

Original Scholarship

The Impact of Choosing Wisely Interventions on Low-Value Medical Services: A Systematic Review

BETSY Q. CLIFF,* ANTON L.V. AVANCEÑA,†
RICHARD A. HIRTH,†
and SHOOU-YIH DANIEL LEE‡

**School of Public Health, University of Illinois Chicago; †School of Public Health, University of Michigan; ‡College of Health Professions, Virginia Commonwealth University*

Policy Points:

- Dissemination of Choosing Wisely guidelines alone is unlikely to reduce the use of low-value health services.
- Interventions by health systems to implement Choosing Wisely guidelines can reduce the use of low-value services.
- Multicomponent interventions targeting clinicians are currently the most effective types of interventions.

Context: Choosing Wisely aims to reduce the use of unnecessary, low-value medical services through development of recommendations related to service utilization. Despite the creation and dissemination of these recommendations, evidence shows low-value services are still prevalent. This paper synthesizes literature on interventions designed to reduce medical care identified as low value by Choosing Wisely and evaluates which intervention characteristics are most effective.

Methods: We searched peer-reviewed and gray literature from the inception of Choosing Wisely in 2012 through June 2019 to identify interventions in the United States motivated by or using Choosing Wisely recommendations. We also included studies measuring the impact of Choosing Wisely on its own, without interventions. We developed a coding guide and established coding agreement. We coded all included articles for types of services targeted, components of each intervention, results of the intervention, study type, and, where

applicable, study quality. We measured the success rate of interventions, using chi-squared tests or Wald tests to compare across interventions.

Findings: We reviewed 131 articles. Eighty-eight percent of interventions focused on clinicians only; 48% included multiple components. Compared with dissemination of Choosing Wisely recommendations only, active interventions were more likely to generate intended results (65% vs 13%, $p < 0.001$) and, among those, interventions with multiple components were more successful than those with one component (77% vs 47%, $p = 0.002$). The type of services targeted did not matter for success. Clinician-based interventions were more effective than consumer-based, though there is a dearth of studies on consumer-based interventions. Only 17% of studies included a control arm.

Conclusions: Interventions built on the Choosing Wisely recommendations can be effective at changing practice patterns to reduce the use of low-value care. Interventions are more effective when targeting clinicians and using more than one component. There is a need for high-quality studies that include active controls.

Keywords: delivery of health care, health services/standards, health services/economics, Choosing Wisely guidelines, health care quality.

REDUCING THE USE OF MEDICAL SERVICES THAT DO NOT improve patients' health is crucial for both the efficiency and quality of the health care system. One of the largest efforts to do so is the Choosing Wisely campaign, which launched in 2012 to reduce utilization of unnecessary tests and procedures. In the past eight years, it has attracted historic levels of engagement from medical societies, health care delivery systems, employers, and patient groups. To date, more than 80 medical specialist organizations have participated in the campaign and generated lists of often unnecessary services pertaining to their specialty. These efforts have produced more than 600 recommendations for ways to reduce overused services and align medical care with clinical value.

Since its inception, the campaign has garnered widespread publicity both within the medical community and beyond. It has been the subject of numerous research papers, journal commentaries, and policy papers.¹⁻⁴ Additionally, it has gained attention through national news coverage, academic and clinical conferences, and dissemination through specialty societies. Despite the attention, a key question remains: has Choosing Wisely reduced use of low-value care?

Recent work suggests low-value services remain prevalent in the US health care system, leading to as much as \$101 billion in additional spending.⁵⁻⁹ Impediments to aligning service use with the goals of the Choosing Wisely campaign include identifying recommendations that are both motivated by evidence and have the potential to significantly impact patient care, physician awareness of the campaign, and measurement of low-value services.^{8,10,11} One key challenge is that while dissemination of the recommendations has been widespread, actual interventions to implement the recommendations have been piecemeal. Efforts to measure and reduce low-value care have primarily fallen to individual health systems, hospitals, or divisions within facilities, with little coordination among them. The literature includes numerous reports of efforts to reduce low-value services identified by the Choosing Wisely campaign; yet, to our knowledge, no review has set out to synthesize the results of these individual efforts. In particular, it is not known what types of strategies are most commonly employed by health care providers and payers to implement the Choosing Wisely recommendations and the effectiveness of these strategies. Although one notable review by Colla et al. looked at low-value services, it encompassed services both within and beyond Choosing Wisely.¹² It contained articles from before the Choosing Wisely campaign and up to early 2015; many interventions tied to the campaign were published after this date.

This literature review provides a focused update to previous work by analyzing existing evidence on interventions that sought to reduce the use of low-value services targeted or motivated by the Choosing Wisely campaign. We describe interventions used by health systems, payers, hospitals, and clinics, and the interventions' impacts on specific low-value medical services from the inception of Choosing Wisely in 2012 through the middle of 2019. We also assess the quality of extant literature. Our aim is to inform policymakers, health system leaders, payers, and clinicians about the components needed for successful implementation of Choosing Wisely recommendations and, as such, generate more widespread reductions in low-value care.

Methods

Search Strategy

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines¹³

and was registered in the PROSPERO registry (study no. CRD 42019140501).¹⁴ With the help of a reference librarian, we developed a search algorithm for both full-text article searches and title/abstract title searches. We aimed for sensitivity over specificity in our initial search, thus using a full range of terms related to low-value or unnecessary care. We searched Medline (via OVID), Scopus, Web of Science, and DimensionsPlus for English-language peer-reviewed and gray literature published from 2012 to June 2019. The full list of keywords and search terms is available in Appendix S1 in the Supplement.

Study Screening and Selection

After removing duplicates from the results, one author screened articles based on titles and abstracts. After the initial screening, we retrieved full-text articles that were assessed for eligibility by two authors and one research assistant. Articles were included if they (1) included US-based patients or health care providers; (2) measured the rate of use over time of at least one low-value service identified or explicitly motivated by the Choosing Wisely campaign; (3) mentioned the Choosing Wisely campaign or the ABIM Foundation in one of the article's fields (e.g., title, abstract, full text, funding source); and (4) included at least one intervention, such as the use of the Choosing Wisely list or a clinician- or consumer-focused strategy to reduce low-value service utilization. Articles were excluded if they (1) only measured opinions about or knowledge of Choosing Wisely recommendations; (2) were descriptions of or opinions about the Choosing Wisely campaign; or (3) did not include an intervention as previously defined. For uncertainties or disagreements on the inclusion of specific articles, the authors discussed until a consensus was reached.

We also hand-searched reference lists of all full-text articles for studies that may have been missed in the original search. Similarly, experts in the field, including those at the ABIM Foundation, were consulted for relevant articles. Unique titles identified from this step were assessed for eligibility based on the aforementioned inclusion and exclusion criteria.

Data Collection and Synthesis

We extracted the following data from the included articles: study type, setting, medical field or specialty, target population, components of the

intervention, length of intervention and follow-up period, other outcome measures (e.g., satisfaction, health outcomes, costs), and up to 10 primary outcomes related to low-value service utilization.

We followed and expanded on the framework of Colla et al. to develop a list of possible clinician- and consumer-focused intervention components.¹² (See Table S1 in the Supplement for a list of intervention components and definitions.) Components included clinician-focused tactics such as point-of-care alerts in the electronic health record (EHR), feedback about service utilization, and clinician education or academic detailing. Consumer-focused tactics included patient cost sharing or patient education materials. We also measured outcomes from the development and dissemination of the Choosing Wisely recommendation lists without any intervention.

Our central interest is the effects of interventions on low-value medical care identified or informed by the Choosing Wisely initiative. Our unit of analysis is a single study; no study described more than one intervention, though many interventions had multiple components. Because many interventions also measured effects on more than one low-value service, we determined whether the intervention produced statistically significant and intended results across all primary outcomes (e.g., reductions in low-value services), no statistically significant results for any primary outcomes, unintended results across all primary outcomes (e.g., an increase in low-value services) or, finally, mixed results, which included any combination of intended, no, and unintended results. For studies in which authors split primary outcomes (such as an aggregate measure of low-value service use in addition to individual service measures), we examined studies on a case-by-case basis, with the default being to split services by types of service. When studies included services that were not on a Choosing Wisely list, we examined only services that were part of the list or, in cases where the intervention was explicitly part of a Choosing Wisely–motivated initiative, all primary outcomes. To compare outcomes across types of interventions, we used chi-squared tests of independence or, to cluster standard errors in service-level comparisons, Wald tests. We assessed all studies for which outcomes and statistical significance could be measured. Then, we separately assessed the subset of articles that included control groups to measure whether studies with more rigor had different rates of success than those with less rigorous designs.

