Bond University

Research Repository

Effect of two behavioural 'nudging' interventions on management decisions for low back pain:
A randomised vignette-based study in general practitioners

Soon, Jason; Traeger, Adrian C.; Elshaug, Adam G.; Cvejic, Erin; Maher, Chris G.; Doust,
Jenny A.; Mathieson, Stephanie; McCaffery, Kirsten; Bonner, Carissa

Link to output in Bond University research repository.

Effect of two behavioural 'nudging' interventions on choice of management options for low back pain: A randomised vignette-based study in general practitioners

Abstract

Objective: 'Nudges' are subtle cognitive cues thought to influence behaviour. We investigated whether embedding nudges in a general practitioner (GP) clinical decision support display can reduce choice of low-value management options.

Methods: Australian GPs completed four clinical vignettes of patients with low back pain. Participants chose from three guideline-concordant and three guideline-discordant (low-value) management options for each vignette, on a computer screen. A 2x2 factorial design randomised participants to two possible nudge interventions: 'partition display' nudge (low-value options presented horizontally, high value options listed vertically), or 'default option' nudge (high-value options presented as the default, low-value options presented only after clicking for more). The primary outcome was the proportion of scenarios where practitioners chose at least one of the low-value care options.

Results: 120 general practitioners (72% male, 28% female) completed the trial (n=480 vignettes). Participants using a conventional menu display without nudges chose at least one low-value care option in 42% of scenarios. Participants exposed to the default option nudge were 44% less likely to choose at least one low-value care option (odds ratio [OR] 0.56, 95% CI 0·37 to 0.85; p=0·006) compared to those not exposed. The partition display nudge had no effect on choice of low-value care (OR 1.08, 95% CI 0·72 to 1.64; p=0.7). There was no interaction between the nudges (OR 0.94, 95% CI 0·41 to 2.15; p=0·89).

Interpretation: A default option nudge reduced the odds that a GP will choose low-value options for low back pain in clinical vignettes. Embedding high value options as defaults in

clinical decision support tools could improve quality of care. More research is needed into

how nudges impact clinical decision-making in different contexts.

INTRODUCTION

Low-value care is healthcare that provides little or no benefit, may cause patient harm, or yields marginal benefits at a disproportionately high cost.[1] This is increasingly recognised as a worldwide problem.[2 3]

Current diagnosis and management of low back pain provides compelling examples of low-value care. Recent research shows that clinical mismanagement of low back pain is an international problem, including in low, middle and high income countries.[4] Imaging of acute low back pain, for example, does not sensibly inform management,[5] and is associated with harms,[6] yet a survey of Australian general practitioners (GP) found low back pain imaging was ordered in 25% of consultations.[7] Trends in analgesic medicine prescribing for low back pain are worsening; one Australian study found that between 2004 and 2014, GPs were prescribing increasing amounts of guideline-discordant complex medicines including opioids and fewer guideline-concordant medicines (e.g. non-steroidal anti-inflammatory drugs).[8]

Traditional knowledge translation tools such as clinical education about current guidelines are one means of changing clinical behaviour to reduce low-value care but are estimated to be effective only 10-20% of the time.[9] For example, one trial found targeted clinical education and reminders reduced low back pain imaging[10] but this has never been replicated. More research is needed into alternative approaches.

One alternative approach is the use of 'nudges' which take advantage of the cognitive biases influencing clinicians' habitual behaviour to make it more difficult to choose low-value options while preserving their freedom to choose .[11] Partition display nudges assume that people are less likely to select items listed horizontally than those listed vertically, for example on a computer screen.[12] One recent randomised study found that embedding a partition display nudge (listing low-value medication options horizontally rather than vertically) in a simulated Electronic Health Record reduced the proportion of respondents choosing aggressive antibiotic medications by 11.2% (43.7% in control group vs. 32.5% in partition display nudge

group).[12] Default option nudges exploit a well-documented cognitive bias towards preferring the default choice option over those that require additional thought or action.[13] There is evidence from non-randomised and quasi-randomised experiments that positioning the preferred choices as the default options in Electronic Health Record systems can influence decision making about medicines [14-16] and tests.[17] [18]

