Early Trends Among Seven Recommendations From the Choosing Wisely Campaign

Alan Rosenberg, MD; Abiy Agiro, PhD; Marc Gottlieb, MPA; John Barron, PharmD; Peter Brady, MBA; Ying Liu, MS; Cindy Li, MSc; Andrea DeVries, PhD

IMPORTANCE The Choosing Wisely campaign consists of more than 70 lists produced by specialty societies of medical practices or procedures of minimal clinical benefit to patients in most situations, with recommendations regarding judicious use.

OBJECTIVE To quantify the frequency and trends of some of the earliest Choosing Wisely recommendations using nationwide commercial health plan population-level data.

DESIGN, SETTING, AND PARTICIPANTS Retrospective analysis of claims data for members of Anthem-affiliated commercial health plans. The low-value services selected were (1) imaging tests for uncomplicated headache; (2) cardiac imaging without history of cardiac conditions; (3) low back pain imaging without red-flag conditions; (4) preoperative chest x-rays with unremarkable history and physical examination results; (5) human papillomavirus testing for women younger than 30 years; (6) use of antibiotics for acute sinusitis; and (7) use of prescription nonsteroidal anti-inflammatory drugs (NSAIDs) for members with hypertension, heart failure, or chronic kidney disease.

MAIN OUTCOMES AND MEASURES The number of members with medical and/or pharmacy claims for the included low-value services was assessed quarterly over a 2- to 3-year span through 2013. Trend changes in recommendations were evaluated across all quarters using Poisson regression with denominators as offsets.

RESULTS Two services had declines: Use of imaging for headache decreased from 14.9% to 13.4% (trend estimate, 0.99 [95% CI, 0.98-0.99]; P < .001), and cardiac imaging decreased from 10.8% to 9.7% (trend estimate, 0.99 [95% CI, 0.99-0.99]; P < .001). Two services had increases: Use of NSAIDs in select conditions increased from 14.4% to 16.2% (trend estimate, 1.02 [95% CI, 1.01-1.02]; P < .001), and human papillomavirus testing in younger women increased from 4.8% to 6.0% (trend estimate, 1.01 [95% CI, 1.00-1.01]; P < .001). Use of antibiotics for sinusitis remained stable (0.8% decrease from 84.5% to 83.7%; trend estimate, 1.00 [95% CI, 1.00-1.00]; P = .16). Use of preoperative chest x-rays (0.2% decrease, ending utilization 91.5%; trend estimate, 1.00 [95% CI, 1.00-1.00]; P = .70) and imaging for low back pain (53.7% utilization throughout study; P = .71) remained high with no statistically significant changes.

CONCLUSIONS AND RELEVANCE For this population-level analysis of 7 low-value services analyzed, changes were modest but showed a desirable decrease for 2 recommendations (imaging for headache, cardiac imaging for low-risk patients). The effect sizes were marginal, however, and although 4 of the 7 lists had statistically significant changes—unsurprising given the large sample size—the clinical significance is uncertain. These results suggest that additional interventions are necessary for wider implementation of Choosing Wisely recommendations.

Published online October 12, 2015. Corrected on October 19, 2015.
Reducing use of unnecessary medical procedures and treatments is a key component to controlling health care expenditures. As much as 30% of US health care expenditures are for interventions with marginal clinical benefit for patients and unnecessary diagnostic procedures that may lead to inappropriate treatment or a cascade of additional tests or procedures.1–4

The Choosing Wisely campaign represents a physician-driven effort to create conversations between physicians and patients around overuse and waste. The campaign was piloted in 2009 by the National Physicians Alliance, funded by the American Board of Internal Medicine (ABIM) Foundation, and was expanded in 2012.3 Choosing Wisely currently consists of more than 70 lists of approximately 400 recommendations produced by specialty societies of frequently used medical practices or procedures that provide minimal clinical benefit to patients in most situations.5,6 Each list includes recommendations regarding judicious use of tests, services, or medications in situations in which there is likely to be low value returned to the patient.7–8

Because each society produced its own list, the processes for choosing the services included in the lists varied from society to society. The ABIM Foundation, however, reviewed the lists and provided general parameters, such as requirements for evidence-based recommendations and inclusion of only frequently performed services.8

Whereas specialty societies have long promoted practice guidelines, typically guideline dissemination occurs within the confines of peer-reviewed journals, teaching hospitals, and conferences. That is, dissemination occurs within a specific physician area of specialty. The Choosing Wisely campaign, in contrast, involves a more public process, for example, partnering with Consumer Reports and including website material specifically directed toward patients.