Risk of Bias Within Studies

We classified studies based on the taxonomy developed by the Oxford Centre for Evidence-based Medicine, which includes, in order of decreasing evidence strength, (1) properly powered and conducted randomized controlled trial (RCT); (2) well-designed controlled trial without randomization or prospective cohort trials; (3) case-control study or retrospective cohort study; and (4) case series or cross-sectional study.¹⁵

To assess study quality and risk of bias in controlled studies, we used an adapted version of the ROBINS-I (Risk of Bias in Non-randomized Studies—of Interventions) tool published by the US Agency for Healthcare Research and Quality.¹⁶ We appraised only RCTs, nonrandomized controlled trials, prospective comparative cohort trials, and case-control studies. The quality of all other study types without controls (e.g., cross-sectional studies, case series) was not assessed, though these studies may be prone to confounding by underlying trends and factors that affect outcomes and are not related to the intervention.¹⁷ Two authors conducted the quality assessment; when disagreements arose, discussions were held until a consensus was reached.

Risk of Bias Across Studies

We qualitatively assessed the risk of two types of dissemination bias in this review: publication bias and reporting bias.¹⁸ Publication bias is the selective publication of large studies and/or studies with positive results, which leads to the underreporting of null results; reporting bias is the selective release of statistically significant outcomes by researchers.¹⁹

We addressed publication bias, in part, through the use of multiple research databases and the inclusion of gray literature in our search. We also searched ClinicalTrials.gov, the largest trial registration database maintained by the National Library of Medicine of the National Institutes of Health, for unpublished trials that could meet our inclusion criteria. We could not use commonly used graphical and statistical approaches to test publication bias in meta-analyses because outcome measures between studies in our final sample were not uniform and did not lend themselves to such methods.^{18,20,21} To assess reporting bias, we followed the PRISMA recommendation and cross-checked RCT results against their published protocols, including those in

ClinicalTrials.gov, to determine deviations in the planned and reported outcomes.²²

Results

Our initial search resulted in 13,313 nonduplicated articles. After an initial screen that excluded articles not relevant to the topic, not in English, or without either a full text or abstract, we were left with 1,095 articles that received full-text review. Most of the articles excluded at the full-text review stage either were commentary about the Choosing Wisely campaign or were not explicitly motivated by the Choosing Wisely campaign or one of its lists. Our final sample included 131 articles. (A PRISMA diagram is found in Figure S1 in the Supplement.)

The majority of our sample included articles published since 2017 (Table 1), though some articles were published as early as 2014. More than 80 percent of included research was designed as cross-sectional or case series studies (e.g., pre-post design without a control group). Many of these studies were health care system quality improvement projects. Most studies were conducted in inpatient settings or in more than one setting or medical specialty. Several studies included more than one hospital system, particularly those that examined consumer-focused interventions.^{23–29} About two-thirds of studies were done in academic medical centers, and 88% targeted just clinicians. Of studies reviewed with an intervention, 40% had one primary outcome ($n = 53$ studies); 16% had two outcomes ($n = 22$ studies), 13% had three outcomes ($n = 17$ studies), and 12% had between four and nine outcomes ($n = 16$). Imaging scans and lab tests were the two most common types of service outcomes examined, followed by procedures. Prescriptions and medical product use (including blood products) were also examined; clinician visits were rarely studied (Table 1). This prevalence of measured service types differs slightly from the Choosing Wisely lists. Of the 667 recommendations included in Choosing Wisely in 2019, procedures make up the highest percentage (26%), followed by imaging (22%), medications (19%), lab tests (17%), blood products (3%), and exams (<1%) (Kelly Rand, ABIM Foundation, written communication, January 2021). Among the studies that reported length of intervention, the range was between two weeks³⁰ and 60 months.³¹ (For full details of each included study by intervention type, see Tables S3-S17 in the Supplement.)

Table 1. Summary of Selected Characteristics of Included Articles		
Study Characteristic	Number	Percent
Type of study		
Case series; cross-sectional study	109	83
Case-control studies; retrospective cohort study	4	3
Randomized controlled trial	8	6
Well-designed controlled trial without randomization; prospective comparative cohort trial	10	8
Publication year		
2014	10	8
2015	20	15
2016	21	16
2017	32	24
2018	31	24
2019	17	13
Type of setting		
Emergency department	15	11
Hospital inpatient	33	25
Hospital outpatient	13	10
Hospital-affiliated lab	1	1
Laboratory (not within a hospital)	1	1
Outpatient medical clinic (either single or multispecialty)	13	10
System-wide or multiple settings	75	57
Veterans Administration system	1	1
Academic medical center		
Yes	81	62
No	33	26
Not stated	16	12

Continued

Table 1. (Continued)		
Study Characteristic	Number	Percent
Medical specialty		
Hospitalist or general inpatient medicine	39	30
Primary care ^b	32	24
Oncology ^c	20	15
Emergency medicine	13	10
Critical care	10	8
Surgery	10	8
Cardiology	9	7
Radiology	9	7
Hematology	7	5
Anesthesiology	4	3
Other ^d	27	20
Social or health system costs considered		
Yes	44	34
Types of low-value services		
Imaging/scanning tests	87	29
Lab test	83	28
Medical product (e.g., blood products)	29	10
Prescription	31	10
Procedure (e.g., urinary catheter, telemetry monitoring)	61	20
Clinician visit	7	2
Number of intervention components		
Choosing Wisely guideline development only ^e	23	18
1 ^f	45	34
2	32	24
3	20	15
4 or more	11	8
Intervention targets		
Clinician-focused	115	88
Consumer-focused	4	3
Both clinicians and consumers	12	9

Continued

Table 1. (Continued)		
Study Characteristic	Number	Percent
Types of intervention components		
<i>Clinician-focused</i>		
Recommendation guideline dissemination only	23	18
Behavioral nudges	6	5
Change to order set or clinical documentation	32	24
Clinical decision support: mandatory or optional utilization review	6	5
Clinical decision support: point-of-care information or alert	36	27
Increasing access or use of health information exchange	2	2
Clinician champions	9	7
Clinician education or academic detailing	58	44
Creation of new clinical pathways or discontinuation criteria	11	8
Clinician feedback or report cards to clinicians	34	26
Creation of organizational change frameworks	9	7
Risk-sharing or alternative payment methods	4	3
<i>Consumer-focused</i>		
Patient cost sharing	1	1
Patient education materials or shared decision making	8	6
Clinician report cards to patients	1	1
^a May not add up to 100% due to rounding or because studies report multiple categories. ^b Includes family or community medicine, internal medicine, pediatrics, and obstetrics and gynecology. ^c Includes medical, surgical, and radiation oncology. ^d Includes specialties mentioned three times or less. ^e Refers to Choosing Wisely recommendation dissemination only; no other interventions are included. ^f Refers to studies with a single intervention that is not Choosing Wisely recommendation dissemination.		

We categorized the studies as no intervention (i.e., development and dissemination of the Choosing Wisely recommendations only; $n = 23$ studies; Table S3), single-component intervention ($n = 45$ studies), or multicomponent intervention ($n = 63$ studies; Table 1). Of those studies with an intervention that included only a single component, a point-of-care information or alert ($n = 14$ studies) was the most common, followed by clinician education or academic detailing ($n = 9$ studies) and change to order sets or documentation requirements ($n = 8$ studies; Table S2). Among studies in which interventions included multiple components, clinician education or academic detailing was the most common component ($n = 49$ studies), followed by clinician feedback, including report cards to clinicians ($n = 30$ studies), change to order sets or clinical documentation required for ordering low-value services ($n = 24$ studies), and point-of-care information or alerts ($n = 22$ studies; Table S2).

