Measuring the effect of Choosing Wisely: an integrated framework to assess campaign impact on low-value care

R Sacha Bhatia,1,2 Wendy Levinson,1 Samuel Shortt,3 Ciara Pendrith,2 Elana Fric-Shamji,2 Marjon Kallewaard,4 Wilco Peul,5 Jeremy Veillard,6 Adam Elshaug,7 Ian Forde,8 Eve A Kerr9

ABSTRACT
The Choosing Wisely campaign began in the USA in 2012 to encourage physicians and patients to discuss inappropriate and potentially harmful tests, treatments and procedures. Since its inception, the campaign has grown substantially and has been adopted by 12 countries around the world. Of great interest to countries implementing the campaign, is the effectiveness of Choosing Wisely to reduce overutilisation. This article presents an integrated measurement framework that may be used to assess the impact of a Choosing Wisely campaign on physician and provider awareness and attitudes on low-value care, provider practice behaviour and overuse of low-value services.

INTRODUCTION
Ensuring the sustainability of healthcare systems is a critical concern in countries around the world. Between 2000 and 2013, mean worldwide healthcare expenditures have risen from 7.7% to 9.3% of Gross domestic product (GDP).1 Some of these healthcare expenditures constitute low-value care practices that are unlikely to benefit patients or which may even harm them.2 3 In response to concerns of unnecessary care, 12 countries developed campaigns modelled after Choosing Wisely, a campaign that began in the USA with the goal of reducing overuse. Choosing Wisely was launched 2012 by the American Board of Internal Medicine (ABIM) Foundation as an attempt to change physician practice behaviour by harnessing physician leadership and increasing patient/public awareness regarding low-value tests, procedures and treatments, and the risks they pose to patients.4 5 While each country is tailoring their campaign to meet the unique needs of their country, a central part of the campaign in the USA and other countries are the evidence-based lists of tests, treatments or medical procedures that physicians and patients should question created by medical specialty societies. In the USA and in Canada, the campaign is also providing educational materials to physicians and patients about low-value testing.

Though interest in Choosing Wisely is high among providers and the public, the effectiveness and sustainability of a campaign like Choosing Wisely will depend in large part on the ability to demonstrate impact in the short term and long term. In the short term, the goals of Choosing Wisely campaigns are to produce clear changes in awareness and attitudes of physicians and trainees regarding stewardship of health system resources. However, to be effective in the longer term, these campaigns must change provider behaviour, increase patient knowledge of overuse and ultimately decrease utilisation of services that provide negligible benefit. Campaigns like Choosing Wisely can also have unintended consequences, like underuse of necessary services, dissatisfaction by patients and negative provider experiences. While researchers have begun reporting baseline rates of Choosing Wisely recommendations,6–8 a comprehensive approach to gauging campaign impact has not yet been undertaken. In this paper, we propose an integrated measurement framework that can be used to assess the impact of efforts to reduce low-value care. This paper also aims to
articulate the barriers faced when measuring overuse and what data elements are necessary for measurement efforts to be successful.

OVERUSE AS A FORGOTTEN MEASURE OF QUALITY: THE NEED TO STUDY THE IMPACT OF CHOOSING WISELY CAMPAIGNS

The Institute of Medicine (IOM) classified all healthcare quality problems into three categories: underuse, overuse and misuse. In recent years, underuse of effective therapies and misuse of services have received the most attention from quality improvement initiatives and patient safety efforts. Interestingly, research on the issue of overuse predates that of underuse, spanning back to the late 1960s and early 1970s. Early studies examining the geographical variation of service delivery suggested that a large proportion of medical procedures were not necessary. During the 1980s, researchers at RAND Corporation developed a method to rate the appropriateness of medical services from appropriate, where expected benefits clearly exceed risks, to inappropriate, where expected harms outweigh expected benefits. It has only been in the past few years that overuse began to garner more attention.