Such lists draw attention to low-value services, but they must be translated into measurable recommendations to assess their effect on changing behavior.5,9 A recent article called for the ABIM Foundation and third-party payers to collaborate in creating high-impact lists whereby payers could estimate the volume, effect on quality, and cost of services listed in Choosing Wisely.5

The present study examined metrics related to a range of recommendations appearing on lists released early in the Choosing Wisely campaign. The goal was to assess the frequency at which use of these “low-value” services occurred, and to determine whether, after several years, use of the low-value services occurred more or less frequently. This study is among the first to evaluate a variety of recommendations from the Choosing Wisely campaign using nationwide population-level data from commercial health plans.

### Methods

#### Data Sources and Study Design

This retrospective analysis used medical and pharmacy claims from Anthem-affiliated Blue Cross and Blue Shield health care plans for approximately 25 million members across the United States. All study data were kept anonymous throughout to safeguard patient confidentiality; researchers only accessed a limited data set, devoid of individual patient identifiers, in full compliance with relevant provisions of the Health Insurance Portability and Accountability Act. This nonexperimental study, conducted under the Research Exception provisions of the Privacy Rule 45 CFR 164.514(e), was thus exempt from institutional review board approval and informed consent was not required.

#### Measures for Seven Choosing Wisely Recommendations

For this analysis, we identified 7 services from the initial Choosing Wisely lists published in April 2012.8,10 Each selected service could be analyzed via commercial insurance claims data using the International Classification of Diseases, Ninth Revision; Current Procedural Terminology codes; and whenever appropriate, National Drug Codes and Logical Observation Identifiers Names and Codes (Table 1 and eAppendix A in the Supplement). They also reflected the variety of services and specialty societies participating in the Choosing Wisely campaign and had been published relatively early in the campaign, allowing time for adequate dissemination, and they were anticipated to have more frequent use.

Four of the 7 recommendations included in the analysis related to diagnostic imaging, consistent with the proportion of Choosing Wisely recommendations pertaining to imaging5: (1) imaging tests for headache with uncomplicated conditions, (2) cardiac imaging for members without a history of cardiac conditions, (3) preoperative chest x-rays with unremarkable history and physical examination results, and (4) low back pain imaging for members without red-flag conditions. One recommendation related to age guidelines for cervical cancer screening: (5) human papillomavirus (HPV) testing for women younger than 30 years. Last, 2 recommendations related to inappropriate medication use: (6) antibiotics for acute sinusitis and (7) prescription nonsteroidal anti-inflammatory drugs (NSAIDs) for members with select chronic conditions (hypertension, heart failure, or chronic kidney disease).11

Whereas cardiac testing represented a compilation of services listed by several specialty organizations (the American Academy of Family Physicians, the American College of Cardiology, the American College of Physicians, and the American Society of Nuclear Cardiology), the other services selected were listed by 1 or 2 societies. Preoperative chest x-rays were listed by the American College of Radiology and the American College of Physicians, whereas imaging for headache was listed only by the American College of Radiology; low back pain imaging, sinusitis antibiotics, and HPV testing were listed only by the American Academy of Family Physicians; and NSAID use was listed only by the American Society of Nephrology.

The rationale for including specific services in the Choosing Wisely lists varied, although most services selected typically yield a low level of clinically relevant information, such as imaging for headache, certain cardiac imaging, preoperative chest x-rays in certain situations, and low back pain imaging (Table 1). Use of NSAIDs to treat musculoskeletal pain in those with hypertension, heart failure, or chronic kidney disease was included because these drugs may increase blood pressure (interfering with antihypertensive medications) or cause fluid retention (which may worsen kidney function).
Table 1. Choosing Wisely Recommendation Details