Comparison of Intervention Characteristics

When assessing the effects of interventions, we removed 15 studies for which the statistical significance of the result was unable to be discerned. In remaining studies, those that measured a single-component or multicomponent intervention produced intended results across all outcomes 64.0% of the time. By contrast, in studies that measured the effect of no intervention (i.e., when the study focused solely on whether development of the Choosing Wisely guidelines themselves had an impact), intended outcomes were generated 12.5% of the time (chi-squared test of independence, $p < 0.001$; Table 2). Studies that included a single component were less likely to return intended results than those with multiple components (46.5% vs 77.2%; $p = 0.002$; Table 2). However, once multicomponent studies were separated into the exact number of components, there was no relationship between the number of components and likelihood of success ($p = 0.4$; Table 2). We also did not see any statistically significant differences among the types of low-value services targeted as outcomes ($p = 0.3$) or among interventions that included systems-based changes (changes in clinical pathways, changes in order sets, or clinical alerts) versus those in which clinicians were encouraged to make affirmative behavior changes ($p = 0.5$).

Of the relatively small number of studies with interventions targeting consumers, 12 targeted both consumers/families and clinicians, while

Table 2. Intended Results by Intervention Characteristics

Intervention Type	Number of Studies/Services (Percent) With Intended Results	Number of Studies/Services (Percent) With Null, Unintended, or Mixed Results	P-Value for Chi-Squared Test of Independence of Results
Overall	66 (56.9)	50 (43.1)	n/a
Recommendation vs Intervention (n = 116)			
Recommendation only	2 (12.5)	14 (87.5)	<0.001
Intervention	64 (64.0)	36 (36.0)	
Single- vs multiple-component intervention (n = 100)			
Single-component intervention	20 (46.5)	23 (53.5)	0.002
Multiple-component intervention	44 (77.2)	13 (22.8)	
Number of components in multiple-component studies (n = 57)			
Intervention with 2 components	23 (79.3)	6 (20.7)	0.4
Intervention with 3 components	13 (68.4)	6 (31.6)	
Intervention with 4-6 components	8 (88.9)	1 (11.1)	

Continued

Table 2. (Continued)

Intervention Type	Number of Studies/Services (Percent) With Intended Results	Number of Studies/Services (Percent) With Null, Unintended, or Mixed Results	P-Value for Chi-Squared Test of Independence of Results
Clinician- vs patient-focused interventions (n = 100)			
Clinician-focused intervention	60 (65.2)	32 (34.8)	0.05
Patient/family-focused intervention	0 (0.00)	3 (100.0)	
Intervention focused on both clinicians and patients/families	4 (80.0)	1 (20.0)	
Type of Service Studied (n = 273)			
Imaging scans ^a	49 (63.6)	28 (36.4)	0.3
Lab tests	58 (74.4)	20 (25.6)	
Medical products, including blood products	24 (82.7)	5 (17.2)	
Prescriptions	19 (65.5)	10 (34.5)	
Procedure	37 (69.8)	16 (30.2)	
Clinician visit	6 (85.7)	1 (14.3)	
System-based vs Active Components (n = 100)			
Systems-based changes implemented in EHR	41 (67.2)	20 (32.8)	0.5
Active behavioral change required	23 (59.0)	16 (41.0)	
Controlled vs Uncontrolled Study (n = 116)			
Includes control arm	11 (50.0)	11 (50.0)	0.5
Does not include control arm	55 (58.5)	39 (41.5)	

Abbreviation: EHR, electronic health record.

Of the 131 articles included in the review, 15 are excluded from this table because their design did not permit statistical inference of results. The total number of studies for each statistical test is noted in the table.

^aEvaluated at the individual service level (n = 273) and statistical significance determined by a Wald test to allow for clustering at the study level.

only four targeted consumers or their families only (Table 1). Among studies for which statistical significance could be assessed, the percentage of interventions generating intended results was similar for interventions that targeted clinicians only (65.2%) and those interventions that targeted both consumers and clinicians (80.0%). Notably, among the few studies targeting consumers only, none generated intended results (Table 2).

There was considerable heterogeneity around which combinations of intervention components were used and how they were implemented, making it difficult to draw firm conclusions about which specific components are more likely to be successful. That said, service utilization review by health professionals generated intended results in both single-component interventions in which it was used (Table 3). Among multicomponent interventions, those that included behavioral nudges, utilization review, clinical education or academic detailing, or the creation of new clinical pathways were among the most likely to generate intended outcomes (Table 3).

Clinician-Focused Intervention Components

Behavioral Nudges. Behavioral nudges refer to the use of behavioral economic principles to steer clinicians or patients toward reductions in low-value care. (Definitions for all interventions are found in Table S1.) Behavioral nudges were used in six total interventions, including as a single component and in conjunction with other intervention components (Table S4). In the one study in which behavioral interventions were the only component of the intervention, the results were mixed.³² This study—an RCT among 45 primary care clinicians in six adult primary care clinics—found that a point-of-care precommitment device was associated with a statistically significant but small decrease in one of three targeted low-value services. Additionally, this study found that alternate orders, a secondary outcome, increased during the study period, suggesting that interventions targeting low-value services could have unintended effects.³² Incorporating behavioral nudges with other interventions, however, was much more successful; all five studies that included behavioral nudges as part of a multicomponent intervention reported intended results across all primary outcomes (Table 3).

Table 3. Single- and Multicomponent Interventions With Statistically Significant Results in the Intended Direction

Intervention	Number of Components in Intervention	Number of Studies Where All Primary Outcomes Change in Intended Direction (%)
Clinician-focused		
Recommendation guideline dissemination only	Single (<i>n</i> = 16)	1 (6)
Behavioral nudges	Single (<i>n</i> = 1) Multi (<i>n</i> = 5)	0 (0) 5 (100)
Change to order set or clinical documentation	Single (<i>n</i> = 8) Multi (<i>n</i> = 19)	4 (50) 15 (79)
Clinical decision support: mandatory or optional utilization review	Single (<i>n</i> = 2) Multi (<i>n</i> = 3)	2 (100) 3 (100)
Clinical decision support: point-of-care information or alert	Single (<i>n</i> = 14) Multi (<i>n</i> = 20)	7 (50) 13 (65)
Increasing access or use of health information exchange	Multi (<i>n</i> = 2)	1 (50)
Clinician champions	Multi (<i>n</i> = 7)	5 (71)
Clinician education or academic detailing	Single (<i>n</i> = 7) Multi (<i>n</i> = 45)	4 (57) 35 (78)
Creation of new clinical pathways or discontinuation criteria	Single (<i>n</i> = 2) Multi (<i>n</i> = 7)	1 (50) 6 (86)
Clinician feedback or report cards to clinicians	Single (<i>n</i> = 4) Multi (<i>n</i> = 30)	1 (25) 24 (80)
Creation of organizational change frameworks	Multi (<i>n</i> = 7)	7 (100)
Risk-sharing or alternative payment methods	Single (<i>n</i> = 1) Multi (<i>n</i> = 3)	0 (0) 2 (67)
Patient-focused		
Patient cost sharing	Single (<i>n</i> = 1)	0 (0)
Patient education materials or shared decision making	Single (<i>n</i> = 2) Multi (<i>n</i> = 5)	1 (50) 3 (60)
Clinician report cards to patients	Single (<i>n</i> = 1)	0 (0)

This table shows the proportion of studies that reported statistically significant changes in the intended direction across all primary outcome measures. Single-component interventions are separated from multicomponent interventions, which combine the stated component in the first column with other components. Categories that are missing from this graph have no studies that meet those criteria. Studies with results that were not able to be statistically evaluated are not included in this table.

