

# Precommitting to choose wisely about low-value services: a stepped wedge cluster randomised trial

Jeffrey Todd Kullgren,<sup>1,2,3,4</sup> Erin Krupka,<sup>5</sup> Abigail Schachter,<sup>6</sup> Ariel Linden,<sup>7</sup> Jacquelyn Miller,<sup>4</sup> Yubraj Acharya,<sup>6</sup> James Alford,<sup>8</sup> Richard Duffy,<sup>9</sup> Julia Adler-Milstein<sup>3,5,6</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2017-006699>).

For numbered affiliations see end of article.

## Correspondence to

Dr Jeffrey Todd Kullgren, VA Ann Arbor Healthcare System and University of Michigan, PO Box 130170, Ann Arbor, Michigan 48113, USA; [jkullgre@med.umich.edu](mailto:jkullgre@med.umich.edu)

Received 14 March 2017  
Revised 8 September 2017  
Accepted 7 October 2017  
Published Online First  
24 October 2017



► <http://dx.doi.org/10.1136/bmjqs-2017-006778>  
► <http://dx.doi.org/10.1136/bmjqs-2017-007477>



**To cite:** Kullgren JT, Krupka E, Schachter A, et al. *BMJ Qual Saf* 2018;**27**:355–364.

## ABSTRACT

**Background** Little is known about how to discourage clinicians from ordering low-value services. Our objective was to test whether clinicians committing their future selves (ie, precommitting) to follow Choosing Wisely recommendations with decision supports could decrease potentially low-value orders.

**Methods** We conducted a 12-month stepped wedge cluster randomised trial among 45 primary care physicians and advanced practice providers in six adult primary care clinics of a US community group practice. Clinicians were invited to precommit to Choosing Wisely recommendations against imaging for uncomplicated low back pain, imaging for uncomplicated headaches and unnecessary antibiotics for acute sinusitis. Clinicians who precommitted received 1–6 months of point-of-care precommitment reminders as well as patient education handouts and weekly emails with resources to support communication about low-value services. The primary outcome was the difference between control and intervention period percentages of visits with potentially low-value orders. Secondary outcomes were differences between control and intervention period percentages of visits with possible alternate orders, and differences between control and 3-month postintervention follow-up period percentages of visits with potentially low-value orders.

**Results** The intervention was not associated with a change in the percentage of visits with potentially low-value orders overall, for headaches or for acute sinusitis, but was associated with a 1.7% overall increase in alternate orders ( $p=0.01$ ). For low back pain, the intervention was associated with a 1.2% decrease in the percentage of visits with potentially low-value orders ( $p=0.001$ ) and a 1.9% increase in the percentage of visits with alternate orders ( $p=0.007$ ). No changes were sustained in follow-up.

**Conclusion** Clinician precommitment to follow Choosing Wisely recommendations was associated with a small, unsustainable decrease in potentially low-value orders for only one of three targeted conditions and may have increased alternate orders.

**Trial registration number** NCT02247050; Pre-results.

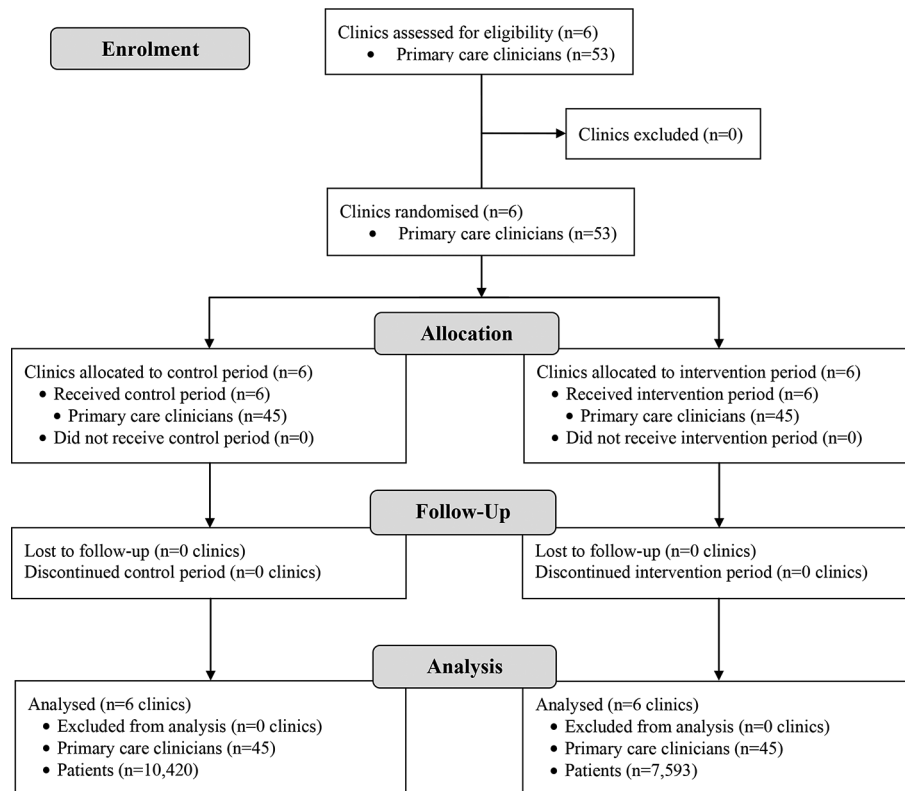
## INTRODUCTION

Low-value health services constitute care that in most cases does not improve patient outcomes and can lead to unnecessary

harms.<sup>1</sup> In the USA and many other countries, delivery of low-value services is common<sup>2</sup> and costly.<sup>3</sup> While the *Choosing Wisely* campaign has drawn substantial attention to evidence about what constitutes low-value care,<sup>4</sup> awareness of evidence alone is rarely enough to change many clinical decisions<sup>5,6</sup> without complementary strategies that target factors that underlie these decisions.<sup>7</sup> However, little is known about how to effectively target factors that drive ordering of low-value services.<sup>8,9</sup>

Two factors that may encourage clinicians to order low-value services are the context in which clinicians typically make these decisions and cognitive biases that can arise in such contexts. First, even when clinicians are aware that certain services do not improve population-level outcomes, decisions to order these services for individual patients are often made quickly during brief clinic visits,<sup>10</sup> in which both multitasking<sup>11</sup> and patient requests for services<sup>12,13</sup> are common and can over-ride intentions to avoid ordering low-value services. Second, making choices in such busy environments often requires ‘fast’ thinking that relies more on heuristics than on ‘slow’ thinking about evidence.<sup>14–17</sup> These contexts and cognitive biases can lead to divergence from evidence-based recommendations to avoid ordering low-value services.