<table>
<thead>
<tr>
<th>Description</th>
<th>Rationale</th>
<th>Denominator and Numerator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagnostic Imaging</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Headache imaging</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do not do imaging for uncomplicated headache (American College of Radiology)</td>
<td>Imaging patients with headache absent specific risk factors for structural disease is not likely to change management or improve outcome. Those patients with a significant likelihood of structural disease requiring immediate attention are detected by clinical screens that have been validated in many settings. Many studies and clinical practice guidelines concur. Also, incidental findings lead to additional medical procedures and expense that do not improve patient well-being.</td>
<td>Denominator: Age ≥18 y and ≤50 y and has current member eligibility. At least 1 outpatient visit for uncomplicated headache during the past 90 d (index event). Index date is earliest (first) date with uncomplicated headache diagnosis during the quarter. Numerator: Members in denominator with ≥1 head computed tomographic scan or ≥1 head magnetic resonance image from index date to index date + 30 d.</td>
</tr>
<tr>
<td>Cardiac testing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do not perform annual EKGs or any other cardiac screening for low-risk patients without symptoms (American Academy of Family Physicians)</td>
<td>There is little evidence that detection of coronary artery stenosis in asymptomatic patients at low risk for coronary heart disease improves health outcomes. False-positive test results are likely to lead to harm through unnecessary invasive procedures, overtreatment, and misdiagnosis.</td>
<td>Denominator: Age ≥30 y at time of visit and ≥1 office visit with a primary care physician or cardiology (index event). Date of office visit in quarter is index date. Excluded members considered to be at cardiac risk and visits with certain unrelated primary diagnoses. Note: A 5% random sample was used to construct this metric. Numerator: Test from members in denominator with ≥1 claim for EKG, echocardiography, or stress test from index date to index date + 30 d.</td>
</tr>
<tr>
<td>Do not perform stress cardiac imaging or advanced noninvasive imaging in the initial evaluation of patients without cardiac symptoms unless high-risk markers are present (American College of Cardiology)</td>
<td>Asymptomatic, low-risk patients account for up to 45% of unnecessary screening. Testing should be performed only when the following findings are present: diabetes mellitus in patients &gt;40 y, peripheral arterial disease, or ≥2% yearly risk for coronary heart disease events.</td>
<td></td>
</tr>
<tr>
<td>Do not perform annual stress cardiac imaging or advanced noninvasive imaging as part of routine follow-up in asymptomatic patients (American College of Cardiology)</td>
<td>Performing stress cardiac imaging or advanced noninvasive imaging in patients without symptoms on a serial or scheduled pattern (eg, every 1-2 y or at a heart procedure anniversary) rarely results in any meaningful change in patient management. This practice may, in fact, lead to unnecessary invasive procedures and excess radiation exposure without any proven effect on patients’ outcomes.</td>
<td></td>
</tr>
<tr>
<td>Do not perform stress cardiac imaging or advanced noninvasive imaging as a preoperative assessment in patients scheduled to undergo low-risk noncardiac surgery (American College of Cardiology)</td>
<td>Noninvasive testing is not useful for patients undergoing low-risk noncardiac surgery (eg, cataract removal). These types of tests do not change the patient’s clinical management or outcomes and will result in increased costs.</td>
<td></td>
</tr>
<tr>
<td>Do not perform EKG as routine follow-up for mild, asymptomatic native valve disease in adults with no change in signs or symptoms (American College of Cardiology)</td>
<td>Patients with native valve disease usually have years without symptoms before the onset of deterioration. An EKG is not recommended yearly unless there is a change in clinical status.</td>
<td></td>
</tr>
<tr>
<td>Do not obtain screening exercise EKG testing in individuals who are asymptomatic and at low risk for coronary heart disease (American College of Physicians)</td>
<td>In asymptomatic individuals at low risk for coronary heart disease (10-y risk &lt;10%), screening for coronary heart disease with exercise EKG does not improve patient outcomes.</td>
<td></td>
</tr>
<tr>
<td>Do not perform stress cardiac imaging or coronary angiography in patients without cardiac symptoms unless high-risk markers are present (American Society of Nuclear Cardiology)</td>
<td>Asymptomatic, low-risk patients account for up to 45% of inappropriate stress testing. Testing should be performed only when the following findings are present: diabetes mellitus in patients &gt;40 y, peripheral arterial disease, and ≥2% yearly coronary heart disease event rate.</td>
<td></td>
</tr>
<tr>
<td>Do not perform cardiac imaging for patients who are at low risk (American Society of Nuclear Cardiology)</td>
<td>Patients with chest pain who are at low risk of cardiac death and myocardial infarction (based on history, physical examination, EKGs, and cardiac biomarkers) do not merit stress radionuclide myocardial perfusion imaging or stress echocardiography as an initial testing strategy if they have a normal EKG result (without baseline ST-segment abnormalities, left ventricular hypertrophy, preexcitation, bundle branch block, intraventricular conduction delay, paced rhythm, or digoxin therapy) and are able to exercise.</td>
<td></td>
</tr>
<tr>
<td>Do not perform radionuclide imaging as part of routine follow-up in asymptomatic patients (American Society of Nuclear Cardiology)</td>
<td>Performing stress radionuclide imaging in patients without symptoms on a serial or scheduled pattern (eg, every 1-2 y or at a heart procedure anniversary) rarely results in any meaningful change in patient management. This practice may lead to unnecessary invasive procedures and excess radiation exposure without any proven effect on patients’ outcomes.</td>
<td></td>
</tr>
<tr>
<td>Do not perform cardiac imaging as a preoperative assessment in patients scheduled to undergo low- or intermediate-risk noncardiac surgery (American Society of Nuclear Cardiology)</td>
<td>Noninvasive testing is not useful for patients undergoing low-risk noncardiac surgery or with no cardiac symptoms or clinical risk factors undergoing intermediate-risk noncardiac surgery. These types of testing do not change the patient’s clinical management or outcomes and will result in increased costs.</td>
<td></td>
</tr>
<tr>
<td>Use methods to reduce radiation exposure in cardiac imaging, whenever possible, including not performing such tests when limited benefits are likely (American Society of Nuclear Cardiology)</td>
<td>The key step to reduce or eliminate radiation exposure is appropriate selection of any test or procedure for a specific person, in keeping with medical society recommendations, such as appropriate use criteria. Clinicians should incorporate new methodologies in cardiac imaging to reduce patient exposure to radiation while maintaining high-quality test results.</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Table 1. Choosing Wisely Recommendation Details (continued)