Change to Order Set or Clinical Documentation. Changes to order sets or to clinical documentation required for orders were a common component of interventions and often used for lab tests. In total, 32 interventions used this component either alone or in combination with other components (Table S5). These interventions often involved changing aspects of the EHR so that low-value services no longer appeared as choices in orders or so that clinicians had to complete additional documentation to justify the order of a low-value service. Often, these interventions were combined with clinician education to inform clinicians why the order set was changed ($n = 16$; Table S5). Used alone, changes to the order set or clinical documentation generated the intended effect across all outcomes 50% of the time; combined with other components it generated the intended effects 79% of the time (Table 3). As an example, a multicomponent intervention in a pediatric hospital included a new order directing nurses to discontinue continuous pulse oximetry and initiate intermittent pulse oximetry at a specified time, with a direction to call a physician if concerns were present.³³ This study found that patient time on continuous pulse oximetry decreased without affecting negative patient outcomes, including discharge time and proportion of patients needing transfers, revisits, or medical emergency teams.

Clinical Decision Support: Mandatory or Optional Utilization Review. Clinical decision support constrains low-value services by promoting compliance and adherence to treatment guidelines and protocols; this can be achieved through utilization review or through point-of-care information or alerts (see the next section). Immediate utilization review of potentially unnecessary service orders is not as well studied as other intervention components ($n = 6$ studies; Table S6). Often, these interventions were done as health system quality improvement projects and the component that included utilization review took the form of certain orders triggering review by another professional typically, though not always, at the time of order. For example, an intervention that required colorectal surgery consult for all patients presenting to the emergency department with peri-anal abscess prior to obtaining CT scans resulted in a reduction in unnecessary scans.³⁴ In our review, this component generated statistically significant results in the intended direction in all studies where statistical inference was possible (Table 3).

Clinical Decision Support: Point-of-Care Information or Alert. Using an alert in the EHR to flag potentially inappropriate care was one of the most common intervention components ($n = 36$ studies; Table S7),

likely because it can be relatively straightforward to implement. The alerts often did not require additional action from clinicians or others and did not aim to censure clinicians in any way; they were primarily used to give clinicians information about service value. When used on their own, these alerts generated intended effects 50% of the time; when used with other intervention components, they generated intended results 65% of the time (Table 3).

As an example, Felcher and coauthors described the implementation of an alert in the Kaiser Permanente Northwest system to reduce unnecessary vitamin D testing. The health system disseminated a new guideline with regards to testing and posted it on the organization's website. Orders for vitamin D tests were removed from laboratory preference lists for all clinicians except endocrinologists, nephrologists, and orthopedists. For all clinicians, any order triggered an alert that included bullet points from the new guidelines and required the clinician to click again to order the test. Vitamin D screening rates decreased overall, including reductions in inappropriate screens.³⁵

In another example, Chien and colleagues describe a randomized controlled trial in a system of outpatient clinics that used a price alert to inform clinicians about the total cost of medical tests, including inappropriate medical tests.³⁶ Unlike other alerts we reviewed, this alert did not give clinicians information about the clinical value of services. This intervention had no effect on number of inappropriate tests ordered.

Increasing Access to the Electronic Health Record. Improving access to EHRs, by increasing informational capacity in EHRs, interoperability between health systems, or moving paper records to electronic, allows clinical and administrative data to be shared throughout a health care setting or system. Only two studies aimed to increase access to the EHR as a way to align service use with Choosing Wisely (Table S8). The low number of studies that used this component is likely related to the widespread use of EHRs within health systems, as well as the paucity of work in low-value care reduction that goes beyond one health system. One case-control study in a rural academic medical center describes a multicomponent intervention that put an enhanced clinician template into an existing EHR system and was successful in reducing unnecessary preoperative testing rates, though the effect was not statistically significant throughout the entire study period.³⁷ In another intervention, clinical notes were automatically transferred into the EHR as part of a multicomponent intervention in a pediatric emergency department,

which reduced inappropriate computed tomography (CT) scans for mild head injury.³⁸

Clinician Champions. Clinician champions, recruiting clinicians to advocate for Choosing Wisely interventions, was used with other components of interventions but never, in our review, on its own (Table S9). Used with other components, it was successful in generating intended results 71% of the time (Table 3). For example, Coronel and coauthors reported on a fellow- and resident-led intervention that aimed to reduce the use of continuous infusion of proton pump inhibitors in patients with upper gastrointestinal bleeding. Trainees targeted change both in hospital systems through clinical decision support and, for some departments, by recruiting faculty leaders to champion the initiative. The group that had clinician champions along with the decision support change—but not the group with decision support change only—saw statistically significant declines in inappropriate use of proton pump inhibitor infusion.³⁹ As another example, a two-week intervention involving nurse practitioner champions was unsuccessful in reducing the number of laboratory tests ordered for intensive care unit patients, though panels of tests did decrease, as did patient costs associated with testing.³⁰

Clinician Education or Academic Detailing. Closely related to clinician champions, clinician education was one of the most common intervention components used in conjunction with others ($n = 49$ studies; Table S10). Indeed, informing and explaining an intervention was often considered by health systems to be a prerequisite for implementing other intervention components and receiving clinician buy-in. Clinician education generated intended results when used on its own 57% of the time and, with other components, 78% of the time (Table 3). Clinician education often took the form of explaining the rationale or evidence for an intervention to clinicians during regular meetings or in grand rounds. Often it was a one-time event with regular reminders such as posters or emails reminding clinicians about the Choosing Wisely campaign and the current aim of an intervention. Wang and coauthors reported on an intervention in three family medicine clinics in which clinicians received hour-long educational presentations about appropriate use of lumbar spine MRIs. In this pre-post study, the authors found that the average number of monthly lumbar spine MRI studies decreased in the 10 months following the presentation.⁴⁰

Creation of New Clinical Pathways or Discontinuation Criteria. Some low-value services result from processes that happen automatically, or

nearly so. For example, children in the hospital with respiratory issues are often monitored with pulse oximetry in the absence of explicit criteria halting it.³³ Creating a new clinical pathway for care or establishing new criteria for discontinuation of a service may therefore reduce its use. When used on its own, which was rare ($n = 2$ studies; Table S11), the creation of new clinical pathways was effective half the time. However, interventions that used this component in conjunction with others ($n = 9$ studies; Table S11), such as education or alerts, were among the most successful, generating the intended results 86% of the time (Table 3). For example, Watnick and coauthors reported on a successful intervention in the emergency and inpatient departments at a children's hospital to reduce the use of chest x-rays for acute asthma hospitalizations.⁴¹ The intervention was multicomponent and included updating guidelines, changing the order sets, and educating clinicians about them. Additionally, the hospital updated its electronic ordering infrastructure to exclude a routine recommendation of chest x-ray and removed some indications for x-rays in the emergency department.

Clinician Feedback/Report Cards to Clinicians. Clinician report cards often take the form of periodic, individual feedback comparing a provider to peers or measuring an individual's progress to alignment with a specific benchmark. For example, a multicomponent intervention in a surgery department that used both department and provider-specific report cards to measure compliance with blood transfusion protocols was associated with improved transfusion practices and decreased costs (Table S12).⁴² Used on its own ($n = 4$ studies), clinician report cards achieved intended results in all outcomes 25% of the time. Used in conjunction with other components, including clinician education or point-of-care alerts, 80% of interventions ($n = 30$ studies) attained intended results (Table 3). Bhatia and coauthors describe an intervention using clinician feedback in the form of monthly emails summarizing each physicians' transthoracic echocardiogram ordering behavior, splitting orders into "appropriate," "maybe appropriate," and "rarely appropriate."⁴³ Prior to beginning the feedback, the authors sent clinicians a 20-minute video about the intervention and its rationale, and gave clinicians access to downloadable appropriate use criteria from the American College of Cardiology. Their study, notable because it was designed as an RCT and done across eight health systems in the United States and Canada, found lower rates of "rarely appropriate"

transthoracic echocardiograms in the group that received feedback, compared with control.