A promising solution to this problem that leverages insights from behavioural economics could be to encourage clinicians to start decisions about low-value services before patient encounters when their thinking is more likely to be ‘slow’ and deliberative.<sup>14–16</sup> While engaged in this slower thinking, clinicians could be asked to commit to following specific *Choosing Wisely* recommendations<sup>18,19</sup> against



**Figure 1** CONSORT flow diagram.

ordering low-value services in future clinical encounters. Behavioural economists term this strategy ‘precommitment’, or the act of committing one’s future self to a course of action.<sup>18 20–22</sup> Precommitment has been shown to improve behaviours that can be influenced by choice environments and cognitive biases such as savings,<sup>23</sup> academic performance,<sup>18</sup> tobacco use<sup>24 25</sup> and weight loss.<sup>26 27</sup> Such precommitment interventions tie consequences to a failure to achieve the course of action.<sup>21</sup> For clinicians who precommit to avoid ordering low-value services, these consequences include abdication of their professional responsibility to steward resources,<sup>28</sup> which could be highlighted through point-of-care supports when ‘fast’ thinking<sup>14</sup> is required and cognitive biases could lead to divergence from their precommitment. Importantly, this approach could be readily integrated into routine care in a range of settings and would preserve clinicians’ autonomy around ordering decisions, which can be integral to sustained intervention acceptance.<sup>29</sup> Our objective was to test whether this behavioural economic strategy of inviting clinicians to precommit to avoid ordering low-value health services and providing supports to promote adherence to precommitment could decrease orders for potentially low-value services for three common conditions in adult primary care clinics.

## METHODS

### Design overview

We conducted a stepped wedge cluster randomised controlled trial between 1 July 2014 and 30 June 2015 in six primary care clinics of Integrated Healthcare

Associates (IHA), a multispecialty group practice in southeast Michigan in the USA. In this study design, the intervention started in a new clinic each month in a randomly assigned order. We chose this design because of its feasibility and because other potential designs would have had critical limitations.<sup>30–32</sup> Specifically, randomisation of individual physicians or patients would have been vulnerable to cross-contamination. A parallel-design, cluster, randomised trial would have been infeasible because resource constraints necessitated starting the intervention in one clinic at a time, and could have been unacceptable to research partners as half of the clinics would have not have received the intervention.

The intervention consisted of inviting clinicians to precommit to follow *Choosing Wisely* recommendations against imaging for uncomplicated low back pain,<sup>33</sup> imaging for uncomplicated headaches<sup>34</sup> and unnecessary antibiotics for acute sinusitis.<sup>35</sup> Clinicians who precommitted received 1–6 months of point-of-care reminders of their precommitment attached to a patient education handout, as well as weekly emails with links to resources to improve communication with patients about low-value services. We used electronic health record (EHR) data to measure the effects of the intervention, and conducted a clinician survey and qualitative analysis of clinician interviews.

### Setting and participants

Eligible participants were primary care clinicians (physicians and advanced practice providers) of the

six study clinics (figure 1). These six clinics (three internal medicine and three family medicine) were chosen because they were among the largest of the 11 IHA primary care clinics and because at the start of the study the 53 clinicians in these six clinics did not practise at other IHA clinics. Within each clinic, clinicians were eligible to participate if, at the time of recruitment in their clinic, they anticipated practising in that clinic for the duration of the study.

### Randomisation and intervention

All six clinics started the study in a control period (figure 2), at the beginning of which all IHA primary care clinicians were informed by email of *Choosing Wisely* recommendations against imaging for uncomplicated low back pain,<sup>33</sup> imaging for uncomplicated headaches<sup>34</sup> and unnecessary antibiotics for acute sinusitis.<sup>35</sup> This served to increase the likelihood that clinicians were aware of the recommendations throughout the control and intervention periods, rather than just at the time of study recruitment. These three recommendations were chosen because the targeted conditions are common in primary care practice and represent situations in which clinicians' intentions to avoid ordering low-value services can potentially be swayed by patient requests for services.<sup>12 13</sup>

After 2 months of the control period, recruitment and the intervention started in a new clinic each month in a random order determined using the user-written RALLOC module<sup>36</sup> in Stata V.14. Clinician recruitment consisted of a 20 min in-person presentation by the principal investigator, followed by informed consent. Clinicians who provided informed consent were given the opportunity to precommit to follow the three *Choosing Wisely* recommendations in future patient encounters by signing a brief document (see online supplementary appendix figure 1). Because the recruitment and precommitment processes could influence practice patterns, the month of the recruitment and precommitment processes for each clinic was

considered a transition period (figure 2) and excluded from analyses.<sup>32</sup>

During the intervention period, medical assistants, whose role in usual care was to bring patients to exam rooms and elicit the reason for their visit, were trained to identify patients presenting with a symptom for which one of the three *Choosing Wisely* recommendations might apply (ie, patients with low back pain, headaches or nasal congestion). For such patients, medical assistants placed for the clinician a Post-it reminder of their precommitment on a patient education handout (see online supplementary appendix figures 2–4) on the way into the exam room. Training of medical assistants occurred at the beginning of the intervention period and consisted of a 10 min individual training and provision of a one-page summary of their role (see online supplementary appendix figure 5). We chose a paper-based reminder over embedding the reminder in the clinics' EHR to ensure clinicians viewed the reminder immediately before the encounter, avoid the challenges of EHR alert fatigue<sup>37</sup> and ensure that the intervention could in the future be implemented in a range of clinical environments irrespective of their resources and EHR adoption.<sup>38</sup> Clinicians who precommitted also received a weekly email with links to *Choosing Wisely* resources to improve communication with patients about low-value services (see online supplementary appendix figure 6).

Depending on the clinic's randomly assigned order, the intervention period lasted for 1–6 months. These varying lengths were included because of the stepped wedge study design and because even a relatively short duration of decision support could be sufficient to change clinicians' practice patterns.<sup>39</sup> After the end of the intervention period, each clinic was followed for an additional 3 months to evaluate the durability of any intervention effects.