Similar to initiatives aimed at decreasing underuse, limiting overuse takes focused improvement initiatives that in many cases are potentially expensive. Evaluating whether these initiatives work is the first step in assessing whether the effort is ‘worth it’. Effective campaigns should have three essential hallmarks: (1) focus on areas that have high baseline rates of overuse and in which limiting overuse is likely to improve outcomes of care (directly or downstream); (2) tailored interventions that address major barriers faced by both providers and patients to motivate change; (3) use clinically meaningful measures to understand the effect of the campaign on physician attitudes, physician behaviour and ordering behaviour, and patient experiences, including both intended and unintended consequences. Although all three key elements are important, measure of overuse and metrics to assess physician and patient attitudes towards overuse are necessary to achieve the first two stated objectives. Therefore, we have chosen to focus this paper on effective measurement of the Choosing Wisely Campaign.

MEASUREMENT TOOLS TO ASSESS THE IMPACT OF CHOOSING WISELY

Table 1 shows a measurement framework to assess the impact of a Choosing Wisely campaign. We identified three broad categories of impact: physician attitudes, physician behaviours and patient engagement and acceptance. Under each of these categories, we have identified various methods for how they may be measured, time frames of expected impact and advantages/disadvantages of each approach. We will expand on each category below.

Assessment of physician attitudes, knowledge and perceptions towards the issue of low-value care and its causes and challenges

One of the early markers of the impact of Choosing Wisely campaigns is provider awareness that more is not always better. Surveying physicians is probably the most straightforward, standardisable and cost-effective way of assessing physicians’ awareness and attitudes. Ideally, one would assess change in perceptions before and after the start of a campaign and compare the change in perceptions with those of physicians in a concurrent control group who are similar in basic characteristics but are not exposed to the campaign at the same time. Detailed survey data can also provide nuanced information about the underlying drivers of low-value care and challenges in reducing low-value care. These data can be subsequently used to help target specific Choosing Wisely recommendations and can guide the development of specific interventions designed to reduce low-value care.

The ABIM Foundation assessed provider attitudes towards low-value services and to the Choosing Wisely campaign. The 2014 survey of 600 US physicians found that 73% (438) of these physicians believed that over-testing was a problem in the health system. Seventy-two per cent (432) of physicians reported that they ordered at least one unnecessary test or treatment per week, and 47% (282) said patients request an unnecessary test or treatment at least once per week.

In this survey, 21% (126) of physicians had heard of the Choosing Wisely campaign. Survey data can not only be used to gauge the magnitude of the problem and the level of awareness, but can also help expose underlying drivers of low-value care. For example, in the same survey 52% (312) of physicians cited malpractice concerns, 36% (216) safety concerns and 28% (168) said patient preference drove the physicians to order a low-value test. Beyond general drivers of low-value care, detailed survey information can provide insight into causes of ordering specific recommendations. For example, in a survey of Ontario family physicians, 79% (1,021) thought patients would have a Q15 difficult time accepting the recommendation to not be prescribed antibiotics for sinusitis, compared with only 33% (426) who thought patients would have a difficult time accepting the recommendation, “Do not repeat colorectal cancer screening (by any method) for 10 years after a high-quality colonoscopy is negative in average-risk individuals”. These results suggest an intervention to reduce antibiotic prescription in sinusitis may require a greater degree of patient education than colorectal screening.

In addition, survey data can illustrate some of the challenges to reducing low-value care. For example, in the same survey of Ontario family physicians, 95% (1,228) said patient requests were a major or minor barrier to implementing the Choosing Wisely recommendations, compared with 64% (836) who cited a
lack of automated decision support and 48% (610) as payment policies as either major or minor barriers. Survey data, both general and specifically geared towards individual recommendations, can guide the selection of interventions (ie, decision support tools, patient or provider education, audit and feedback, policy interventions, etc) and also aid in intervention design. Survey data presented to different physician groups (generalists vs specialists) can also help address critical differences in attitude, knowledge and perceptions that would also aid in developing specific interventions. Finally, standardised survey tools distributed by organisations like The Commonwealth Fund can help assess similarities across countries regarding ordering practices and also discover country-specific differences that may relate to payment schemes.