<table>
<thead>
<tr>
<th>Description</th>
<th>Rationale</th>
<th>Denominator and Numerator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preoperative chest x-ray</strong></td>
<td>In the absence of cardiopulmonary symptoms, preoperative chest radiography rarely provides any meaningful changes in management or improved patient outcomes</td>
<td>Denominator: Any operating room procedure (inpatient or outpatient) for patients ≥18 y. Excluded certain procedures including those related to cardiopulmonary disease. Numerator: Members in the denominator with a chest x-ray within 14 d prior to the index date</td>
</tr>
<tr>
<td><strong>Avoid admission or preoperative chest x-rays for ambulatory patients with unremarkable history and physical examination results (American College of Radiology)</strong></td>
<td>Performing routine admission or preoperative chest x-rays is not recommended for ambulatory patients without specific reasons suggested by the history and/or physical examination findings. Only 2% of such images lead to a change in management. Obtaining a chest radiograph is reasonable if acute cardiopulmonary disease is suspected or there is a history of chronic stable cardiopulmonary disease in a patient &gt;70 y who has not had a chest radiograph within 6 mo.</td>
<td></td>
</tr>
<tr>
<td><strong>Low back pain imaging</strong></td>
<td>Red flags include, but are not limited to, severe or progressive neurological deficits or when serious underlying conditions such as osteomyelitis are suspected. Imaging of the lower spine before 6 weeks does not improve outcomes but does increase costs. Low back pain is the fifth most common reason for all physician visits.</td>
<td>Denominator: Age between 19 and 50 y and ≥1 claim for low back pain as the primary diagnosis from the emergency department or office visit. First occurrence is index date. Excluded members with previous low back pain diagnoses in past 180 d and certain diagnoses such as injuries and cancers. Numerator: Members in denominator with ≥1 claim for low back pain imaging from index date to index date +28 d</td>
</tr>
<tr>
<td><strong>Sinusitis antibiotics</strong></td>
<td>Symptoms must include discolored nasal secretions and facial or dental tenderness when touched. Most sinusitis in the ambulatory setting is due to a viral infection that will resolve on its own. Despite consistent recommendations to the contrary, antibiotics are prescribed in more than 80% of outpatient visits for acute sinusitis. Sinusitis accounts for 16 million office visits and $5.8 billion in annual health care costs. Viral infections cause the majority of acute rhinosinusitis and only 0.5% to 2% progress to bacterial infections. Most acute rhinosinusitis resolves without treatment in 2 weeks. Antibiotics are not recommended for patients with uncomplicated acute rhinosinusitis who have mild illness and assurance of follow-up. If a decision is made to treat, amoxicillin should be first-line antibiotic treatment for most acute rhinosinusitis.</td>
<td>Denominator: Age ≥18 y and has current member eligibility and prescription coverage. At least 1 outpatient claim for acute sinusitis in the first diagnosis position during past 90 d. Excluded members with acute sinusitis visit in prior 365 d or visiting an ear, nose, and throat specialist. Numerator: Members in denominator with ≥1 claim for antibiotics from index date to index date +30 d</td>
</tr>
<tr>
<td><strong>NSAID use</strong></td>
<td>The use of NSAIDS, including cyclo-oxygenase type 2 inhibitors, for the pharmacological treatment of musculoskeletal pain can elevate blood pressure, make antihypertensive drugs less effective, cause fluid retention, and worsen kidney function in these individuals. Other agents such as acetaminophen, tramadol, or short-term use of narcotic analgesics may be safer and as effective as NSAIDs.</td>
<td>Denominator: Age ≥18 y and has current member eligibility for both medical and pharmacy. At least 1 outpatient visit for musculoskeletal pain during the past 90 d (index event). Index date is the earliest (first) date with musculoskeletal pain diagnosis during the quarter. At least 1 diagnosis code over the prior 365 d from index date that indicates that member has hypertension, kidney disease, or heart failure. Note: An 8% random sample was used to construct this metric. Numerator: Members in denominator with ≥1 prescription for NSAID from index date to index date +30 d</td>
</tr>
<tr>
<td><strong>Cervical Cancer Screening</strong></td>
<td>There is adequate evidence of moderate harms from HPV testing in women &lt;30 y, including additional testing and invasive diagnostic procedures plus psychological stress or anxiety associated with abnormal screening results</td>
<td>Denominator: Females &lt;30 y. Members counted in each quarter actively enrolled. Age based on the member’s age at the end of the current quarter. Numerator: Members in denominator with ≥1 claims for HPV testing during the past 12 mo</td>
</tr>
</tbody>
</table>