### Outcomes and analyses of ordering data

All study outcomes and analytic methods were prespecified. The primary outcome was the difference between

	Study Month												
	1	2	3	4	5	6	7	8	9	10	11	12	
Clinic 1													
Clinic 2													
Clinic 3													
Clinic 4													
Clinic 5													
Clinic 6													

Key	Control Month	Transition Month*	Intervention Month	Follow-Up Month

**Figure 2** Study design. \*Data for the months of the recruitment and precommitment processes (ie, the 'transition months') in each clinic were excluded from analyses. The vertical line in the boxes for these months indicates the approximate time of the recruitment and precommitment processes in each respective clinic.

the control and intervention periods in the percentage of visits with an applicable potentially low-value order: a lumbar spine X-ray, CT or MRI order in visits for low back pain<sup>33</sup>; a head CT or MRI order in visits for headaches<sup>34</sup>; and an antibiotic order in visits for acute sinusitis.<sup>35</sup> The main secondary outcome was the difference in these percentages between the control and follow-up periods to assess whether any intervention effect was sustained in the short term following the intervention.

During intervention pretesting and study implementation, clinicians noted that efforts to reduce low-value services could lead to substitution with referrals to specialists who might order the targeted low-value service, or alternate services of questionable value. Consequently, another secondary outcome was the difference between the control and intervention periods in the percentage of visits with a potential alternate order, which we defined as an order for an opiate,<sup>40</sup> or an orthopaedics, neurosurgery, spine clinic, neurology or pain clinic referral in visits for low back pain; an order for an opiate or butalbital,<sup>41</sup> or a neurology or pain clinic referral in visits for headaches; and an order for a sinus X-ray,<sup>42</sup> sinus CT,<sup>43</sup> or ear, nose and throat referral in visits for acute sinusitis.

We identified applicable visits in EHR data using the International Classification of Diseases, Ninth Revision (ICD-9) codes developed by Premera Blue Cross and the Washington State Choosing Wisely Task Force to measure rates of potentially low-value services for low back pain, headaches and acute sinusitis in administrative data.<sup>44 45</sup> The lists of ICD-9 codes are shown in online supplementary appendix table 1.

For each outcome we estimated differences in percentages using linear mixed models with random effects for providers nested in practices, adjusted for patient age, patient gender, patient Charlson comorbidity score at the visit,<sup>46</sup> month and condition. We fitted linear mixed models for dichotomous outcome variables because generalised linear models did not converge.<sup>47</sup> The choice of model terms and random effects was aided by reviewing likelihood ratio tests and Akaike's and Schwarz's Bayesian information criteria among competing models. We used robust SEs clustered at the clinic level. All hypothesis tests used a two-sided  $\alpha$  of 0.05 as the threshold for statistical significance. Data were analysed in Stata V.14.

In our power calculation, we used IHA EHR data to estimate each clinician would have 29 applicable visits per month, approximately 20% of these visits would have an order for a potentially low-value service, and an intraclass correlation coefficient of 0.034. Based on these data, a sample size of eight clinicians per clinic, six clinics, and a two-sided  $\alpha$  of 0.05 would provide 80% power to detect a difference in percentages of at least 3.7% between the control and intervention periods.

### Postintervention surveys and interviews of clinicians

During the 3-month follow-up period, we asked study clinicians to complete a web-based survey about their demographic characteristics, patient panel characteristics and perceptions of the intervention components (see online supplementary appendix figure 7). Each survey participant received a \$25 gift card.

During the follow-up period we also conducted semi-structured telephone interviews with a random sample of study clinicians, stratified by clinic, to explore how the intervention affected their conversations with patients and clinical decision making (see online supplementary appendix figure 8). Each interview participant received a \$25 gift card. We conducted 24 interviews, with at least two interviews per clinic, stopping when there were no novel responses to interview questions. Interviews were audio-recorded and transcribed for subjects who consented to recording; for subjects who did not consent to recording, detailed notes were taken. Two members of the study team independently coded all transcripts and notes and identified themes using template analysis.<sup>48</sup>

### Sensitivity analyses

We conducted three sensitivity analyses. The first was to gauge the robustness of the primary outcome results to the exclusion of one clinician from clinic 1 who started to also practise in clinic 4 before the intervention started in clinic 4. The second was to gauge the robustness of the primary outcome results to alternative ICD-9 codes to identify applicable visits that could be more sensitive for identifying use of low-value services<sup>3</sup> and capture any shifts in coding between the control and intervention periods. These more sensitive codes included the ICD-9 codes used for the main analyses and additional codes used in previous research to identify each condition in administrative data<sup>49–60</sup> (see online supplementary appendix table 2). The third was to examine the robustness of our results by also using cluster-bootstrapped and conventional SEs.

## RESULTS

Fifty-three clinicians were eligible for the study, and 45 (85%) consented to participate (figure 1). All 45 participating clinicians (100%) chose to precommit to follow the three *Choosing Wisely* recommendations. Clinician characteristics are shown in table 1.

### Orders for potentially low-value services

During the control period, 10.0% of 10 420 visits had an order for a potentially low-value service (table 2); trends by clinic are shown in online supplementary appendix figure 9. During the 7593 intervention period visits, there was no statistically significant decrease in the overall percentage of visits with an order for a potentially low-value service (−1.4%, 95% CI −2.9% to 0.1%;  $p=0.06$ ). However, for low