| Table 1 Measurement sources for assessment of Choosing Wisely initiatives |
|-----------------------------|------------------|------------------|
| **Measurement area**        | **Measurement tools** | **Time frame for Choosing Wisely impact** |
|                             |                   | **Advantages of approach**                                                                 |
|                             |                   | **Disadvantages of approach**                                                                 |
| Provider attitudes and awareness | Physician surveys | Short term | Inexpensive Can reach large number of physicians Tailor questions for specific physician groups Can resample over time to assess change Can compare physicians response across various health systems |
|                             | Structured interviews | Short term | More detailed information about attitudes May expose factors that can be used to improve campaign |
| Provider ordering behaviour: overuse of low-value services | Administrative databases | Medium to long term | Direct measurements of overuse at the individual patient level |
|                             | Electronic health record (EHR)/chart data | Medium to long term | More clinical details (past medical history, previous testing) to assess appropriateness Can assess a greater variety of recommendations than with administrative data As well as measuring the ‘index’ low-value care event, enables measurement of downstream effects (eg, harms and side effects) Can provide regional variation on overuse Can assess claims across health systems Can reassess over time to assess change Can monitor for unintended consequences such as underuse |
| Patient perceptions and outcomes | Patient-reported experience measures (PREMs; survey tools) | Short to medium term | Standardised data collection Patient perceptions/drivers could help design interventions to reduce low-value care Impact of interventions on patient perceptions could monitor for unintended consequences Comparable across jurisdictions |
|                             | Patient-reported outcome measures (PROMs; validated survey tools) | Medium term | Validated patient tools assist providers in developing care plans, including tests and treatments Results can monitor for unintended consequences of interventions |

delivery systems, policy differences or cultural factors. This can help countries identify best practices in each of these areas and help guide policy makers, particularly in the areas of medical education, payment reform and health human resource planning.

**Measurement of provider behaviours**

Most important to payors and health policy makers is the impact of the Choosing Wisely campaign on utilisation of low-value care, which is predominantly related to physician practising behaviours and to a lesser extent patient demand. To determine the impact of a Choosing Wisely campaign on ordering behaviour, the recommendations must be translated into measurable indicators to determine the baseline utilisation of these indicators prior to and after the initiation of the Choosing Wisely campaign. Unfortunately, the clinical complexity of the Choosing Wisely recommendations means that the recommendations often do not lend themselves easily to measurement. Nonetheless, successful attempts have been made to determine the baseline utilisation rate of a number of the Choosing Wisely recommendations, and measuring utilisation has become an emerging area of research.

**Administrative data**

Linked population-based administrative databases, such as insurance claims or hospital discharge databases, are commonly used to measure low-value care across a health system. Administrative databases contain patient-level data, which allow for direct measurement of appropriateness compared with indirect measures, such as studies of geographic variation. There have been numerous recent studies using administrative data that highlight the frequency of low-value care as defined by Choosing Wisely recommendations, mainly in the USA. For example, Schwartz and colleagues developed 26 claims-based measures for low-value services, 16 of which were based on Choosing Wisely recommendations, for assessment in US Medicare population. They found that low-value services affected 42% of Medicare beneficiaries and were responsible for 0.6–2.7% of overall Medicare spending, depending on whether a more specific or sensitive coding algorithm was used. Another study used Medicare claims data to assess the prevalence of cardiac testing in low-risk patients and found that 13% received cardiac tests without a clear indication, while Kerr et al. showed that stress tests before low-risk surgery are very rarely performed. Morden and colleagues also used claims data to show that 1 out of 10 bone mineral density tests was done more frequently than recommended by Choosing Wisely.