Creation of Organizational Change Frameworks. Organizational change frameworks can take a number of forms, but all have as an aim system-level cultural change and, often, assessment of progress toward organizational goals. These frameworks are typically used to support other intervention components and as part of a suite of larger changes within a system ($n = 9$ studies; Table S13). As such a support, the use of change frameworks generated intended results in all cases where results could be ascertained (Table 3). However, because this component can take different forms depending on the specific organization, it may be hard to replicate. For example, one initiative, in an inpatient general medicine unit at an academic medical center, provided support and encouraged individual faculty to lead targeted pilot projects to reduced unnecessary testing within their specific department.⁴⁴ Another, set within the Veterans Health Administration, used a pilot study with one VA health system to test an intervention to deintensify treatment for hypoglycemia risk. As it rolled out nationally, this initiative brought together workgroups with experts and other stakeholders and involved the creation of shared decision-making tools and new clinical pathways of care to identify patient candidates.³¹

Alternative Payment Methods. Because alternative payment mechanisms aim to create efficiency within health systems, they can be aligned with the goals of decreasing unnecessary medical services, including those targeted by Choosing Wisely. However, few studies of alternative payment specifically cite alignment with Choosing Wisely recommendations ($n = 4$; Table S14). Of the four studies we identified that explicitly identified Choosing Wisely, primarily by measuring change in services identified by Choosing Wisely lists, two generated intended results across all primary outcomes (Table 3). One of these reduced laboratory costs by 16%, while the other reduced daily charges for telemetry monitoring in an academic medical center by 69%.^{45,46} A third study, by Schwartz and coauthors, examined myriad low-value services and found overall reductions in low-value services among organizations participating in the Medicare accountable care organization program, though heterogeneity in reductions in individual types of services. For example, while low-value cancer testing dropped by 2.4% relative to the mean, preoperative services experienced no statistically significant change in use.⁴⁷

Consumer-Focused Intervention Components

Patient Cost Sharing. Just one study examined an intervention that used patient cost sharing to explicitly reduce use of Choosing Wisely services (Table S15), measuring whether people who switched to plans with high deductibles used fewer low-value services compared with prior to the switch.²⁶ Using data from a large commercial insurer on more than 300,000 patients, the researchers matched patients who switched plans with those who stayed in a traditional plan. Although the switch was associated with decreases in overall health care spending, the study did not find any effects on spending on low-value services, either in absolute terms or relative to overall decreases in spending.²⁶

Patient Education Materials or Facilitation of Informed Decision Making. Patient education materials or informed decision making, whether targeting patients only or both patients and providers, can be used to help patients understand when medical services may not be necessary. By empowering patients with information, patients may choose not to use potentially low-value services. On its own ($n = 2$ studies; Table S16), this intervention component was effective 50% of the time, and as a component of other interventions ($n = 6$ studies) it was effective 60% of the time (Table 3). In one study, which also included multiple clinician-focused components, handouts and videos were developed and disseminated in order to educate families of pediatric patients who were being treated for bronchiolitis.⁴⁸ That multicomponent intervention was successful in aligning care at the institution with Choosing Wisely guidelines for bronchiolitis. In another intervention, Engineer and coauthors reported on a successful multicomponent intervention that included clinician-focused elements and a structured parent discussion tool to guide discussion in cases of mild head injury in the emergency department with the aim of reducing head CT scans.³⁸ Head CT utilization in the emergency department was reduced from 63% of patients to 22% in that study.

Clinical Report Cards/Quality Reporting Directly to Patients. Instead of providing periodic feedback, or report cards, to clinicians about their utilization of low-value services, similar reports could be provided directly to patients. The idea is that patients might take overuse of unnecessary services as a sign of poor quality and steer away from such providers. This idea does not have much evidence to bolster it, at least in our review. We found one study that examined such an intervention

using Medicare's public reporting of physician imaging rates of low back pain (Table S17); its results are described in the section that follows.²⁹

Unintended Effects

Six total studies reported unintended effects for primary outcomes, defined as outcomes that were both statistically significant and in the opposite direction of outcomes congruent with Choosing Wisely guidelines. Four of those studies measured the impacts of the dissemination of Choosing Wisely guidelines and two measured an intervention. Of the two intervention-based studies, one of the analyzed studies generated only unintended results. In that study, researchers examined the impact of Medicare's public reporting of physician rates of imaging for low back pain prior to attempting conservative therapy in a cohort of Texas patients.²⁹ They found very little overall change in imaging rates (the statistically significant increase was small in magnitude) but did note, without hypothesizing why, that hospitals that had previously had lower inappropriate imaging rates increased their imaging rates after reporting started. The other study, which measured adherence to appropriate transfusion protocols, found mixed results, with an unexpected increase in plasma orders outside of hospital guidelines.⁴⁹ In that study, a multicomponent intervention in an academic hospital that included education, clinical decision support in the EHR, and provider report cards allowing for peer comparison, overall red blood cell use decreased and adherence to guidelines improved. Study authors hypothesized that the ubiquity of anticoagulant drugs and stronger evidence for red blood cell transfusion protocols compared to evidence for plasma guidelines may have led to the differences in outcomes.

Quality Assessment

Risk of Bias Within Studies

Twenty-two studies, or about 17% of the total sample, included a control group—four retrospective case-control studies, 10 nonrandomized trials, and eight RCTs (Table 1). The remaining 109 studies (84%) without controls were a mix of cross-sectional studies and case series that were framed as quality improvement studies involving single hospitals, provider groups, or entire health systems.

Table 4. Controlled Studies With Full Articles (*N* = 21) That Meet the Full ROBINS-I Bias Assessment Tool, by Study Type

Type of Study	Number of Studies That Met 100% of ROBINS-I Tool (%)
Case-control studies (<i>n</i> = 4)	3 (75)
Nonrandomized controlled trials (<i>n</i> = 9)	8 (89)
Randomized controlled trials (<i>n</i> = 8)	8 (100)

Study quality was assessed using an adapted version of the ROBINS-I tool published by the US Agency for Healthcare Research and Quality.¹⁶ The ROBINS-I tool includes seven bias domains: three (bias due to confounding, bias in selection of participants into the study, and bias in classification of interventions) that occur before or at the time of an intervention, and four (bias due to deviations from intended interventions, bias due to missing data, bias in measurement of outcomes, and bias in selection of the reported result) that occur after an intervention.⁶⁶ Only studies that met the complete ROBINS-I tool criteria for each study type were included in this table; one controlled study was excluded because it was described only via abstract.

We excluded one nonrandomized controlled study from the quality assessment because only the abstract was available,⁵⁰ bringing the final number of studies whose quality was assessed to 21. The quality of controlled studies was generally high, with the majority of studies (19 out of 21; 90%) meeting all the study design specifications needed to reduce bias based on the ROBINS-I tool (Table 4; study specifics in Table S18). Of the two studies that did not meet the entire bias assessment tool, one RCT failed to account for potential biases that arise from the randomization process and one case-control study failed to account for potential biases from missing data.

Overall, controlled studies were statistically just as likely to generate intended changes in alignment with Choosing Wisely guidelines as those that were not controlled (*p* = 0.6; Table 2). However, within different types of interventions there was some heterogeneity (Table S19). For example, whereas patient education materials were effective more than half the time across the full sample, the two studies that used a control found the intervention ineffective. In one of these studies, patients who were potential candidates for prostate-specific antigen screening were randomized to (1) usual care, (2) a decision aid without clinician interaction, or (3) a decision aid plus shared decision making, which included a discussion of the decision aid with a clinician. That

study found no differences between treatment and control arms.⁵¹ Conversely, controlled studies that used clinician education as part of a multicomponent intervention ($n = 8$ studies) found intended results 75% of the time, and controlled studies that gave clinicians periodic performance feedback found intended results in 5 out of the 7 studies (71%). Although the low number of studies makes it hard to draw firm conclusions, the analysis of controlled studies bolsters our conclusion that clinician-focused interventions tend to be more successful than patient-focused interventions, and points toward some specific interventions with solid evidence of effectiveness.

Risk of Bias Across Studies

Publication Bias. We found 25 studies in ClinicalTrials.gov that explicitly included Choosing Wisely in their trial description; 15 studies did not meet our inclusion criteria because they were not conducted in the United States ($n = 10$); were ongoing studies ($n = 3$); were withdrawn ($n = 1$); or had results that were published after our search ($n = 1$). Of the 10 remaining trials that were reported as completed, six trials had no final results reported in the database and no publications could be linked to the trial identification after extensive searching. Four registered trials from this verification process were included in our final sample.^{32,35,50,52}

Reporting Bias. Eight RCTs were included in our final literature sample, including two that mentioned Choosing Wisely in their trial registration noted above.^{32,35} Two RCTs included in our analysis were neither registered nor had published protocols^{51,53}; therefore, their final reported outcomes could not be verified or compared to any publicly available source. Two nonrandomized trials were registered,^{50,52} and they were assessed for selective reporting bias along with the six registered RCTs.