**Table 1** Characteristics of study clinicians, clinics and patients

	Clinic 1	Clinic 2	Clinic 3	Clinic 4	Clinic 5	Clinic 6	Total
<b>Study clinicians, n*</b>	<b>8</b>	<b>8</b>	<b>9</b>	<b>7</b>	<b>6</b>	<b>8</b>	<b>45</b>
Age in years, mean (SD)	48.6 (7.8)	50.0 (11.3)	48.7 (11.4)	53.3 (1.1)	46.2 (7.8)	46.0 (10.0)	48.5 (9.7)
Female, n (%)	3 (42.9)	4 (50.0)	6 (66.7)	4 (57.1)	4 (66.7)	4 (50.0)	25 (56.8)
Clinician type, n (%)							
Physician (MD or DO)	6 (85.7)	7 (87.5)	8 (88.9)	6 (87.5)	6 (100.0)	7 (87.5)	40 (90.9)
Physician assistant	0 (0.0)	1 (12.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.3)
Nurse practitioner	1 (14.3)	0 (0.0)	1 (11.1)	1 (14.3)	0 (0.0)	1 (12.5)	3 (6.8)
Years since training, n (%)							
Less than 5 years	0 (0.0)	1 (12.5)	0 (0.0)	0 (0.0)	1 (16.7)	2 (25.0)	4 (9.1)
5–9 years	0 (0.0)	3 (37.5)	2 (22.2)	1 (14.3)	0 (0.0)	1 (12.5)	7 (15.9)
10–19 years	4 (57.1)	2 (25.0)	4 (44.4)	2 (28.6)	4 (66.7)	2 (25.0)	18 (40.9)
20 or more years	3 (42.9)	2 (25.0)	3 (33.3)	4 (57.1)	1 (16.7)	3 (37.5)	15 (34.1)
Payer mix, n (%)†							
Private insurance	7 (87.5)	8 (100.0)	7 (77.8)	6 (85.7)	6 (100.0)	8 (100.0)	41 (95.3)
Medicaid	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (16.7)	1 (12.5)	2 (4.9)
Medicare	2 (28.6)	3 (37.5)	5 (62.5)	0 (0.0)	2 (33.3)	2 (25.0)	14 (34.1)
Visits, n‡							
Low back pain	3466	1758	3769	2207	1978	3014	16 192
Headaches	689	336	610	596	493	802	3526
Acute sinusitis	364	331	581	365	617	1920	4178
Overall	4519	2425	4960	3168	3088	5736	23 896
Patients§							
Mean age (SD)	50.1 (14.9)	51.4 (17.3)	55.9 (16.5)	52.2 (16.1)	47.7 (15.6)	49.9 (16.5)	51.4 (16.3)
Female, n (%)¶	2862 (63.3)	1436 (59.2)	3245 (65.4)	2456 (77.5)	2108 (68.3)	3640 (63.5)	15 747 (65.9)
Charlson score, mean (SD)**	0.13 (0.40)	0.15 (0.47)	0.22 (0.52)	0.18 (0.48)	0.16 (0.45)	0.15 (0.42)	0.17 (0.46)

\*Number of clinicians who consented to study participation. Column counts sum to greater than the total counts for clinician characteristics because after the intervention was implemented in clinic 1, one clinician from clinic 1 started to also practise in clinic 4 due to operational needs. All of the clinic 4 visits for this clinician were conservatively treated as being in the control period 4 until the intervention was implemented in clinic 4. Data on clinicians are from the postintervention survey (completed by 44 of the 45 study clinicians).

†Clinicians' self-report that more than 30% of their patients had each type of health insurance coverage. One survey participant did not complete this item.

‡The number of applicable visits across the control, intervention and follow-up periods using lists of the International Classification of Diseases, Ninth Revision (ICD-9) codes previously developed by Premera Blue Cross and the Washington State Choosing Wisely Task Force to measure rates of potentially low-value services in visits for which each respective *Choosing Wisely* recommendation could apply. All lists of ICD-9 codes are shown in online supplementary appendix table 1.

§Characteristics of patients in visits with an ICD-9 code from the lists of ICD-9 codes previously developed by Premera Blue Cross and the Washington State Choosing Wisely Task Force to measure rates of potentially low-value services in visits for which each respective *Choosing Wisely* recommendation could apply.

¶Number and per cent of patients in each clinic who were female. Data on gender were missing for 35 patients.

\*\*Charlson scores based on ICD-9 codes used at each study visit. The weighted Charlson sum ranges from 0 to 8. The index is 0 for observations whose sum is 0, 1 for observations whose sum is 1, and 2 for observations whose sum is between 2 and 8.

back pain there was a statistically significant decrease in the percentage of visits with an order for a potentially low-value service (−1.2%, 95% CI −2.0% to −0.5%;  $p=0.001$ ).

In sensitivity analyses, when we excluded the one clinician from clinic 1 who started to also practise in clinic 4 before the intervention started in clinic 4, we found nearly identical results (data not shown). When we used the more sensitive sets of ICD-9 codes to identify applicable visits, the sample size for the control and intervention periods combined increased from 18 013 to 95 477, and the changes in the percentage of visits for low back pain with an order for a potentially low-value service were similar (see online supplementary appendix

table 3). Additionally, there were statistically significant decreases in the percentages of visits with an order for a potentially low-value service overall (−1.8%, 95% CI −2.9% to −0.7%;  $p=0.001$ ), for headaches (−0.7%, 95% CI −1.3% to −0.2;  $p=0.006$ ) and for acute sinusitis (−3.2%, 95% CI −5.1% to −1.3%;  $p=0.001$ ). We found similar results using cluster-bootstrapped SEs (see online supplementary appendix table 4) and a statistically significant decrease in the overall percentage of visits with an order for a potentially low-value service (95% CI −2.7% to −0.2%;  $p=0.02$ ) when using conventional SEs.

During the 5883 follow-up period visits, there was no statistically significant decrease in the overall or

**Table 2** Changes in percentages of visits with orders for potentially low-value services

Condition	Control period		Intervention period			Follow-up period		
	n*	%†	n*	Δ in % (95% CI)‡	p	n*	Δ in % (95% CI)‡	p
Low back pain	6669	4.0	5295	-1.2 (-2.0 to -0.5)	0.001	4228	-0.3 (-1.3 to 0.8)	0.62
Headaches	1661	0.7	1074	0.7 (-0.7 to 2.1)	0.34	791	0.3 (-0.6 to 1.2)	0.52
Acute sinusitis	2090	36.6	1224	-3.4 (-8.2 to 1.4)	0.16	864	-2.7 (-6.6 to 1.3)	0.19
Any of the three conditions§	10 420	10.0	7593	-1.4 (-2.9 to 0.1)	0.06	5883	-0.7 (-2.0 to 0.7)	0.33

\*n=number of visits with an International Classification of Diseases, Ninth Revision (ICD-9) code for each respective condition during each study period. The lists of ICD-9 codes were previously developed by Premera Blue Cross and the Washington State Choosing Wisely Task Force to measure rates of potentially low-value services in visits for which each respective *Choosing Wisely* recommendation could apply. All lists of ICD-9 codes are shown in online supplementary appendix table 1.