While there has been growing interest in the use of nudges to reduce low-value care,[19-21] few randomised controlled trials have investigated the effectiveness of default option and partition display nudges in a clinical context. We are aware of only one (previously mentioned) randomised controlled trial[12] of partition display nudges and none involving default option nudges. The study of partition display nudges[12] found that the nudges could influence decisions about antibiotic medicines but did not examine effects on test ordering behaviour e.g. ordering diagnostic imaging. Thus, it remains unclear whether partition display or default option nudges can influence testing and treatment decisions in other common conditions such as low back pain.

We aimed to determine whether the use of partition display and default option nudges in a simulated clinical decision support display can reduce low-value clinical management choices related to imaging and complex medicines for low back pain.

METHODS

Study design and participants

This was a double blind (participant and assessor), 2x2 factorial, parallel group randomised trial. The trial was registered at the Australian New Zealand Clinical Trials Registry (ACTRN12618000769280; trial protocol and statistical analysis plan available in full at https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=375033&isReview=true). Ethics approval was also obtained from Royal Prince Alfred Hospital Ethics Review Committee (Approval number HREC/18/RPAH/138). No amendments to study protocol were made after the enrolment of the first participant.

Participants were qualified general practitioners aged 18 years and over located in Australia. They were recruited by a market research company and informed in the recruitment email that they would receive $40 to complete a 15-minute survey. The company relies on an extensive database of general practitioners who are required to opt-in to their database and therefore have already indicated a willingness to participate in online research. Recruitment began on 21 May 2018 and halted on 4 June 2018.

Procedures and intervention

Participants completed four clinical vignettes of patients with low back pain. The vignettes were developed by experienced clinicians and based on current guidance from international clinical guidelines.[22 23] The first two vignettes described a low back pain scenario paired with six imaging options, while the last two vignettes described a low back pain scenario with six medication options (see **Supplementary File 1** for complete vignettes).

For each vignette, participants had to select at least one, and up to three, of the six options provided. Each participant was randomised to one of four experimental conditions in the presentation of these options across all vignettes:

1. All options displayed conventionally, that is, fully visible and listed vertically (**No nudge**)
2. All options visible but three low-value options listed horizontally rather than vertically (**Partition display nudge**)
3. Three guideline-concordant options visible on the initial view, with message stating 'click to see more options' below third option. Upon one click, low-value options fully visible and listed vertically (**Default option nudge**)
4. Three guideline-concordant options visible on initial view, with message stating 'click to see more options' below third option. Upon one click, low-value options fully visible but listed horizontally (i.e. a Partition display nudge). (**Double nudge i.e. Default option nudge + Partition display nudge**).

Figure 1 shows the visual appearance of the partition display nudge and default option nudge on a computer screen. See **Supplementary File 2** for computer screenshots showing how the menu options were displayed to participants in each of the four experimental conditions.

The 2x2 factorial design underlying this experiment is summarised in **Supplementary File 3**.

Randomisation and blinding

The survey was hosted on the Qualtrics survey platform. Randomisation of each participant to one of our four experimental conditions was conducted by the Qualtrics Block Randomizer function set to 'Evenly present elements'. This meant that the randomiser assigned each new participant to those conditions with the lowest display counts using a random number generator. Participants were blinded to the study hypothesis, and outcome assessors were blinded to group allocation until study completion.

Primary and secondary outcome measures

The primary outcome measure was the proportion of scenarios where participants chose at least one of the low-value options.

The secondary outcome measure was the likelihood participants would recommend each of the six options using a Likert scale based on the following: '*How likely would you be to recommend each of the below options for this case? (1 = very unlikely and 5 = very likely)'.* This information was collected after participants had made their initial choice of management option/s. Likert score questions were displayed conventionally (i.e. all options were visible and listed vertically) without nudges.

Other measures and data collected

Information was also collected about participant characteristics, namely sex, practice management software used, years of clinical experience, location of practice and estimated percentage of working hours spent teaching or in academia. This demographic data was collected before participants were allocated to the intervention.