The planned and actual outcomes of the eight trials with published protocols are compared in Table S20. In three out of the eight studies, primary outcomes differed between the protocols and the final manuscripts; for example, one study intended to measure the 30-day equivalent of drug prescriptions but reported the total prescription days instead.⁵⁴ Additionally, the predetermined time frame for primary outcomes in at least four out of the eight studies also differed between protocol and final manuscript. Some studies acknowledged changes

between protocol and their final analysis,⁵² though most did not. Although changes between protocol and study implementation are not uncommon, the deviations we observed can lead to biases in the reported effects.

Discussion

We identified a significant number of studies that tested the impact of interventions to leverage the Choosing Wisely guidelines to reduce low-value services. The vast majority of interventions implemented and evaluated in empirical research were focused on changing the behaviors of clinicians and health care organizations. Consumer-oriented interventions had a small representation in the studies we reviewed.

Several interesting patterns emerged from our systematic review. First, for many intervention components, the majority of reviewed studies showed statistically significant effects in the intended direction. Second, the success rate was notably higher for studies of multicomponent interventions versus single-component interventions. Although the number of components past two did not make a significant difference in success rate, the complexity of the individual components was suggestive. Specifically, interventions that sought to create organizational change to support implementation of Choosing Wisely recommendations or involved multiple health care providers (arguably among the most complex interventions) had high success rates. This result aligns with most of the findings in the quality improvement and implementation science literature and adds important evidence to the ongoing debate about the effectiveness of single- versus multiple-component interventions.⁵⁵

Third, only about one in six studies had controls, underscoring the need for methodological rigor in future research. Overall, there was no indication that studies with controls were less likely to yield positive results than those without controls, though there was some heterogeneity within individual interventions.

Fourth, although the number of interventions targeting patients is low, the impact of clinician-focused interventions appeared to be more pronounced than that of consumer-focused interventions. Patient interventions were found largely ineffective. This finding differs from work in insurance design, which finds that increasing out-of-pocket

costs, whether through a value-based insurance design framework or a high-deductible health plan, can reduce low-value services.^{56–58} Within high-deductible plans, this reduction sometimes comes at the cost of also reducing high-value services.^{56,59} Consistent with this pattern, our review notes one study in which moving to a high-deductible plan lowered overall spending but not low-value service spending.²⁶ We did not review any studies that included value-based insurance design, though recent work has focused on using—and communicating—targeted increases in out-of-pocket costs as a way to reduce low-value service use.⁵⁷

Results of our review suggest dissemination of Choosing Wisely guidelines alone produces little success in reducing low-value care. Conversely, a number of interventions to implement Choosing Wisely guidelines, particularly those that are clinician-focused and multicomponent, have significant effects and produce desirable results. This echoes the recent and increased emphasis on implementation and the recognition that the broader context into which a guideline is introduced has substantial influence on whether the guideline can be successfully integrated into routine care. Implementation, as many have suggested, is as important as—or even more important than—dissemination of guidelines, because commitment to delivering high-value and cost-efficient care requires health care organizations put in place compatible interventions and allocate resources to fundamentally shift the practice patterns of physicians and other health care clinicians.^{60,61} However, guideline implementation in complex health care organizations is challenging and rarely follows a rational and linear pathway. Given the myriad factors that may influence health care delivery (e.g., ambiguities of evidence, multiple lines of authority, fragmentation of reimbursement), the common practice of implementation typically favors multifaceted approaches.^{55,62} Our review bears this out by showing that interventions with multiple components are more likely to be successful than interventions with a single component.

The literature on implementation also emphasizes the importance of broad social, economic, and political contexts outside health care organizations and the internal context of the medical practice. Our review, however, does not include an assessment of those contextual factors, which could be an interesting topic for future evaluations.

The literature review has several notable limitations related to the underlying literature. First, we report the possibility of publication and reporting bias that could lead to an overrepresentation of positive

studies in the literature. Though we undertook several steps to minimize the risk of publication bias (e.g., use of multiple research databases, inclusion of gray literature),^{63,64} our review of one widely used trial database found six registered trials that could potentially meet our inclusion criteria but whose results have not yet been reported or published. Unfortunately, mitigating sources of publication or reporting bias is largely beyond our control; these include prospective trial registration and improvements in the peer-review process and journal acceptance policies.⁶⁵ Given the potential for publication bias, the treatment effects and success rates reported here should be considered upper bounds.

Second, most studies were done in only one health system and reported with short follow-up times, limiting their generalizability and knowledge of long-term effects. Third, low-value service trials that did not explicitly mention Choosing Wisely were not included. This criteria may have led to the exclusion of some interventions in which the motivation was unstated, or to the exclusion of services on a Choosing Wisely list but not noted as such. Fourth, the initial screening of titles and abstracts of more than 13,000 potential articles was done by one author. To mitigate this limitation, we additionally consulted with a number of experts and did rigorous bibliographic tracing. Finally, to avoid the complication of interpreting results in multiple national contexts with varying health systems, our study was limited to studies within the United States. Although this limits our ability to draw conclusions about the prevalence or effectiveness of interventions in other nations, this review can provide a template for future research that seeks to assess the evidence for Choosing Wisely within or across nations. These comparisons are a fertile area for future research.

Overall, this review should fuel optimism among health care systems that thoughtful interventions can produce meaningful changes within their organizations. The Choosing Wisely initiative has been praised for involving multiple stakeholders in recommendation development. Health systems and payers should consider interventions to support these recommendations to improve quality and value within the health care system.

References

1. Cassel CK, Guest JA. Choosing Wisely: helping physicians and patients make smart decisions about their care. *JAMA*.

- 2012;307(17):1801-1802. <https://doi.org/10.1001/jama.2012.476>.
2. Levinson W, Born K, Wolfson D. Choosing Wisely campaigns: a work in progress. *JAMA*. 2018;319(19):1975-1976. <https://doi.org/10.1001/jama.2018.2202>.
 3. Rocque GB, Williams CP, Jackson BE, et al. Choosing Wisely: opportunities for improving value in cancer care delivery? *J Oncol Pract*. 2017;13(1):e11-e21. <https://doi.org/10.1200/JOP.2016.015396>.
 4. Patashnik EM, Dowling CM. Realizing the promise of Choosing Wisely will require changes both in the culture of specialty societies and in public policy. *Health Affairs Blog*. November 7, 2017. <https://www.healthaffairs.org/doi/10.1377/hblog20171107.649027/full/>. Accessed September 21, 2020.
 5. Colla CH, Morden NE, Sequist TD, Schpero WL, Rosenthal MB. Choosing Wisely: prevalence and correlates of low-value health care services in the United States. *J Gen Intern Med*. 2015;30(2):221-228. <https://doi.org/10.1007/s11606-014-3070-z>.
 6. Schwartz AL, Zaslavsky AM, Landon BE, Chernew ME, McWilliams JM. Low-value service use in provider organizations. *Health Serv Res*. 2018;53(1):87-119. <https://doi.org/10.1111/1475-6773.12597>.
 7. Cliff BQ, Hirth RA, Mark Fendrick A. Spillover effects from a consumer-based intervention to increase high-value preventive care. *Health Aff (Millwood)*. 2019;38(3):448-455. <https://doi.org/10.1377/hlthaff.2018.05015>.
 8. Kerr EA, Kullgren JT, Saini SD. Choosing Wisely: how to fulfill the promise in the next 5 years. *Health Aff (Millwood)*. 2017;36(11):2012-2018. <https://doi.org/10.1377/hlthaff.2017.0953>.
 9. Shrank WH, Rogstad TL, Parekh N. Waste in the US health care system: estimated costs and potential for savings. *JAMA*. 2019;322(15):1501. <https://doi.org/10.1001/jama.2019.13978>.
 10. Morden NE, Colla CH, Sequist TD, Rosenthal MB. Choosing Wisely—the politics and economics of labeling low-value services. *N Engl J Med*. 2014;370(7):589-592. <https://doi.org/10.1056/NEJMp1314965>.
 11. Colla CH, Malnor AJ. Choosing Wisely campaign: valuable for providers who knew about it, but awareness remained constant, 2014-17. *Health Aff (Millwood)*. 2017;36(11):2005-2011. <https://doi.org/10.1377/hlthaff.2017.0945>.