Administrative data have utility in not only determining utilisation rates of low-value services across a region but more importantly by highlighting variation between institutions and groups of providers. For example, Colla and colleagues found rates of cardiac testing in low-risk patients varied more than threefold between hospital referral regions (HRRs) in the USA. Another study quantifying overuse using Medicare claims data from 1,451,142 patients found 14% (217,672) of Medicare patients experienced at least one overuse event a year, but this rate varied from 8% to 27% across HRR.

There are large gaps in the overuse literature, likely due to the challenges of developing standardised measures. Korenstein and colleagues found that relatively few low-value services have been extensively studied, including antibiotics for upper respiratory tract infections, coronary artery bypass grafting, coronary angiography and some diagnostic imaging. Low-value services that are convenient and feasible to measure reliably are most frequently studied, but the lack of clinical granularity in administrative databases precludes measuring more clinically relevant recommendations. Schwartz and colleagues highlight the challenge of defining specific exclusions for low-value care metrics using administrative data. Sensitive measures that do not capture important exclusions in the population eligible for a service (the denominator) may label services as being overused that are actually appropriately used for some individuals (false-negative overuse); specific measures that attempt to limit such false positives may miss important instances of overuse (false-negative overuse). Appropriately specifying the denominator to include only those patients who clearly did not need the test, procedure or treatment measured in the numerator is the main challenge of using administrative data for measuring overuse because many clinically relevant factors are not coded.

Despite the limitations of administrative data to accurately capture utilisation of some of the detailed Choosing Wisely recommendations, understanding trends in physician ordering nationally and regionally has important value. It allows policy makers to identify trends in utilisation of low-value services over time and identify problem areas to focus possible policy-related interventions. Furthermore, it allows institutions and providers to identify regional, institutional and perhaps even provider-level variation in the utilisation of low-value services that could serve as the catalyst for local quality improvement efforts, particularly if integrated as part of a broader quality improvement reporting programme. Finally, standardised measures of low-value care could be used to evaluate the effectiveness of implementation efforts.

**Clinical data**

Accessing clinical data through manual chart reviews or electronic health records (EHRs) may address some of the pitfalls of using administrative data to measure overutilisation. Rather than measuring aggregate rates of potentially inappropriate use from administrative
databases, medical records contain the clinical details necessary to assess appropriateness on an individual level. Much of the available evidence on overuse pre-dates the Choosing Wisely campaign and comes from manual chart reviews assessing the appropriateness of a given service at a single or few sites. For example, Brook and colleagues found that 23% of coronary angiographies and 64% of carotid endarterectomies performed at several hospitals were for inappropriate indications. Clinical information available from records also enables the measurement of a broader range of indicators that could not be measured by administrative data alone.

Despite having several advantages over administrative databases, there are challenges of using medical records to measure low-value services. Chart review is a resource-intensive process requiring several abstractors, and the lack of standardised definitions for qualifiers such as ‘low risk’ may limit inter-rater reliability. In addition, there are still significant gaps in the types of data available. For example, the easily accessible coded fields of EHRs rarely contain a comprehensive listing of symptoms, which often drives testing. Furthermore, while administrative data can assess systematic overuse, the laborious nature of chart abstraction may limit the capacity to sample broadly.

**Patient experience and outcomes**

While much of the initial focus of the Choosing Wisely campaign has been on providers, a key consideration of these campaigns must be on their impact on patients, and that impact should be measured in a structured way. Patient-reported experience measures (PREMs) and patient-reported outcome measures (PROMs) are two measurement dimensions that are growing in importance. The measurement of the patient experience is important to assess both the potential positive aspects of the Choosing Wisely campaign on overuse (ie, less inconvenience and/or expense from fewer unnecessary tests) and the potential negative consequences (ie, damage to the physician–patient relationship). Though currently not well established as a measurement dimension in overuse research, PREMs and PROMs will likely take on a greater role assessing a Choosing Wisely campaign’s impact over time.