To explore acceptability of the menu formats we planned to examine the intervention on dropout rate; if drop-outs were disproportionately high in one intervention group we felt this could indicate poor acceptability of the menu format.

We also asked participants to complete a Cognitive Reflection Test at the end of the survey. The Cognitive Reflection Test is designed to measure a person's tendency to override their cognitive biases and engage in further reflection to find a correct answer.[24] It is based on three questions (for example, *'A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?'*) and is measured on a 0 to 3 scale (0 = 0 questions answered correctly, 3 = all 3 questions answered correctly). A low test score indicates that a person may be thinking 'fast' (i.e. more susceptible to cognitive biases). The Cognitive Reflection Test has been validated as a measure of 'reflective thinking.' [24] [25] We defined a low test score as less than or equal to one (see **Supplementary File 4** for explanation of test and cutoff score).

Sample size

We based our initial sample size estimate on a single vignette per participant. We assumed a conservative baseline rate of 0.45 (i.e., at least one low-value option chosen in 45% of vignettes) in the no nudge group based on a study of general practitioner use of imaging and medicines for low back pain at first consultation.[7] With an allocation proportion of 0.5 to each of the nudges (0.25 to each of the four experimental groups), we estimated a sample of 120 observations would provide 80% power at the 0.025 significance level (to account for multiple comparisons) to detect a difference between no nudge and nudge groups of approximately 25% due to exposure to either nudge. To obtain more reliable estimates and gain additional power to detect smaller effects, each participant was exposed to four vignettes. To account for the likelihood of correlated outcomes from participants completing multiple vignettes, we utilised the sample size formula for clustered binomial outcomes given in Lee and Dubin.[26] Assuming that within-subject responses would follow a beta-binomial distribution (with parameters a=0.5 and b=0.5), with 60 participants per cluster (480 observations in total), we estimated that we could detect between-group differences in the proportion of vignettes where participants choose any low-value option as small as 15% to test the main effect of either of the nudges with more than 80% power and an alpha of 0.025.

We undertook a planned, blinded interim analysis after 20 participants had provided complete responses to ensure sample size assumptions were reasonable.

Statistical plan

To determine the main effects of the default option and partition display nudges on the odds of choosing low-value options we fitted a logistic regression model. The binary independent variables were default option nudge (yes, no) and partition display nudge (yes, no). The outcome variable for the primary model was defined as at least one low-value option selected (yes, no). We specified 'at least one low-value option selected' (no) as the reference category.

For our secondary outcome measure (the Likert score assigned to each option), to estimate the effect of the intervention on the likelihood of recommending low-value options we fitted a linear regression model. The binary independent variables were the same as those for the primary outcome analysis above, but the outcome variable was the mean of the Likert scores assigned to low-value options (continuous scale).

For both these models (for the primary and secondary outcomes), we used generalised estimating equations (GEE) model to account for repeated (correlated) observations across vignettes. To assess the robustness of our results we ran these models as a univariable regression (i.e. with the default option nudge as the sole independent variable, omitting the partition display nudge, and vice versa) and as a multivariable regression (with both nudges included in the model). We also checked for the presence of interaction effects between nudges. Our exploratory hypothesis was that if there was an interaction effect present, it would be likely to be positive (i.e. combining the two nudges would be super-additive) rather than negative.

All descriptive statistics for the statistical analysis were undertaken in Excel 2016 and all regression models were undertaken in SPSS Statistics Version 24.

RESULTS

We invited 183 GPs to participate. 120 GPs agreed to participate in the trial and provided complete outcome responses (n=480 scenarios), a response rate of 65.6%. The full participant flow is summarised in **Figure 2**. Demographic characteristics of the participants are summarised in **Table 1**. Participants were predominantly male, practised in urban areas and spent less than 15% of working hours teaching or in academia. The majority had practised clinically for more than 20 years. None of the participants had fewer than 5 years' experience. Participants mostly reported using MedicalDirector and Best Practice software. The majority had a high Cognitive Reflection Test score (i.e. they were thinking 'slow') meaning they were theoretically more likely to override their cognitive biases.

