Evidentiary Rationales for the Choosing Wisely Top 5 Lists

Aiming to reduce wasteful medical care, the American Board of Internal Medicine (ABIM) Foundation’s Choosing Wisely initiative asks leading physician specialty societies to create a “Top 5” list of medical services that provide no overall benefit to patients in most situations.1,2 As of August 2013, 25 participating specialty societies had produced 1 or more Top 5 lists containing a total of 135 services.

The goal of this study was to evaluate the role that evidence on benefits, risks, and costs plays in selecting a service for the Top 5 lists. As Choosing Wisely continues to grow, clarity on the evidentiary justifications for the lists will be crucial for the overall credibility of the campaign.

Methods | We analyzed the evidentiary rationales provided by specialty societies using a framework of different types of marginal medicine (Table 1).1,3 These categories are based on the level of certainty in the evidence regarding risks and benefits, how the risks and benefits of the service compare with other alternatives, and the comparative cost or cost-effectiveness of the service.

Categorization was performed independently by both authors, and disagreements were resolved through discussion. We relied solely on the information given explicitly in the published rationales for each Top 5 list.

Table 1. Categories of Evidentiary Justifications for Top 5 List Services

<table>
<thead>
<tr>
<th>Category*</th>
<th>Description</th>
<th>Example From Choosing Wisely Top 5 Lists⁶</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insufficient evidence to evaluate comparative benefit for any indication</td>
<td>This category includes medical services that lack sufficient evidence demonstrating that their net benefits (over any risks that the services entail) are equal to or better than other options, for any indication. This is not the same thing as having solid evidence that a service provides no benefit or an inferior net benefit compared with other services; instead, it is meant to encompass experimental services, or anything that otherwise lacks an adequate evidence base to evaluate its benefits and risks.</td>
<td>American Academy of Hospice and Palliative Medicine recommendation: “Don’t use topical lorazepam (Ativan), diphenhydramine (Benadryl), or haloperidol (Haldol) (ABH) gel for nausea.” Rationale includes: “[W]hile topical gels are commonly prescribed in hospice practice, anti-nausea gels have not been proven effective in any large, well-designed or placebo-controlled trials.”</td>
</tr>
<tr>
<td>Insufficient evidence to evaluate comparative benefit for use beyond the boundaries of established indications, frequency, intensity, or dosage</td>
<td>This category includes services for which there is inadequate evidence supporting use outside established boundaries. The prototypical example is unsupported off-label drug use, although this category can encompass any treatment (even those with strong evidence of superior clinical benefits for certain indications) for which there is insufficient evidence to evaluate its use for other conditions or at different dosage or intensity levels.</td>
<td>American Urological Association recommendation: “Don’t prescribe testosterone to men with erectile dysfunction who have normal testosterone levels.” Rationale includes: “The information available in studies to date is insufficient to fully evaluate testosterone’s efficacy in the treatment of men with erectile dysfunction who have normal testosterone levels.”</td>
</tr>
<tr>
<td>Adequate evidence demonstrating equivalent benefit with higher risk, higher cost, or both</td>
<td>This category encompasses services whose benefits are equivalent but not superior to other options, and which are more expensive or involve increased risk to patients. Additional risks may be serious adverse events documented in the clinical literature, or potential risks, such as future cancers from radiation.</td>
<td>American Geriatrics Society recommendation: “Don’t use antimicrobials to treat bacteriuria in older adults unless specific urinary tract symptoms are present.” Rationale includes: “Antimicrobial treatment studies for asymptomatic bacteriuria in older adults demonstrate no benefits and show increased adverse antimicrobial effects.”</td>
</tr>
<tr>
<td>Adequate evidence demonstrating a small comparative benefit not large enough to justify higher risk to patients, higher cost, or both</td>
<td>This category includes services for which evidence demonstrates marginal benefits that do not warrant the magnitude of increased risks or higher costs. Common examples of this category include the first-line use of treatments with risks of serious adverse effects for minor dermatological conditions, and expensive drugs that provide short average gains in life expectancy for advanced cancer.</td>
<td>American Academy of Family Physicians recommendation: “Don’t use dual-energy x-ray absorptiometry (DEXA) screening for osteoporosis in women younger than 65 or men younger than 70 with no risk factors.” Rationale includes: “DEXA is not cost effective in younger, low-risk patients, but is cost effective in older patients.”</td>
</tr>
<tr>
<td>Adequate evidence demonstrating improved comparative benefit, lower risk, lower cost, or both when using the intervention⁷</td>
<td>This category is meant to represent services that when used tend to reduce waste. Waste is therefore failure to use that service. Examples include more precise diagnostic tests or first-line treatments that are more effective, less costly, or both, than other options.</td>
<td>American Academy of Allergy, Asthma, and Immunology recommendation: “Don’t diagnose or manage asthma without spirometry.” Rationale includes: “[R]ecent guidelines highlight spirometry’s value in stratifying disease severity and monitoring control. History and physical exam alone may over- or underestimate asthma control. Beyond the increased costs of care, repercussions of misdiagnosing asthma include delaying a correct diagnosis and treatment.”</td>
</tr>
</tbody>
</table>