**Patient-reported experience measures (PREMs)**

Patient-reported data on the experience of the care they received and interaction with the healthcare team is often overlooked, yet vital part of evaluating the impact of Choosing Wisely campaigns on low-value care, and potentially assessing for some of the unintended consequences of initiatives designed to reduce overuse. A PREM is a measure of a patient’s perception of the healthcare experience they have received and often focuses on the elements of that experience that matter to them. PREMs are most often validated survey tools that can be used to provide both a snapshot of the patient experience and a longitudinal assessment to determine progress on these metrics. While there are currently no tools that specifically discuss low-value care, many of the questions asked in these surveys can identify underlying issues that may either address root causes of overuse (such as trust and communication) or may help inform potential interventions to reduce low-value care. For example, if an institution with a high rate of antibiotic prescriptions for viral infections scored lower on the patient question, “provider explained things in a way that was easy to understand,” an intervention might be designed to enhance patient education about unnecessary antibiotics and to improve patient–provider communication on the topic.

**Patient-reported outcome measures (PROMs)**

PROMs are standardised methods, often through survey instruments, of determining patients’ views of their symptoms, functional status and health-related quality of life. PROMs may either be disease specific (ie, the Kansas City Cardiomyopathy Questionnaire) or generic (ie, *EuroQol EQ-5D*). PROMs can assist clinicians and patients make more patient-centred treatment decisions, can be used to compare providers and institutions and can evaluate the impact of health delivery interventions over time. Similar to the PREMs, specific questions about overuse or low-value care have not been addressed in the majority of these questionnaires; however, structured patient input can help providers and patients make care decisions that may limit unnecessary tests and treatments. PROMs (and PREMs) should be integrated into the decision support for clinicians when setting out care individual plans and also by health systems as ways of tracking the aggregate patient perceptions with care as a response to initiatives such as Choosing Wisely campaigns. For example, in deciding on whether a patient with back pain needs low back pain imaging, the use of the Bournemouth Questionnaire (BQ) for low back pain would provide clinicians with an objective measure of back pain severity over time. Integrated into a decision support tool, a lack of improvement in this PROM could indicate the need for low back pain imaging, such as a CT or MRI.

**ASSESSING UNINTENDED CONSEQUENCES FOR CHOOSING WISELY CAMPAIGNS**

As recommendations to limit overuse are being developed and implemented into practice as part of Choosing Wisely type initiatives, we must be diligent to assess for the unintended consequences of such campaigns. In particular, focus on reducing overuse of low-value care may lead to underuse of high-value care. In order to assess for underuse, measures must be specific, ensure data for exclusions are readily available and importantly also routinely monitor for...
Table 2  Integrative framework for measuring overuse (using low back pain as an example)