12. Colla CH, Mainor AJ, Hargreaves C, Sequist T, Morden N. Interventions aimed at reducing use of low-value health services: a systematic review. *Med Care Res Rev.* 2017;74(5):507-550. <https://doi.org/10.1177/1077558716656970>.
13. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol.* 2009;62(10):1006-1012. <https://doi.org/10.1016/j.jclinepi.2009.06.005>.
14. Cliff E, Lee S-YD, Hirth R, Avanceña ALV. *The impact of Choosing Wisely interventions on unnecessary medical services: a systematic review.* PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=140501. Published September 24, 2019. Accessed January 17, 2020.
15. OCEBM Levels of Evidence Working Group. *The Oxford 2011 levels of evidence.* Oxford, England: Oxford Centre for Evidence-Based Medicine; 2011. <https://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-2.1.pdf>. Accessed November 1, 2019.
16. Viswanathan M, Patnode CD, Berkman ND, et al. *Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions.* Rockville, MD: Agency for Healthcare Research and Quality; 2017. <https://doi.org/10.23970/AHRQEPDMETHGUIDE2>.
17. Campbell D, Stanley J. *Experimental and Quasi-Experimental Designs for Research.* Chicago, IL: Rand McNally; 1963.
18. Jin Z-C, Zhou X-H, He J. Statistical methods for dealing with publication bias in meta-analysis. *Stat Med.* 2015;34(2):343-360. <https://doi.org/10.1002/sim.6342>.
19. Joobar R, Schmitz N, Annable L, Boksa P. Publication bias: what are the challenges and can they be overcome? *J Psychiatry Neurosci.* 2012;37(3):149-152. <https://doi.org/10.1503/jpn.120065>.
20. Murad MH, Chu H, Lin L, Wang Z. The effect of publication bias magnitude and direction on the certainty in evidence. *BMJ Evid Based Med.* 2018;23(3):84-86. <https://doi.org/10.1136/bmjebm-2018-110891>.
21. Lin L, Chu H. Quantifying publication bias in meta-analysis. *Biometrics.* 2018;74(3):785-794. <https://doi.org/10.1111/biom.12817>.
22. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med.* 2009;6(7):e1000100. <https://doi.org/10.1371/journal.pmed.1000100>.

23. Drees M, Fischer K, Consiglio-Ward L, et al. Statewide antibiotic stewardship: an eBrightHealth Choosing Wisely initiative. *Del J Public Health*. 2019;5(2):50-58.
24. Ip IK, Raja AS, Gupta A, Andruchow J, Sodickson A, Khorasani R. Impact of clinical decision support on head computed tomography use in patients with mild traumatic brain injury in the ED. *Am J Emerg Med*. 2015;33(3):320-325. <https://doi.org/10.1016/j.ajem.2014.11.005>.
25. Hong AS, Ross-Degnan D, Zhang F, Wharam JF. Small decline in low-value back imaging associated with the 'Choosing Wisely' campaign, 2012-14. *Health Aff (Millwood)*. 2017;36(4):671-679. <https://doi.org/10.1377/hlthaff.2016.1263>.
26. Reid RO, Rabideau B, Sood N. Impact of consumer-directed health plans on low-value healthcare. *Am J Manag Care*. 2017;23(12):741-748.
27. Rosenberg A, Agiro A, Gottlieb M, et al. Early trends among seven recommendations from the Choosing Wisely campaign. *JAMA Intern Med*. 2015;175(12):1913-1920. <https://doi.org/10.1001/jamainternmed.2015.5441>.
28. Encinosa W, Davidoff AJ. Changes in antiemetic overuse in response to Choosing Wisely recommendations. *JAMA Oncol*. 2017;3(3):320-326. <https://doi.org/10.1001/jamaoncol.2016.2530>.
29. Ganduglia CM, Zezza M, Smith JD, John SD, Franzini L. Effect of public reporting on MR imaging use for low back pain. *Radiology*. 2015;276(1):175-183. <https://doi.org/10.1148/radiol.15141145>.
30. Jefferson BK, King JE. Impact of the acute care nurse practitioner in reducing the number of unwarranted daily laboratory tests in the intensive care unit. *J Am Assoc Nurse Pract*. 2018;30(5):285-292. <https://doi.org/10.1097/JXX.0000000000000050>.
31. Wright SM, Hedin SC, McConnell M, et al. Using shared decision-making to address possible overtreatment in patients at high risk for hypoglycemia: the Veterans Health Administration's Choosing Wisely Hypoglycemia Safety Initiative. *Clin Diabetes*. 2018;36(2):120-127. <https://doi.org/10.2337/cd17-0060>.
32. Kullgren JT, Krupka E, Schachter A, et al. Precommitting to choose wisely about low-value services: a stepped wedge cluster randomised trial. *BMJ Qual Saf*. 2018;27(5):355-364. <https://doi.org/10.1136/bmjqs-2017-006699>.
33. Schondelmeyer AC, Simmons JM, Statile AM, et al. Using quality improvement to reduce continuous pulse oximetry use in children with wheezing. *Pediatrics*. 2015;135(4):e1044-e1051. <https://doi.org/10.1542/peds.2014-2295>.

34. Mukerji A, Konz B, Gantt G, et al. Choosing Wisely: reductions in CT-scans for perianal abscesses. *Dis Colon Rectum*. 2019;62(6):e99.
35. Felcher AH, Gold R, Mosen DM, Stoneburner AB. Decrease in unnecessary vitamin D testing using clinical decision support tools: making it harder to do the wrong thing. *J Am Med Inform Assoc*. 2017;24(4):776-780. <https://doi.org/10.1093/jamia/ocw182>.
36. Chien AT, Lehmann LS, Hatfield LA, et al. A randomized trial of displaying paid price information on imaging study and procedure ordering rates. *J Gen Intern Med*. 2017;32(4):434-448. <https://doi.org/10.1007/s11606-016-3917-6>.
37. Matulis J, Liu S, Mecchella J, North F, Holmes A. Choosing Wisely: a quality improvement initiative to decrease unnecessary preoperative testing. *BMJ Qual Improv Rep*. 2017;6:bmjqr.u216281.w6691. <https://doi.org/10.1136/bmjquality.u216281.w6691>.
38. Engineer RS, Podolsky SR, Fertel BS, et al. A pilot study to reduce computed tomography utilization for pediatric mild head injury in the emergency department using a clinical decision support tool and a structured parent discussion tool. *Pediatr Emerg Care*. May 15, 2018. <https://doi.org/10.1097/PEC.0000000000001501>.
39. Coronel E, Bassi N, Donahue-Rolfe S, et al. Evaluation of a trainee-led project to reduce inappropriate proton pump inhibitor infusion in patients with upper gastrointestinal bleeding: skip the drips. *JAMA Intern Med*. 2017;177(11):1687-1689. <https://doi.org/10.1001/jamainternmed.2017.4851>.
40. Wang KY, Yen CJ, Chen M, et al. Reducing inappropriate lumbar spine MRI for low back pain: radiology support, communication and alignment network. *J Am Coll Radiol*. 2018;15(1):116-122. <https://doi.org/10.1016/j.jacr.2017.08.005>.
41. Watnick CS, Arnold DH, Latuska R. Successful chest radiograph reduction by using quality improvement methodology for children with asthma. *Pediatrics*. 2018;142(2):e20174003. <https://doi.org/10.1542/peds.2017-4003>.
42. Hicks CW, Liu J, Yang WW, et al. A comprehensive Choosing Wisely quality improvement initiative reduces unnecessary transfusions in an Academic Department of Surgery. *Am J Surg*. 2017;214(4):571-576. <https://doi.org/10.1016/j.amjsurg.2017.06.020>.
43. Bhatia RS, Ivers NM, Yin XC, et al. Improving the appropriate use of transthoracic echocardiography: the Echo WISELY Trial. *J Am Coll Cardiol*. 2017;70(9):1135-1144. <https://doi.org/10.1016/j.jacc.2017.06.065>.
44. Stinnett-Donnelly JM, Stevens PG, Hood VL. Developing a high value care programme from the bottom up: a programme of