Table 1. Characteristics of participants

	All participants (n=120)	No nudge (n=30)	Partition display nudge (n=31)	Default option nudge (n=30))	Double nudge (n=29)
Gender					
Male	86 (71.7%)	21 (70%)	21 (67.7%)	25 (83.3%)	19 (65.5%)
Female	34(28.3%)	9 (30%)	10 (32.3%)	5 (16.7%)	10 (34.5%)
Practice Management Software					
Medical Director	55 (45.8%)	12 (40%)	16 (51.6%)	15 (50%)	12 (41.4%)
Best Practice	52 (43.3%)	15 (50%)	11 (35.5%)	10 (33.3%)	16 (55.2%)
Other	13 (10.8%)	3 (10%)	4 (12.9%)	5 (16.7%)	1 (3.4%)
Years of clinical practice					
<5	0	0	0	0	0

5 to 10	4 (3.3%)	1 (3.3%)	0	1 (3.3%)	2 (6.9%)
11 to 15	22 (18.3%)	11 (36.7%)	3 (9.7%)	5 (16.7%)	3 (10.3%)
16 to 20	16 (13.3%)	3 (10.0%)	5 (16.1%)	4 (13.3%)	4 (13.8%)
>20	78 (65%)	15 (50.0%)	23 (74.2%)	20 (66.7%)	20 (69.0%)
Location of clinical practice					
Urban	107 (89.2%)	28 (93.3%)	28 (90.3%)	27 (90.0%)	24 (82.8%)
Rural	12 (10%)	1 (3.3%)	3 (9.7%)	3 (10.0%)	5 (17.2%)
Remote	1 (0.8%)	1 (3.3%)	0	0	0
% of working hours teaching/academia					
0 to 15%	102 (85%)	26 (86.7%)	27 (87.1%)	28 (93.3%)	21 (72.4%)
16 to 25%	5 (4.2%)	1 (3.3%)	1 (3.2%)	1 (3.3%)	2 (6.9%)
26 to 50%	0	0	0	0	0
51 to 75%	0	0	0	0	0
More than 75%	13 (10.8%)	3 (10.0%)	3 (9.7%)	1 (3.3%)	6 (20.7%)
Cognitive Reflection Test Scores					
Mean Cognitive Reflection Test score (0 to 3)	1.8	1.7	1.7	2.0	1.8
High Cognitive Reflection Test score (%)	76 (63.3%)	16 (53.3%)	18 (58.1%)	23 (76.7%)	19 (65.5%)
Low Cognitive Reflection Test score (%)	44 (36.7%)	14 (46.7%)	13 (41.9%)	7 (23.3%)	10 (34.5%)

Primary outcome measure

Participants in the conventional display (no nudge) condition chose at least one low-value option in 42.5% of scenarios. The proportion of scenarios where participants chose at least one low-value option was lower in the default option nudge and double nudge conditions than for those allocated to the no nudge and partition display nudge conditions (**Figure 3**). The proportion of scenarios where participants chose at least one low-value option for the default option and double nudge (which included the default option nudge) condition groups were similar (30% for the default option nudge and 31% for the double nudge). The default option nudge effects appeared to apply to both imaging and medication vignettes (**Supplementary file 5**).

Our multivariable logistic regression model estimated that the partition display nudge had no effect on choice of low-value care (odds ratio [OR] 1.08, 95% CI 0·72 to 1.64; p=0.7). Participants exposed to the default option nudge condition were 44% less likely to choose at least one low-value option (odds ratio [OR] 0.56, 95% CI 0·37 to 0.85; p=0·006) compared to those not exposed to the default option nudge. Results were robust to whether the effects of nudges were tested in univariable (single nudge) or multivariable (account for both nudges) models. There was no statistically significant interaction effect (odds ratio [OR] 0.94, 95% CI 0·41 to 2.15; p=0·89).