†The raw percentage of visits with an order for a potentially low-value service: antibiotics for acute sinusitis, lumbar spine X-ray, CT or MRI for low back pain, and head CT or MRI for headaches.

‡Mean absolute differences between the control period and the intervention or follow-up period, respectively, in the percentages of visits with orders for a potentially low-value service (defined above) as estimated from linear mixed models with random effects for providers nested in practices, adjusted for patient characteristics, time and comorbidities. SEs are robust SEs clustered at the clinic level.

§Composite measure in which the denominator is all visits for one of the three conditions, and the numerator is an order for a potentially low-value service for that respective condition.

condition-specific percentage of visits with orders for potentially low-value services relative to the control period (table 2).

### Orders for potential alternate services

During the intervention period there was a 1.7% increase in the overall percentage of visits with orders for potential alternate services relative to the control period (95% CI 0.3% to 3.1%; p=0.01) (table 3). Low back pain was the only individual condition for which there was a statistically significant increase in the percentage of visits with a potential alternate order (1.9%, 95% CI 0.5% to 3.3%; p=0.007).

### Postintervention surveys and interviews

Twenty-one of the 44 clinicians (47.7%) who completed the survey found the precommitment itself helpful, 14 (31.8%) found the precommitment reminders helpful, 28 (63.6%) found the patient education handouts

helpful, and 9 (20.4%) found the online resources to improve communication with patients helpful. Of the 24 clinicians we interviewed, 14 felt the intervention improved their conversations with patients about low-value services. One clinician, for example, said the intervention “provided extra support in my decision [to avoid ordering a low-value service] and something else for [patients] to read and understand.” Another said, “I was able to use those handouts and point to specific parts...saying this is what we believe to be evidence-based and this is the right way...to practice.” Ten clinicians felt the intervention positively changed their overall clinical decision making. For example, one clinician said, “the more difficult patients where I would have been tempted to not follow guidelines...I would more consistently follow guidelines in those tougher cases and spend the extra time.” Another said, “I think I probably [now] lean much further in the direction of not treating too early.” Additional

**Table 3** Changes in percentages of visits with potential alternate orders

Condition	Control period		Intervention period			Follow-up period		
	n*	%†	n*	Δ in % (95% CI)‡	p	n*	Δ in % (95% CI)‡	p
Low back pain	6669	13.2	5295	1.9 (0.5 to 3.3)	0.007	4228	1.5 (-0.5 to 3.6)	0.14
Headaches	1661	9.5	1074	4.2 (-1.0 to 9.3)	0.11	791	3.4 (-1.7 to 8.6)	0.19
Acute sinusitis	2090	0.7	1224	-0.2 (-0.8 to 0.4)	0.44	864	-0.5 (-1.3 to 0.2)	0.14
Any of the three conditions§	10 420	10.1	7593	1.7 (0.3 to 3.1)	0.01	5883	1.4 (-0.4 to 3.1)	0.12

\*n=number of visits with an International Classification of Diseases, Ninth Revision (ICD-9) code for each respective condition during each study period. The lists of ICD-9 codes were previously developed by Premera Blue Cross and the Washington State Choosing Wisely Task Force to measure rates of potentially low-value services in visits for which each respective *Choosing Wisely* recommendation could apply. All lists of ICD-9 codes are shown in online supplementary appendix table 1.

†The raw percentage of visits with a potential alternate order of questionable value: sinus X-ray, sinus CT or ear, nose and throat referrals for acute sinusitis; opiate prescriptions, or orthopaedics, neurosurgery, spine clinic, neurology or pain clinic referrals for low back pain; or opiate or butalbital prescriptions, or neurology or pain clinic referrals for headaches.

‡Mean absolute differences between the control period and the intervention or follow-up period, respectively, in the percentages of visits with potential alternate orders of questionable value (defined above) as estimated from linear mixed models with random effects for providers nested in practices, adjusted for patient characteristics, time and comorbidities. SEs are robust SEs clustered at the clinic level.

§Composite measure in which the denominator is all visits for one of the three conditions, and the numerator is a potential alternate order of questionable value for that respective condition.

representative quotes are shown in online supplementary appendix table 5.

## DISCUSSION

We found that a behavioural economic strategy of asking clinicians to precommit to avoid ordering specific low-value services, followed by supports to promote adherence to this precommitment, was only associated with a statistically significant but small decrease in orders for potentially low-value services in visits for one of three targeted conditions: low back pain. Further, these intervention effects were not sustained in the 3 months after the intervention ended, and may have led to alternate orders of questionable value.

Our study's strengths include the translation of multiple insights from behavioural economics into a novel and highly scalable intervention that targeted multiple low-value health services, the randomised design, the conduct of our trial in a community (rather than academic) practice, the measurement of unintended consequences, and the use of mixed methods to better understand the impact of the intervention on clinicians' decision making. Weaknesses of our study include our inability to determine whether appropriateness of ordering changed as a result of the intervention because our data lacked detailed clinical information. Yet this was consistent with the main goal of our study, which was to first determine whether the intervention could change clinicians' ordering behaviours. Future studies should supplement ordering data with detailed clinical data to examine how the intervention affects appropriateness. Our study was of a relatively short duration and stepped wedge trials are not well-suited to estimate long-term intervention effects. Thus it will be important in future research to measure longer term effects using other study designs. We tested a multicomponent intervention<sup>9</sup> and our study was not designed to disentangle the effects of each of the precommitment intervention components, although our survey and interview results provide insight into clinicians' perceptions of these components. As with any surveys or interviews, our findings about clinicians' views of the intervention could have been subject to recall or social desirability biases. Our study did not examine patient experiences or economic impact, and it will be important in future work to examine how this intervention affects these outcomes. Although we conducted our study in a community group practice to enhance the external validity of our findings, our results may not generalise to all clinical settings.