Abbreviations: EKG, electrocardiogram; HPV, human papillomavirus; NSAID, nonsteroidal anti-inflammatory drug.

**Target Population Identification Criteria**

Each of the 7 recommendations analyzed had separate target, or study, populations. The denominators and numerators for each recommendation (Appendix B in the Supplement) were determined by specific criteria relevant to that recommendation (Table I). For 5 of the 7 recommendations analyzed, we strived to include all of the target population that met our study criteria (Table I and Appendix A in the Supplement). For 2 measures, however, the number of patients meeting denominator criteria was too large to include the entire target population. Consequently, we used a 5% random sample for cardiac imaging.
without a history of cardiac conditions and an 8% random sample for prescription NSAIDs for members with select chronic conditions (hypertension, heart failure, or chronic kidney disease). We selected a random sample percentage that optimized the analysis while capturing as much of the target populations as possible.

The number and percentage of members with claims for low-value services for each recommendation were assessed quarterly for at least 10 quarters through the third quarter of 2013. The starting point for each measure ranged from 2010 through 2011 on the basis of data availability. The 7 Choosing Wisely lists were published in the second quarter of 2012.

The HPV testing recommendation was a population-based measure (ie, the denominator definition depended on age and sex only); consequently, a member could be counted more than once in the denominator (or numerator) calculation each instance the criteria were met. The remainder of the recommendations were event-driven measures, so a member would qualify for the numerator or denominator on the basis of an event that took place in a given quarter. The same member could be recounted in future quarters but only if that member had a new qualifying event in a subsequent quarter. For both population- and event-based measures, members were counted only once per quarter.

Statistical Analysis
The descriptive statistics (numerator, denominator, and frequency) of each recommendation is provided in Appendix B in the Supplement. We assessed trend changes in each recommendation across all quarters using Poisson regression with offsets. The dependent variable consisted of numerator counts in each quarter while the independent variable, as a measure of trend effect. The log of denominators in each quarter served as an offset. Poisson regression with offsets assumes a linear change in trends over time. Poisson regression with population size (denominator) offsets offers a natural way of analyzing aggregate data in the absence of patient-level data sets. The trend estimates should be interpreted as exponentiated rates per quarter as exponentiation was applied to estimates and confidence intervals. All analyses were conducted using SAS, version 9.4. Statistical significance was set at $P < .05$.