<table>
<thead>
<tr>
<th>Goal</th>
<th>Example</th>
<th>Data source</th>
<th>Specific measure example</th>
<th>Action items</th>
<th>Additional data needed for robust measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change physician attitudes to clinical practice</td>
<td>Assessment of provider awareness/attitudes of Choosing Wisely recommendation on low back pain imaging</td>
<td>Provider survey</td>
<td>Validated physician survey on Choosing Wisely awareness and attitudes</td>
<td>Understanding of provider awareness/understanding of Choosing Wisely recommendation</td>
<td>Data from a representative sample of physicians to avoid non-response bias</td>
</tr>
<tr>
<td></td>
<td>Assessment of provider satisfaction with care provided for low back pain</td>
<td>Provider interviews</td>
<td>Semi-structured interviews</td>
<td>Tailor provider education on low back pain imaging</td>
<td>Linked responses from same physicians to track changes in attitudes over time</td>
</tr>
<tr>
<td>Foster patient engagement and acceptance</td>
<td>Determine patient perceptions regarding care for low back pain</td>
<td>Patient-reported experience measures</td>
<td>Consumer Assessment of Healthcare Providers and Systems (CAHPS) Clinician and Group Survey (with expanded questions for low-value care)</td>
<td>Monitor for unintended consequences (eg, clinician–patient trust, provider satisfaction)</td>
<td>Measures from patients in areas not exposed to Choosing Wisely to assess campaign impact and monitor for unintended consequences</td>
</tr>
<tr>
<td></td>
<td>Assessment of patient-reported outcomes such as pain, function, work loss</td>
<td>Patient-reported outcome measures</td>
<td>Oswestry Low Back Pain Disability Questionnaire (OSW)</td>
<td>Provide patient input in developing interventions to reduce low-value care</td>
<td>Linked experience and outcome measures to assess the relationship between patient care experience and outcomes</td>
</tr>
<tr>
<td>Change key clinical practices</td>
<td>Determine practice-specific rates of imaging for uncomplicated low back pain</td>
<td>Clinical data</td>
<td>Rate of CT/MRI excluding red flags (ie, history of cancer)</td>
<td>Feedback to providers on ordering behaviour—identification of outliers Validation of administrative data on true degree of overuse Data to administrators to implement local interventions to reduce low-value care Assessment of impact on local Choosing Wisely efforts</td>
<td>Data from other practices to provide context of rate and similar practices not exposed to Choosing Wisely to assess impact of local interventions Clinical data from all points of access to healthcare system (eg, walk-in clinics, ER visits)</td>
</tr>
<tr>
<td></td>
<td>Determine practice-specific rates of high-value imaging for low back pain</td>
<td>Clinical data</td>
<td>Rate of indicated CT/MRI (red flags)</td>
<td>Monitor for unintended consequences (eg, underuse of high-value services)</td>
<td></td>
</tr>
</tbody>
</table>
underuse of high-value services or the overuse of alternative services. For example, high compliance with the Choosing Wisely recommendation, "do not perform positron emission tomography (PET), CT, and radionuclide bone scans in the staging of early prostate cancer at low risk for metastasis," could result in underuse of bone scans among patients who are at high risk for metastasis. Finally, campaigns may put pressure on the clinician–patient relationship in ways that deteriorate trust, cause patient dissatisfaction and cause provider stress. Such unintended effects need to be assessed with specific survey questions directed at patients and physicians.

AN INTEGRATIVE APPROACH TO MEASURING CHOOSING WISELY

Table 2 shows an integrative approach to measurement of a Choosing Wisely campaign, using the recommendation on imaging for low back pain as an example. This approach shows the multiple areas that should be assessed simultaneously to fully assess the intended and unintended consequences of a campaign on physician attitudes, provider ordering behavior and patient experience and outcomes.

<table>
<thead>
<tr>
<th>Goal</th>
<th>Example</th>
<th>Data source</th>
<th>Specific measure example</th>
<th>Action items</th>
<th>Additional data needed for robust measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promote alignment with the healthcare system</td>
<td>Determine regional rates of imaging for uncomplicated low back pain imaging (CT/MRI scans)</td>
<td>Administrative data</td>
<td>Rate of CT/MRI utilisation per 1000 population</td>
<td>Data to policy makers to inform policy development (i.e., payment reforms)</td>
<td>Utilisation rates from areas not exposed to Choosing Wisely to assess campaign effects</td>
</tr>
<tr>
<td></td>
<td>Determine regional rates of alternate imaging for uncomplicated low back pain imaging</td>
<td>Administrative data</td>
<td>Rate of X-ray utilisation per 1000 population</td>
<td>Monitor for unintended consequences (e.g., increased use of alternate procedures)</td>
<td></td>
</tr>
</tbody>
</table>

WHAT IS NEEDED FOR THIS APPROACH TO BE SUCCESSFUL?

First, Choosing Wisely campaigns should create standard survey tools that could allow for appropriate comparisons of provider knowledge and attitudes and patient experiences within and between countries. Second, campaigns should engage with the research and policy communities to create standard data on regional variation of low-value services estimated from administrative data and identifying 'hot spots' of high utilisation to local administrators and policy makers. Finally, campaigns should engage with the research and policy communities to create standard data on regional variation of low-value services estimated from administrative data and identifying 'hot spots' of high utilisation to local administrators and policy makers.