Other studies have tested interventions to improve clinicians' decisions about low-value care<sup>9</sup> and guideline adherence,<sup>61</sup> yet we are aware of only two trials that have tested behavioural economic strategies to discourage ordering of low-value services. The first trial tested poster-sized letters from clinicians that were placed in clinic examination rooms for 12 weeks. These

posters, which stated clinicians' commitment to avoid inappropriate antibiotic prescribing for acute respiratory infections (ARIs) and included their photographs and signatures, led to a 19.7% absolute reduction in inappropriate antibiotic prescriptions for ARIs.<sup>62</sup> The second trial found that 18 months of 'accountable justification' (prompting clinicians to enter free-text justifications for prescribing antibiotics into patients' EHRs) and 'peer comparison' (emails to clinicians that compared their own antibiotic prescribing rates with peers with the lowest inappropriate prescribing rates) led to 7.0% and 5.2% absolute reductions, respectively, in inappropriate antibiotic prescriptions.<sup>63</sup> In contrast to these studies, ours tested a multicomponent intervention that aimed to reduce potentially low-value orders for multiple conditions, including an acute condition (ie, acute sinusitis) and conditions that are often chronic (ie, low back pain and headaches). For these latter two conditions the orders we targeted were for diagnostic imaging, which per order is typically more costly than an antibiotic. Further, our study examined the effects of the intervention on ordering of potentially low-value services and on clinicians' experiences and intended consequences.

Our results have important implications for practitioners and policymakers. First, while our intervention was associated with a decrease in the percentage of visits for low back pain with an order for a potentially low-value service, the magnitude of this change was small (just 12 fewer potentially low-value orders per 1000 visits) and not sustained in the near term after the intervention ended. Nonetheless, our intervention required minimal resources and minor adjustments to clinical processes, and the results for low back pain are comparable in magnitude to the effects of the Medicare Pioneer Accountable Care Organization programme on use of low-value care.<sup>64</sup> It is possible that sustained changes in ordering could be achieved by continuously integrating the intervention into routine care, for example as an EHR clinical decision support algorithm in higher resource settings or as a paper-based approach in lower resource settings. Second, for headaches and acute sinusitis, we only found a decrease in ordering potentially low-value services in sensitivity analyses using more sensitive ICD-9 codes. For these two conditions the increases in sample size in the sensitivity analyses were particularly large, which could have enhanced our ability to detect a statistically significant effect for these conditions. Alternatively, because the more specific codes for these two conditions were used less commonly in the intervention period than the control period (see online supplementary appendix table 6), shifts in coding or seasonal trends could have contributed to these sensitivity analysis findings. Another possibility is that the more sensitive ICD-9 codes for headaches and acute sinusitis could have less specificity for identifying use of low-value services,<sup>3</sup> and thus when used

in sensitivity analyses may have detected decreases in appropriate orders. Future studies should resolve these differences between the main and sensitivity analyses by examining how this intervention changes the appropriateness of care, which could be accomplished using measures that combine ordering and clinical data.<sup>4</sup> Third, our survey and interview findings suggest our results were driven at least in part by changes in clinician–patient interactions. At its core, the *Choosing Wisely* campaign encourages clinicians and patients to question services that may not yield benefits and could lead to harms,<sup>65</sup> and an effective way to facilitate such conversations could be provision of point-of-care handouts by clinicians who had precommitted to avoiding low-value care. Fourth, the intervention may have led to more referrals to other clinicians who could order the targeted services or other services of questionable value, particularly in visits for low back pain. Because the urge to ‘do something’ can be strong in clinical practice,<sup>66</sup> our results suggest it will be crucial to monitor for and seek to ameliorate such unintended consequences in efforts to reduce low-value care.

In conclusion, a behavioural economic strategy of asking clinicians to precommit to specific *Choosing Wisely* recommendations paired with decision supports, although theoretically promising and highly scalable, only yielded a small and unsustainable decrease in potentially low-value orders for one of three targeted conditions. The intervention also likely led to a small increase in alternate orders for services of questionable value. These results highlight the potential of behavioural economic strategies to be integrated into clinical workflows with minimal resources as well as some of the important challenges to be confronted in interventions to reduce ordering of low-value services. As momentum to reduce delivery of low-value care continues to grow worldwide,<sup>4 66</sup> more such research is needed to identify low-cost, scalable strategies to achieve this goal while minimising the potential for unintended consequences.

#### Author affiliations

<sup>1</sup>Center for Clinical Management Research, VA Ann Arbor Healthcare System, Ann Arbor, Michigan, USA

<sup>2</sup>Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA

<sup>3</sup>Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, USA

<sup>4</sup>Center for Bioethics and Social Sciences in Medicine, University of Michigan Medical School, Ann Arbor, MI, USA

<sup>5</sup>School of Information, University of Michigan, Ann Arbor, MI, USA

<sup>6</sup>School of Public Health, University of Michigan, Ann Arbor, MI, USA

<sup>7</sup>Linden Consulting Group, Ann Arbor, Michigan, USA

<sup>8</sup>Division of Family Medicine, IHA, Ann Arbor, Michigan, USA

<sup>9</sup>Division of Quality and Performance Improvement, IHA, Ann Arbor, MI, USA

**Acknowledgements** We thank A Jay Holmgren, Eric Pfeifer and Dori Cross for assistance with coding the interview data and Pavel Vinarov for assistance with assembling the administrative data.

**Funding** This work was funded by the Robert Wood Johnson Foundation (grant number 71475). Support was also provided by the US Department of Veterans Affairs (VA). Dr Kullgren was supported by a VA Health Services Research & Development Career Development Award (grant number 13-267). The funding sources had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; or preparation, review or approval of the manuscript.

**Competing interests** JTK has received research grants from the US Department of Veterans Affairs Health Services Research & Development (HSR&D) Service, Robert Wood Johnson Foundation, Donaghue Foundation, and Center for Medicare & Medicaid Services (CMS); received consulting fees from SeeChange Health and HealthMine; and received speaking honoraria from AbilTo. EK has received research grants from the Robert Wood Johnson Foundation, Donaghue Foundation and National Science Foundation. JM has received research grants from the Donaghue Foundation, Agency for Healthcare Research and Quality, Commonwealth Fund, CMS, John A Hartford Foundation, California Health Care Foundation, DHHS – Office of the National Coordinator for Health IT, and Blue Cross Blue Shield of Michigan Foundation. JM is on the advisory board of QPID Health.