Statistical Process Control Charts
In addition to the regression analysis, we plotted the quarterly frequencies using statistical process control charts. These charts are often used as a tool for continuous improvement or quality control in monitoring process measures. In this case, quarterly data (ie, frequency of low-value services from the 7 recommendations) were plotted along with 3 reference lines (Figure). The raw mean of frequencies across all quarters for each recommendation was graphed as the “average” reference line. The standard deviation of frequencies across all quarters for each recommendation was calculated and used to define upper and lower limits (Appendix B in the Supplement).

Results
During the study, the percentage of members with imaging for headaches with uncomplicated conditions decreased from 14.9% to 13.4%, a relative reduction of 10.1%. On the basis of Poisson regression, the trend effect was decreasing to a statistically significant extent (trend estimate, 0.99 [95% CI, 0.98-0.99]; $P < .001$). The percentage of members with cardiac imaging in the absence of cardiac disease (using a 5% random sample) decreased from 10.8% to 9.7, a 10.2% relative reduction. The Poisson trend effect was decreasing to a statistically significant extent (trend estimate, 0.99 [95% CI, 0.99-0.99]; $P < .001$).

In contrast, the percentage of preservative chest x-rays and low back pain imaging in the absence of red-flag conditions remained stable although high: the percentage for preservative chest x-rays began at 91.3% and ended at 91.5%, a relative change of 0.2%; the percentage for low back pain imaging began and ended at 53.7%. There was no Poisson trend effect for preservative chest x-rays (trend estimate, 1.00 [95% CI, 1.00-1.00]; $P = .70$) and low back pain imaging (trend estimate, 1.00 [95% CI, 1.00-1.00]; $P = .71$). Use of antibiotics for sinusitis decreased slightly, from 84.5% to 83.7%, a relative reduction of 0.9%. The Poisson trend effect was not statistically significant (trend estimate, 1.00 [95% CI, 1.00-1.00]; $P = .16$). Sinusitis antibiotic use exhibited minor seasonality that did not alter the overall stable trend.

The percentage of women younger than 30 years receiving HPV testing increased from 4.8% to 6.0%, a 25.0% relative increase. The Poisson trend effect was increasing to a statistically significant extent (trend estimate, 1.01 [95% CI, 1.00-1.01]; $P < .001$). Last, the use of prescription NSAIDs by members with select chronic conditions increased from 14.4% to 16.2%, a 12.5% relative increase. The Poisson trend effect was increasing to a statistically significant extent (trend estimate, 1.02 [95% CI, 1.01-1.02]; $P < .001$).

The statistical process control charts shown in the Figure support the general findings from the regression analysis in terms of trend directionality. For example, imaging for headache showed a modestly consistent decline with all 3 quarterly measures for 2013 falling below the lower limit. In contrast, use of NSAIDs by members with select chronic conditions showed an increasing trend in which all 3 quarterly measures for 2013 were at or above the upper limit.

Discussion
The Choosing Wisely campaign represents an innovative approach toward changing physician and patient attitudes, via the greater emphasis on including the patient in the conversation. However, the campaign currently lacks a structural mechanism to evaluate its influence on clinical practice or to gauge whether additional efforts are necessary. Accessing claims data allowed an early evaluation of Choosing Wisely recommendations on low-value services based on nationwide population-level data. Prior studies on Choosing Wisely recommendations either focused on 1 or 2 recommendations or consolidated different recommendations into a single summary measure. To our knowledge, this study encompassed the largest number of patients studied to date and was the first to separately analyze a number of recommendations.
Figure. Trends for Selected Low-Value Services