* Categories adapted from those presented by Hoffman and Pearson.³
⁶ All examples from Choosing Wisely.¹
⁷ In this category, waste is framed as a failure to use a service. This category was necessitated by the rationales introduced by several specialty societies.
Table 2. Distribution of Top 5 List Services in Different Evidentiary Categories

<table>
<thead>
<tr>
<th>Category</th>
<th>No. (%) of Services (n = 135)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insufficient evidence to evaluate comparative benefit for any indication</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Insufficient evidence to evaluate comparative benefit for use beyond the boundaries of established indications, frequency, intensity, or dosage</td>
<td>18 (13)</td>
</tr>
<tr>
<td>Adequate evidence demonstrating equivalent benefit with higher risk, higher cost, or both</td>
<td>102 (76)</td>
</tr>
<tr>
<td>Adequate evidence demonstrating a small comparative benefit not large enough to justify higher risk to patients, higher cost, or both</td>
<td>8 (6)</td>
</tr>
<tr>
<td>Adequate evidence demonstrating improved comparative benefit, lower risk, lower cost, or both, when using the intervention*</td>
<td>5 (4)</td>
</tr>
</tbody>
</table>

* In this category, waste is framed as a failure to use a service. This category was necessitated by the rationales introduced by several specialty societies.

Results | Of the 135 services, 49 (36%) were for patient diagnosis, prognosis, or monitoring; 46 (34%) for patient treatment; and 40 (30%) for population screening. Initial evidence categorization was identical for 128 (95%) of 135 services.

Inclusion of 102 services (76%) was justified by claims that adequate evidence demonstrated no additional benefit with higher risk, higher cost, or both, compared with other options (Table 2). The second most common rationale, given for 18 services (13%), was that there was insufficient evidence to evaluate comparative benefit for use of the service beyond the evidentiary boundaries of established indications, frequency, intensity, or dosage. Other evidentiary rationales shown in Table 2 were used infrequently.

Overall, 66 (49%) of all 135 rationales mentioned greater risks to patients as a consideration in selecting the service, 33 (24%) mentioned higher costs, 21 (16%) mentioned both greater risk and higher cost, and 57 (42%) mentioned neither. Of the 25 specialty societies, 15 (60%) had at least 1 service whose inclusion was justified in part by higher costs.

Discussion | Most services were included in the Top 5 lists on the basis of evidence considered to be adequate to demonstrate equivalent but not superior benefit with higher risk or higher costs, or both, compared with other options. The specialty societies did not emphasize emerging or experimental interventions.

Our data show that the issue of cost was almost always raised in the context of a service being judged as good as other options but more expensive. We believe that specialty societies should seek greater opportunities to include within their Top 5 lists services that offer only small incremental benefits at much higher prices.

Less than half (49%) of the rationales mentioned specific patient risks. However, we only noted explicit mentions of risk, so it is likely that some risks (such as those from radiation exposure) were an implicit concern.

Our analysis has additional limitations. We did not evaluate the procedures through which each specialty society created its Top 5 list or analyze the evidence cited in their rationales. However, we believe the rationales given by the specialty societies need to stand on their own in justifying the selection of services. Specialty societies can enhance trust in the Choosing Wisely campaign by defining more clearly the types of potentially wasteful medical care they seek to eliminate, and by providing a clear evidentiary justification for the selection of each service.

Catherine Gliwa, BA
Steven D. Pearson, MD, MSc

Author Affiliations: Department of Bioethics, National Institutes of Health, Bethesda, Maryland.

Corresponding Author: Steven D. Pearson, MD, MSc, National Institutes of Health, 10 Center Dr, Bldg 10, Room 1C118, Bethesda, MD 20892 (pearsonsd@cc.nih.gov).

Author Contributions: Ms Gliwa and Dr Pearson had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Both authors.

Acquisition, analysis, or interpretation of data: Both authors.

Drafting of the manuscript: Both authors.

Critical revision of the manuscript for important intellectual content: Both authors.

Administrative, technical, or material support: Gliwa.

Study supervision: Pearson.

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding/Support: This research was supported by the Intramural Research Program of the National Institutes of Health.

Role of the Sponsor: The Intramural Research Program of the National Institutes of Health had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The opinions expressed are those of the authors and do not reflect the position or policy of the National Institutes of Health, the Public Health Service, or the Department of Health and Human Services.

Additional Contributions: We thank our colleagues in the Department of Bioethics at the National Institutes of Health for their thoughtful comments and advice during the early stages of this project.


COMMENT & RESPONSE

Gastric Acid–Inhibiting Medications and Vitamin B12 Deficiency

To the Editor Mr Lam and colleagues1 reported that previous and current gastric acid–inhibiting medication use for 2 or more years (proton pump inhibitors [PPIs] and histamine 2 receptor antagonists [H2RAs]) was associated with a greater risk of vitamin B12 deficiency; this association became stronger with increasing daily consumption of PPIs or H2RAs. The study is, to date, the largest study suggesting a relationship between PPIs or H2RAs and vitamin B12 deficiency. Lam et al1 confirmed and