**Ethics approval** The protocol was approved by the institutional review board of the University of Michigan Medical School (HUM00087820).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** Patient-level data, statistical code and the trial protocol are available from the corresponding author at jkullgre@med.umich.edu. Consent for data sharing was not obtained from participants but the presented data are anonymised and risk of identification is low.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

#### REFERENCES

- 1 Blumenthal-Barby JS. "Choosing wisely" to reduce low-value care: a conceptual and ethical analysis. *J Med Philos* 2013;38:559–80.
- 2 Colla CH, Morden NE, Sequist TD, *et al*. Choosing wisely: prevalence and correlates of low-value health care services in the united states. *J Gen Intern Med* 2015;30:221–8.
- 3 Schwartz AL, Landon BE, Elshaug AG, *et al*. Measuring low-value care in medicare. *JAMA Intern Med* 2014;174:1067–76.
- 4 Bhatia RS, Levinson W, Shortt S, *et al*. Measuring the effect of choosing wisely: an integrated framework to assess campaign impact on low-value care. *BMJ Qual Saf* 2015;24:523–31.
- 5 Cabana MD, Rand CS, Powe NR, *et al*. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;282:1458–65.
- 6 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;362:1225–30.
- 7 Armstrong D. Clinical autonomy, individual and collective: the problem of changing doctors' behaviour. *Soc Sci Med* 2002;55:1771–7.
- 8 Colla CH. Swimming against the current-what might work to reduce low-value care? *N Engl J Med* 2014;371:1280–3.
- 9 Colla CH, Mainor AJ, Hargreaves C, *et al*. Interventions Aimed at Reducing Use of Low-Value Health Services: A Systematic Review. *Med Care Res Rev* 2016.
- 10 Shaw MK, Davis SA, Fleischer AB, *et al*. The duration of office visits in the United States, 1993 to 2010. *Am J Manag Care*



- 2014;20:820-6 <http://www.ajmc.com/journals/issue/2014/2014-vol20-n10/the-duration-of-office-visits-in-the-united-states-1993-to-2010>.
- 11 Baron RJ. What's keeping us so busy in primary care? a snapshot from one practice. *N Engl J Med* 2010;362:1632-6.
  - 12 Kravitz RL, Bell RA, Azari R, *et al*. Request fulfillment in office practice: antecedents and relationship to outcomes. *Med Care* 2002;40:38-51.
  - 13 Marple RL, Kroenke K, Lucey CR, *et al*. Concerns and expectations in patients presenting with physical complaints. Frequency, physician perceptions and actions, and 2-week outcome. *Arch Intern Med* 1997;157:1482-8.
  - 14 Kahneman D. *Thinking, Fast and Slow*. 1st edn. New York: Farrar, Straus and Giroux, 2013.
  - 15 Stanovich KE, West RF. Individual differences in reasoning: implications for the rationality debate? *Behav Brain Sci* 2000;23:645-65.
  - 16 Elqayam S, Over D. Probabilities, beliefs, and dual processing: the paradigm shift in the psychology of reasoning. *Mind & Society* 2012;11:27-40.
  - 17 Cohen DA, Babey SH. Contextual influences on eating behaviours: heuristic processing and dietary choices. *Obes Rev* 2012;13:766-79.
  - 18 Ariely D, Wertenbroch K. Procrastination, deadlines, and performance: self-control by precommitment. *Psychol Sci* 2002;13:219-24.
  - 19 Loewenstein G. *Out of control: visceral influences on behavior. in: advances in behavioral economics*. Princeton, NJ: Princeton University Press, 2004:689-723.
  - 20 Thaler R. Toward a positive theory of consumer choice. *J Econ Behav Organ* 1980;1:39-60.
  - 21 Rogers T, Milkman KL, Volpp KG. Commitment devices: using initiatives to change behavior. *JAMA* 2014;311:2065-6.
  - 22 Halpern SD, Asch DA, Volpp KG. Commitment contracts as a way to health. *BMJ* 2012;344:e522.
  - 23 Thaler RH, Benartzi S. Save more Tomorrow™: using behavioral economics to increase employee saving. *J Polit Econ* 2004;112:S164-S187.
  - 24 Giné X, Karlan D, Zinman J. Put your money where your butt is: a commitment contract for smoking cessation. *Am Econ J Appl Econ* 2010;2:213-35.
  - 25 Halpern SD, French B, Small DS, *et al*. Randomized trial of four financial-incentive programs for smoking cessation. *N Engl J Med* 2015;372:2108-17.
  - 26 Volpp KG, John LK, Troxel AB, *et al*. Financial incentive-based approaches for weight loss: a randomized trial. *JAMA* 2008;300:2631-7.
  - 27 Kullgren JT, Troxel AB, Loewenstein G, *et al*. A randomized controlled trial of employer matching of employees' monetary contributions to deposit contracts to promote weight loss. *Am J Health Promot* 2016;30:441-52.
  - 28 Cassel CK, Guest JA. Choosing wisely: helping physicians and patients make smart decisions about their care. *JAMA* 2012;307:1801-2.
  - 29 Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000;55:68-78.
  - 30 Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol* 2006;6:54.
  - 31 Hemming K, Girling A, Martin J, *et al*. Stepped wedge cluster randomized trials are efficient and provide a method of evaluation without which some interventions would not be evaluated. *J Clin Epidemiol* 2013;66:1058-9.
  - 32 Hemming K, Haines TP, Chilton PJ, *et al*. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* 2015;350:h391.
  - 33 ACP - Avoid imaging for low back pain | Choosing Wisely. <http://www.choosingwisely.org/clinician-lists/american-college-physicians-imaging-for-non-specific-low-back-pain/> (accessed 30 Jun 2016).
  - 34 ACR - Avoid imaging for uncomplicated headache | Choosing Wisely. <http://www.choosingwisely.org/clinician-lists/american-college-radiology-imaging-for-uncomplicated-headache/> (accessed 30 Jun 2016).
  - 35 AAFP - Antibiotics for acute mild-to-moderate sinusitis | Choosing Wisely. <http://www.choosingwisely.org/clinician-lists/american-academy-family-physicians-antibiotics-for-sinusitis/> (accessed 30 Jun 2016).
  - 36 Ryan P. RALLOC:Stata module to design randomized controlled trials. 2011. <http://econpapers.repec.org/software/bocbocode/s319901.htm> (accessed 5 Aug 2017).
  - 37 Shah NR, Seger AC, Seger DL, *et al*. Improving acceptance of computerized prescribing alerts in ambulatory care. *J Am Med Inform Assoc* 2006;13:5-11.
  - 38 Jamoom E, Hing E. Progress with electronic health record adoption among emergency and outpatient departments: United States, 2006-2011. *NCHS Data Brief* 2015:1-8.
  - 39 Chauhan BF, Jeyaraman M, Mann AS, *et al*. Behavior change interventions and policies influencing primary healthcare professionals' practice—an overview of reviews. *Implement Sci IS* 2017:12.
  - 40 AAPMR - Opiates for low back pain | Choosing Wisely. <http://www.choosingwisely.org/clinician-lists/aapmr-opiates-for-low-back-pain/> (accessed 20 Jul 2016).
  - 41 American Headache Soc - Treating recurrent headaches | Choosing Wisely. <http://www.choosingwisely.org/clinician-lists/american-headache-society-opioid-butalbital-containing-meds-for-recurrent-headaches/> (accessed 20 Jul 2016).
  - 42 Aring AM, Chan MM. Current concepts in adult acute rhinosinusitis. *Am Fam Physician* 2016;94:97-105.
  - 43 AAAAI - For uncomplicated acute rhinosinusitis | Choosing Wisely. <http://www.choosingwisely.org/clinician-lists/american-academy-allergy-asthma-immunology-uncomplicated-acute-rhinosinusitis/> (accessed 20 Jul 2016).
  - 44 Less waste. Less harm. *Choosing Wisely in Washington state*. Seattle, WA: Washington Health Alliance, 2014. <http://wahealthalliance.org/wp-content/uploads.php?link-year=2014&link-month=09&link=Choosing-Wisely-in-Washington-state.pdf>. (accessed 18 Jul 2016).
  - 45 Choosing Wisely Claims-Based Technical Specifications. Seattle, WA: Washington State Choosing Wisely Task Force, 2014. [http://wahealthalliance.org/wp-content/uploads.php?link-year=2013&link-month=11&link=Choosing\\_Wisely\\_Specifications\\_2014.pdf](http://wahealthalliance.org/wp-content/uploads.php?link-year=2013&link-month=11&link=Choosing_Wisely_Specifications_2014.pdf).
  - 46 Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45:613-9.
  - 47 Leontjevas R, Gerritsen DL, Smalbrugge M, *et al*. A structural multidisciplinary approach to depression management in nursing-home residents: a multicentre, stepped-wedge cluster-randomised trial. *Lancet* 2013;381:2255-64.
  - 48 Brooks J, McCluskey S, Turley E, *et al*. The utility of template analysis in qualitative psychology research. *Qual Res Psychol* 2015;12:202-22.
  - 49 Cherkin DC, Deyo RA, Volinn E, *et al*. Use of the international classification of diseases (ICD-9-CM) to