- faculty-resident improvement projects targeting harmful or unnecessary care. *BMJ Qual Saf.* 2016;25(11):901-908. <https://doi.org/10.1136/bmjqs-2015-004546>.
45. Edholm K, Kukhareva P, Ciarkowski C, et al. Decrease in inpatient telemetry utilization through a system-wide electronic health record change and a multifaceted hospitalist intervention. *J Hosp Med.* 2018;13:531-536. doi:10.12788/jhm.2933.
 46. Yarbrough PM, Kukhareva PV, Horton D, Edholm K, Kawamoto K. Multifaceted intervention including education, rounding checklist implementation, cost feedback, and financial incentives reduces inpatient laboratory costs. *J Hosp Med.* 2016;11(5):348-354. <https://doi.org/10.1002/jhm.2552>.
 47. Schwartz AL, Chernew ME, Landon BE, McWilliams JM. Changes in low-value services in year 1 of the Medicare Pioneer accountable care organization program. *JAMA Intern Med.* 2015;175(11):1815-1825. <https://doi.org/10.1001/jamainternmed.2015.4525>.
 48. Tyler A, Krack P, Bakel LA, et al. Interventions to reduce over-utilized tests and treatments in bronchiolitis. *Pediatrics.* 2018;141(6):e20170485. <https://doi.org/10.1542/peds.2017-0485>.
 49. Thakkar RN, Kim D, Knight AM, Riedel S, Vaidya D, Wright SM. Impact of an educational intervention on the frequency of daily blood test orders for hospitalized patients. *Am J Clin Pathol.* 2015;143(3):393-397. <https://doi.org/10.1309/AJCPJS4EEM7UAUBV>.
 50. Mafi JN, Trotzky R, Wei E, et al. Evaluation of a Choosing Wisely intervention to reduce low-value antibiotic prescribing at a large safety net medical center. *J Gen Intern Med.* 2018;33(Suppl 2):S186-S187.
 51. Stamm AW, Banerji JS, Wolff EM, et al. A decision aid versus shared decision making for prostate cancer screening: results of a randomized, controlled trial. *Can J Urol.* 2017;24(4):8910-8917.
 52. Mafi JN, Godoy-Travieso P, Wei E, et al. Evaluation of an intervention to reduce low-value preoperative care for patients undergoing cataract surgery at a safety-net health system. *JAMA Intern Med.* 2019;179(5):648-657. <https://doi.org/10.1001/jamainternmed.2018.8358>.
 53. Raja AS, Ip IK, Dunne RM, Schuur JD, Mills AM, Khorasani R. Effects of performance feedback reports on adherence to evidence-based guidelines in use of CT for evaluation of pulmonary embolism in the emergency department: a randomized trial. *AJR Am J Roentgenol.* 2015;205(5):936-940. <https://doi.org/10.2214/AJR.15.14677>.

54. Sacarny A, Barnett ML, Le J, Tetkoski F, Yokum D, Agrawal S. Effect of peer comparison letters for high-volume primary care prescribers of quetiapine in older and disabled adults: a randomized clinical trial. *JAMA Psychiatry*. 2018;75(10):1003-1011. <https://doi.org/10.1001/jamapsychiatry.2018.1867>.
55. Harvey G, Kitson A. Translating evidence into healthcare policy and practice: single versus multi-faceted implementation strategies—is there a simple answer to a complex question? *Int J Health Policy Manag*. 2015;4(3):123-126. <https://doi.org/10.15171/ijhpm.2015.54>.
56. Brot-Goldberg ZC, Chandra A, Handel BR, Kolstad JT. What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics. *Q J Econ*. 2017;132(3):1261-1318. <https://doi.org/10.1093/qje/qjx013>.
57. Gruber J, Maclean JC, Wright B, Wilkinson E, Volpp KG. The effect of increased cost-sharing on low-value service use. *Health Econ*. 2020;29(10):1180-1201. <https://doi.org/10.1002/hec.4127>.
58. Zheng S, Ren ZJ, Heineke JD, Geissler KH. Reductions in diagnostic imaging with high deductible health plans. *Med Care*. 2016;54(2):110-117. <https://doi.org/10.1097/MLR.0000000000000472>.
59. Rabideau B, Eisenberg MD, Reid R, Sood N. Effects of employer-offered high-deductible plans on low-value spending in the privately insured population. *J Health Econ*. 2021;76,102424. <https://doi.org/10.1016/j.jhealeco.2021.102424>.
60. Rapport F, Clay-Williams R, Churruca K, Shih P, Hogden A, Braithwaite J. The struggle of translating science into action: foundational concepts of implementation science. *J Eval Clin Pract*. 2018;24(1):117-126. <https://doi.org/10.1111/jep.12741>.
61. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? *Psychiatry Res*. 2020;283:112376. <https://doi.org/10.1016/j.psychres.2019.04.025>.
62. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ*. 1998;317(7156):465-468. <https://doi.org/10.1136/bmj.317.7156.465>.
63. Parekh-Bhurke S, Kwok CS, Pang C, et al. Uptake of methods to deal with publication bias in systematic reviews has increased over time, but there is still much scope for improvement. *J Clin Epidemiol*. 2011;64(4):349-357. <https://doi.org/10.1016/j.jclinepi.2010.04.022>.
64. Knobloch K, Yoon U, Vogt PM. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and

- publication bias. *J Cranio-Maxillofac Surg.* 2011;39(2):91-92. <https://doi.org/10.1016/j.jcms.2010.11.001>.
65. Carroll HA, Toumpakari Z, Johnson L, Betts JA. The perceived feasibility of methods to reduce publication bias. *PLoS One.* 2017;12(10):e0186472. <https://doi.org/10.1371/journal.pone.0186472>.
66. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* 2016;355:i4919. <https://doi.org/10.1136/bmj.i4919>.

Funding/Support: The authors acknowledge funding for this project from the ABIM Foundation.

Conflict of Interest Disclosures: All authors completed the ICMJE Form for Disclosure of Potential Conflicts of Interest. No conflicts were reported.

Acknowledgments: We are grateful for research assistance from Judith Smith at the University of Michigan and Khat Naing at the University of Illinois Chicago.

Address correspondence to: Betsy Q. Cliff, University of Illinois Chicago School of Public Health, 1603 W. Taylor St, MC 923, Chicago, IL 60612 (email: bqcliff@uic.edu).

Supplementary Material

Additional supporting information may be found in the online version of this article at [http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1468-0009](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1468-0009):

Appendix S1. Search strategy

Figure S1. PRISMA diagram

Table S1. Types of interventions to reduce low-value service use

Table S2. Summary of interventions

Table S3. Articles that include recommendation guideline dissemination only

Table S4. Studies that include behavioral nudges

Table S5. Articles that include changes to order set or clinical documentation

Table S6. Articles that include clinical decision support: mandatory or optional utilization review

Table S7. Articles that include clinical decision support: point of care information or alert

Table S8. Articles that include increasing access or use of health information exchange

Table S9. Articles that include clinician champions

Table S10. Articles that include clinician education or academic detailing

Table S11. Articles that include creation of new clinical pathways or discontinuation criteria

Table S12. Articles that include clinician feedback or report cards to clinicians

Table S13. Articles that include creation of organizational change frameworks

Table S14. Articles that include risk-sharing or alternative payment methods

Table S15. Studies that include patient cost-sharing

Table S16. Studies that include patient education materials or informed decision-making

Table S17. Studies that include clinician report cards to patients

Table S18. Summary of quality assessment of controlled studies

Table S19. Controlled studies* with statistically significant results in the intended direction

Table S20. Risk of reporting bias among trials registered in Clinical-Trials.gov