cassel CK, Guest JA. Choosing wisely: Helping physicians and patients make smart decisions about their care. *JAMA.* 2012;307(17):1801-1802.
 4. Bulger J, Nickel W, Messler J, Goldstein J, O'Callaghan J, Auron M, et al. Choosing wisely in adult hospital medicine: Five opportunities for improved healthcare value. *J Hosp Med.* 2013;8(9):486-492.
 5. Quinonez RA, Garber MD, Schroeder AR, Alverson BK, Nickel W, Goldstein J, et al. Choosing wisely in pediatric hospital medicine: Five opportunities for improved healthcare value. *J Hosp Med.* 2013;8(9):479-485.
 6. Prasad V, Ioannidis JP. Evidence-based de-implementation for contradicted, unproven, and aspiring healthcare practices. *Implement Sci.* 2014;9:1.
 7. Rosenberg A, Agiro A, Gottlieb M, Barron J, Brady P, Liu Y, et al. Early trends among seven recommendations from the Choosing Wisely campaign. *JAMA Intern Med.* 2015;175(12):1913-1920.
 8. Zikmund-Fisher BJ, Kullgren JT, Fagerlin A, Klamerus ML, Bernstein SJ, Kerr EA. Perceived barriers to implementing individual Choosing Wisely® recommendations in two national surveys of primary care providers. *J Gen Intern Med.* 2017;32(2):210-217.
 9. Bishop TF, Cea M, Miranda Y, Kim R, Lash-Dardia M, Lee JI, et al. Academic physicians' views on low-value services and the choosing wisely campaign: A qualitative study. *Healthc (Amsterdam, Netherlands).* 2017;5(1-2):17-22.
 10. Prochaska MT, Hohmann SF, Modes M, Arora VM. Trends in Troponin-only testing for AMI in academic teaching hospitals and the impact of Choosing Wisely®. *J Hosp Med.* 2017;12(12):957-962.
 11. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA.* 1999;282(15):1458-1465.
 12. ABIM Foundation. *ChoosingWisely.org Search Recommendations.* 2014.
 13. Institute of Medicine (US) Committee on Standards for Developing Trust-worthy Clinical Practice Guidelines. *Clinical Practice Guidelines We Can Trust.* Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, editors. Washington, D.C.: National Academies Press; 2011.
 14. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: Advancing guideline development, reporting, and evaluation in health care. *Prev Med (Baltim).* 2010;51(5):421-424.
 15. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. Development of the AGREE II, part 2: Assessment of validity of items and tools to support application. *CMAJ.* 2010;182(10):E472-E478.
 16. He Z, Tian H, Song A, Jin L, Zhou X, Liu X, et al. Quality appraisal of clinical practice guidelines on pancreatic cancer. *Medicine (Baltimore).* 2015;94(12):e635.
 17. Isaac A, Saginur M, Hartling L, Robinson JL. Quality of reporting and evidence in American Academy of Pediatrics guidelines. *Pediatrics.* 2013;131(4):732-738.
 18. Lin KW, Yancey JR. Evaluating the Evidence for Choosing Wisely™ in Primary Care Using the Strength of Recommendation Taxonomy (SORT). *J Am Board Fam Med.* 2016;29(4):512-515.
 19. McAlister FA, van Diepen S, Padwal RS, Johnson JA, Majumdar SR. How evidence-based are the recommendations in evidence-based guidelines? *PLoS Med.* 2007;4(8):e250.
 20. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA.* 2009;301(8):831-841.
 21. Feuerstein JD, Gifford AE, Akbari M, Goldman J, Leffler DA, Sheth SG, et al. Systematic analysis underlying the quality of the scientific evidence and conflicts of interest in gastroenterology practice guidelines. *Am J Gastroenterol.* 2013;108(11):1686-1693.
 22. Robert G, Harlock J, Williams I. Disentangling rhetoric and reality: an international Delphi study of factors and processes that facilitate the successful implementation of decisions to decommission healthcare services. *Implement Sci.* 2014;9:123.