Secondary outcome measure

Mean Likert scores (*1 = very unlikely to recommend and 5 = very likely to recommend*) for the three low-value choices were 2.27 (95% CI 2.07 to 2.47) for participants in the partition display nudge condition and 2.17 (95% CI 1.92 to 2.42) for participants in the default option nudge condition compared to 2.34 (95% CI 2.18 to 2.5) for participants in the no nudge condition. Mean Likert scores for the three guideline concordant choices were 2.78 (95% CI 2.56 to 3.01) for participants in the partition display nudge condition and 3.01 (95% 2.84 to 3.18) for participants in the default option nudge condition compared to 2.72 (95% CI 2.56 to

2.87) for participants in the no nudge condition. Our linear regression model did not find any statistically significant effect of the partition display nudge (-0.12, 95% CI -0.36 to 0.12; p=0·33) or the default option nudge (-0.23, 95% CI -0.47 to 0.02; p=0·07) on the secondary outcome measure. Results were robust to whether the effects of nudges were examined in univariable (single nudge) or multivariable (account for both nudges) models. There was no statistically significant interaction effect (-0.11, 95% CI -0.6 to 0.37; p=0·65).

Other exploratory hypotheses

Because post-randomisation drop-outs were negligible (there were a total of three incompletely answered surveys) we did not proceed with the planned analysis of the effect of the interventions on drop-out rate.

Table 2. Cognitive characteristics of participants selecting low-value options in default option nudge vs no default option nudge conditions

	Default option nudge (n=59)	No default option nudge (n=61)
	% of scenarios where low-value options selected	% of scenarios where low-value options selected
All participants	30.5	43.9
Participants with low Cognitive Reflection Test score (thinking 'fast')	25	44.4
Participants with high Cognitive Reflection Test score (thinking 'slow')	32.7	43.4

We hypothesised that participants with low Cognitive Reflection Test scores (i.e. those who were thinking 'fast') would be more susceptible to nudges than participants with high test scores. **Table 2** shows that in those allocated to a default option condition, the effect of the

default option nudge appeared to be higher for those with low Cognitive Reflection Test scores (in 25% of scenarios at least one low-value option was chosen) compared to those with a high Cognitive Reflection Test score (in 32.7% of scenarios at least one low-value option was chosen). The majority (71.6%) of participants exposed to the default option nudge condition 'clicked through' to see the non-default options in all vignettes.

DISCUSSION

Embedding a default option nudge in a simulated electronic decision support display–where high-value options were presented as the default options and low-value options were revealed only after clicking for more–reduced the odds of general practitioners selecting low-value care by 44%. Embedding a partition display nudge had no effect on the odds of choosing low-value options. Combining default option and partition display nudges in the decision support had no additional effect on the odds of choosing low-value options.

Comparison to other research findings

Three non-randomised studies[14 15 16] have found that use of a default option nudge in Electronic Health Record systems where generic medications were presented as default options increased the odds these medications would be prescribed in outpatient and primary care clinical settings. One non-randomised study and a quasi-randomised study found that embedding default option nudges in electronic order sets could influence test ordering behaviour. Munigala[17] found that removing urine tests other than the default test from the 'frequently ordered' window of the emergency department electronic order set decreased ordering of urine cultures. Leis[18] found that whether or not a test was included as part of a standard admission order set strongly influenced whether the test was ordered at all even if there was no explicit prohibition on ordering it.. Our study adds important data to this literature; it is the only randomised controlled trial that we are aware of to study the use of

default option nudges in a clinical context, albeit through the use of vignettes. It provides further robust evidence of the impact of default option nudges on prescribing of medicines and suggests these effects could extend to test ordering behaviour.

Importantly, our study failed to replicate the findings of Tannenbaum et al (2017)[12] who found partition display nudges could reduce low value clinical decision making about antibiotics.[12] We were surprised by this finding, given the similarities between our methodology and that of Tannenbaum et al.[12] Both studies used a randomised factorial design, and vignettes where physicians had to decide between aggressive/complex medicines and simple medicines.