A. All 7 low-value services

B. Low back pain imaging without red-flag conditions

C. Headache imaging for uncomplicated conditions

D. Use of antibiotics for acute sinusitis

E. Cardiac imaging without history of cardiac conditions

F. Use of NSAIDs for members with selected conditions

G. Preoperative chest x-ray with unremarkable history

H. Testing for HPV in women younger than 30 y

HPV indicates human papillomavirus; NSAID, nonsteroidal anti-inflammatory drug.
For the 7 low-value services analyzed, trend changes were modest but showed a desirable decrease for 2 recommendations (imaging for headache, cardiac imaging for low-risk patients). These results were statistically significant, unsurprising given the large sample size. The effect sizes were marginal, however, and may not represent clinically significant changes. Increasing trends were observed for HPV testing in women younger than 30 years and in 1 list focused on medication use (NSAIDs for people with certain chronic conditions), an undesirable trend in light of the Choosing Wisely recommendations. There was no significant change in the remaining 3 metrics: antibiotic use for sinusitis, preoperative chest x-rays, and low back pain imaging in the absence of red-flag conditions. Although the trend changes were statistically significant for 4 of the 7 lists analyzed, the clinical significance is uncertain. Our mixed results highlight the need for interventions beyond the current level of promotion, such as data feedback, physician communication training, systems interventions (eg, clinical decision support in electronic medical records), clinician scorecards, patient-focused strategies, and financial incentives.15

Although HPV testing was used by fewer than 10% of the members in the target population, other metrics in this study showed much higher use. Antibiotics for sinusitis were used by 8 of 10 members in the target population, preoperative chest x-rays were performed in 9 of 10 of the target, and low back pain imaging was performed in 5 of 10 members of the target population. The fact that some procedures or treatments were used quite often (eg, preoperative chest x-rays) while others were relatively rare (eg, HPV testing) highlights the importance of establishing frequency when creating recommendations and priorities. We initially looked at 2 additional recommendations (Papanicolaou tests for women younger than 21 or with hysterectomy for noncancer disease and sinusitis imaging for acute mild-to-moderate sinusitis or uncomplicated acute rhinitis) but found them to be used so infrequently as to be unsuitable for trend analysis. In fact, a recent publication showing that preoperative cardiac stress testing was infrequently used in the target population before Choosing Wisely lists were published also highlighted the need for specialty societies to focus Choosing Wisely lists on procedures with a high baseline frequency of use.13

Frequency of the most highly used services either remained stable (preoperative chest x-rays and low back pain imaging) or decreased slightly (sinusitis antibiotics) during the study period, which underscores the view that simple publication of recommendations—such as the Choosing Wisely lists—is insufficient to produce major changes to practice.16 The continued evaluation of low-value services will allow payers to use validated lists to inform quality improvement programs, as well as coverage, payment, and utilization management decisions.

This analysis did not examine patient-level risk differences and regional variations. Because our population-level data analysis will not be able to account for confounding, the potential effects of individual or clinician-level variations could be the focus of future studies. By examining each recommendation separately, this analysis preserved the effects of factors such as payment incentives, malpractice concerns, and patient demand on each specific recommendation.16 It is possible that assessing the local implementation of Choosing Wisely recommendations at the delivery system level could be a fruitful area for future studies.

Our study has several limitations. Our analysis is based on administrative claims data that do not adequately capture the clinical circumstances that led to ordering a service, which may be especially important for recommendations on NSAIDs. Therefore, the recommendations that we defined may include some services that might be appropriate for an individual patient. In addition, claims data are subject to coding inaccuracies in terms of diagnostic designation and the reported timing of services. Because our trend analysis was based on quarterly data, we may have excluded some patients because workup and treatments could span more than 1 calendar quarter. However, as a population-level analysis, this will probably not affect the modest findings in this study.
It is likely that several factors in addition to the Choosing Wisely campaign may have been responsible for changes or lack of changes in frequency of low-value services. The first is secular trends and preexisting patterns. For example, during the study period, guidelines from different organizations regarding Papanicolaou testing in young women converged around the notion that those younger than 21 years should not get Papanicolaou tests. This might explain the relatively rare use of Papanicolaou tests. At the same time, HPV testing became more widely available in the marketplace and its use increased over time.

In fact, the changes or lack of changes in frequency of use of low-value services after publication of Choosing Wisely recommendations could be affected by changes in the nature and size of the target population over time, or in criteria used for the underlying clinical diagnoses. Therefore, creating and validating measures for Choosing Wisely recommendations is another area for future studies.

Conclusions

This analysis of 7 Choosing Wisely recommendations provides a starting point for further evaluation of the influence of the initiative on changing behavior by analyzing changes in volume of these services from the early years of the initiative. Our population-level analysis showed both desirable and undesirable modest trends in use of low-value services. The relatively small use changes suggest that additional interventions are necessary for wider implementation of Choosing Wisely recommendations in general practice. Some of the additional interventions needed include data feedback, physician communication training, systems interventions (eg, clinical decision support in electronic medical records), clinician scorecards, patient-focused strategies, and financial incentives.