Table 2

<table>
<thead>
<tr>
<th>Goal</th>
<th>Example</th>
<th>Data source</th>
<th>Specific measure example</th>
<th>Action items</th>
<th>Additional data needed for robust measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promote alignment with the healthcare system</td>
<td>Determine regional rates of imaging for uncomplicated low back pain imaging (CT/MRI scans)</td>
<td>Administrative data</td>
<td>Rate of CT/MRI utilisation per 1000 population</td>
<td>Data to policy makers to inform policy development (i.e., payment reforms)</td>
<td>Utilisation rates from areas not exposed to Choosing Wisely to assess campaign effects</td>
</tr>
<tr>
<td></td>
<td>Determine regional rates of alternate imaging for uncomplicated low back pain imaging</td>
<td>Administrative data</td>
<td>Rate of X-ray utilisation per 1000 population</td>
<td>Monitor for unintended consequences (e.g., increased use of alternate procedures)</td>
<td></td>
</tr>
</tbody>
</table>

Narrative review


BMJ Qual Saf: first published as 10.1136/bmjqs-2015-004070 on 19 June 2015. Downloaded from

http://qualitysafety.bmj.com/
definitions for low-value services that could be measured based on the priority of the recommendations. The measurement definitions can be derived from administrative data, when possible, but should encompass EHR or other clinical data sources when necessary. In this way, evaluations are driven by what is important to measure, rather than what is simply expedient. This also allows for future administrative databases and information technology (IT) infrastructure to incorporate these definitional elements. Additionally, Choosing Wisely campaigns should work with the research community to try and expand the breadth of services that are measurable by either administrative or clinical data, particularly recommendations where more clinical details are necessary to assess overuse. In particular, in table 2 we list current data gaps in overutilisation measurement. Finally, Choosing Wisely campaigns should work with research funders and payors to encourage research into overutilisation, particularly in the area of measurement and implementation science. This will aid Choosing Wisely campaigns in understanding the challenge of overuse and encourage the development and testing of solutions that could be generalised across multiple jurisdictions.

CONCLUSIONS
We have proposed a multi-pronged approach to measurement of overuse. We recognise that there are significant challenges to implementing this approach for hospitals and health systems, including a lack of expertise in data management and information systems, and quality improvement infrastructure. We also recognise that hospitals and health systems face many day-to-day challenges that they must prioritise. Undertaking this comprehensive measurement approach, however, will help these hospitals understand local practice patterns and areas of potential unnecessary use of services that increase costs, and may reduce the capacity to serve needed patients. Finally, by systematically measuring potential areas of overuse, it allows health systems to target high-yield recommendations first and avoid significant time and resources tackling recommendations that may not be a significant local problem. Thus, there is both a financial argument and a quality of care argument for health systems to engage in this endeavour. Therefore, we think a comprehensive measurement approach, if implemented in a thoughtful and systematic way, could result in improved health system efficiency and better patient outcomes.

Author affiliations
1Department of Medicine, University of Toronto, Toronto, Canada
2Institute for Health System Solutions and Virtual Care, Women’s College Hospital, Toronto, Canada
3Health Policy Consultant, Ottawa, Canada
4Dutch Association of Medical Specialists, Utrecht, The Netherlands
5Department of Neurosurgery, Leiden University Medical Centre, Leiden, The Netherlands
6Canadian Institute for Health Information, Toronto, Canada
7Menzies Centre for Health Policy, University of Sydney, Sydney, New South Wales, Australia
8OECD, London, UK
9VA Center for Clinical Management Research, VA Ann Arbor Healthcare System, and Department of Internal Medicine and Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, USA

Funding This work has been funded by The Commonwealth Fund and the Canadian Institutes of Health Research.

Contributors All authors participated in the conception and design of the article. RSB led the drafting of the article, but all authors participated in critical revisions and granted final approval of the submitted manuscript.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


Elshaug AG, McWilliams JM, Landon BE. The value of low-value lists. JAMA 2013;309:775–6.