- identify hospitalizations for mechanical low back problems in administrative databases. *Spine* 1992;17:817–25.
- 50 Licciardone JC. The epidemiology and medical management of low back pain during ambulatory medical care visits in the United States. *Osteopath Med Prim Care* 2008;2:11.
- 51 Mafi JN, McCarthy EP, Davis RB, *et al.* Worsening trends in the management and treatment of back pain. *JAMA Intern Med* 2013;173:1573–81.
- 52 National Quality Measures Clearinghouse. Use of imaging studies for low back pain: percentage of members with a primary diagnosis of low back pain who did not have an imaging study (plain x-ray, MRI, CT scan) within 28 days of the diagnosis. <http://www.qualitymeasures.ahrq.gov/content.aspx?id=48635> (accessed 13 Aug 2015).
- 53 Callaghan BC, Kerber KA, Pace RJ, *et al.* Headache neuroimaging: routine testing when guidelines recommend against them. *Cephalalgia* 2015;35:1144–52.
- 54 Mafi JN, Edwards ST, Pedersen NP, *et al.* Trends in the ambulatory management of headache: analysis of NAMCS and NHAMCS data 1999–2010. *J Gen Intern Med* 2015;8.
- 55 Smitherman TA, Burch R, Sheikh H, *et al.* The prevalence, impact, and treatment of migraine and severe headaches in the United States: a review of statistics from national surveillance studies. *Headache* 2013;53:427–36.
- 56 Bhattacharyya N, Grebner J, Martinson NG. Recurrent acute rhinosinusitis: epidemiology and health care cost burden. *Otolaryngol Head Neck Surg* 2012;146:307–12.
- 57 Cadieux G, Tamblyn R. Accuracy of physician billing claims for identifying acute respiratory infections in primary care. *Health Serv Res* 2008;43:2223–38.
- 58 Fairlie T, Shapiro DJ, Hersh AL, *et al.* National trends in visit rates and antibiotic prescribing for adults with acute sinusitis. *Arch Intern Med* 2012;172:1513–4.
- 59 Gill JM, Fleischut P, Haas S, *et al.* Use of antibiotics for adult upper respiratory infections in outpatient settings: a national ambulatory network study. *Fam Med* 2006;38:349.
- 60 Sharp HJ, Denman D, Puumala S, *et al.* Treatment of acute and chronic rhinosinusitis in the united states, 1999–2002. *Arch Otolaryngol Head Neck Surg* 2007;133:260–5.
- 61 Gagliardi AR, Alhabib S. members of Guidelines International Network Implementation Working Group. Trends in guideline implementation: a scoping systematic review. *Implement Sci* 2015;10:10.
- 62 Meeker D, Knight TK, Friedberg MW, *et al.* Nudging guideline-concordant antibiotic prescribing: a randomized clinical trial. *JAMA Intern Med* 2014;174:425–31.
- 63 Meeker D, Linder JA, Fox CR, *et al.* Effect of behavioral interventions on inappropriate antibiotic prescribing among primary care practices: a randomized clinical trial. *JAMA* 2016;315:562–70.
- 64 Schwartz AL, Chernew ME, Landon BE, *et al.* Changes in low-value services in year 1 of the medicare pioneer accountable care organization program. *JAMA Intern Med* 2015;175:1815.
- 65 Choosing Wisely. Things physicians and patients should question | Choosing Wisely. <http://www.choosingwisely.org/doctor-patient-lists/>. (accessed 12 Jul 2013).
- 66 Malhotra A, Maughan D, Ansell J, *et al.* Choosing wisely in the UK: the academy of medical royal colleges' initiative to reduce the harms of too much medicine. *BMJ* 2015;350:h2308.