However, there were also important differences between our study and Tannenbaum's.[12] Their study was focused on decisions about antibiotics whereas ours were about management of low back pain. It remains possible that partition display nudges are more effective for some clinical management decisions than others. For example, these nudges may be more effective when tested in contexts where providers do not hold strong preconceived notions about the importance of a given test or treatment. In Tannenbaum's study[12], where the focus was on use of antibiotics, providers may have had fewer preconceived notions regarding specific antibiotic types. Clinicians assessing patients with low back pain, on the other hand, may believe imaging is a valuable test, and be committed to a treatment plan that includes diagnostic workup. This could make them less susceptible to the effect of partition display nudges when provided with vignettes about low back pain. Curiously the physician sample in Tannenbaum's study were mostly academic. In our study 85% spent less than 15% of working hours teaching or in academia. It is unclear whether time spent in academia could influence susceptibility to nudges. A final difference between the two studies is that the low-value options grouped by Tannenbaum under a partition display nudge formed clearly discrete clinical categories e.g. broad spectrum (as opposed to narrow spectrum) antibiotics and prescription (as opposed to over the counter) drugs. By

contrast, the low-value options grouped under the partition display nudges in our study comprised a more heterogenous set of choices.

Strengths, weaknesses and clinical implications

Key strengths of our study are the randomised, factorial design to provide evidence on the effects of two nudges. We pre-registered our trial protocol and conducted all procedures and analyses as planned. Our analysis was adequately powered and based on 480 patient scenarios. We used planned statistical techniques to ensure that repeated measures from the same participant were accounted for. Our vignettes were developed by experienced clinicians to ensure that our scenarios and management options questions were realistic. The management options provided in the vignettes reflected the current best evidence and clinical practice pathways in our geographical area.[22 23 27] Another strength is the application of nudges to diagnostic imaging decisions. Given the cost of these tests and the prevalence of low back pain, using default option nudges to reduce unnecessary imaging has the potential for substantial cost savings if shown to be effective in a real-world setting.

This study provides insights on the importance of the 'timing' of nudges and their interaction with the psychology of practitioners being nudged. In particular, we found that (i) nudges did not have an effect on measures of low-value decision making collected shortly after exposure; (ii) participants with a low Cognitive Reflection Test score (who therefore think 'fast') may be more susceptible to a default option nudge. These findings suggest that nudges may be most effective in situations where thinking 'slow' is difficult or discouraged. They also raise the question of when it is most effective to introduce a nudge during a GP consultation with patients, as there is a possibility that the nudge may wear off after a lengthy conversation with the patient.

This study also has limitations. It is reasonable to ask whether findings from hypothetical vignettes can be fully indicative of clinicians' behaviour in a 'real world' setting. Nonetheless,

15

vignette-based studies are widely used and accepted in research into medical choices.[28 29] There may be important differences between our sample and the broader Australian GP population e.g. our sample had a lower share of early career physicians compared to the Australian sample. [30] However, it is unclear whether this means the nudges would be more or less effective in a more representative GP sample.

An important limitation of the current study is that it does not directly reflect current GP interaction with existing Electronic Health Record systems. For example, current practice management software systems use drop down menus and predictive texts to locate tests and treatments. However, as software systems and decision support tools evolve, there may be scope for our results to inform how management choices are presented on screen. Default option nudges are not necessarily limited to the specific format or clinical management choices we used in this study and could be implemented in new ways as these systems and tools evolve.

Moreover, while it may not necessarily reflect current workflow we believe there are important reasons to consider the results of nudging studies such as this one. First, despite being a potentially powerful tool for enhancing the safety, quality and effectiveness of healthcare without compromising clinician autonomy,[19 21] clinician nudges are still relatively understudied. Second, there is uncertainty around the specific contexts in which these nudges are most effective and the myriad factors which may modify these effects.[31 32] We believe the evidence base for quality improvement interventions would benefit from an inventory of well-researched examples of nudging in different contexts. These would include the target (e.g. medications, pathology tests, imaging tests), desired actions (avoiding something unnecessary versus doing something recommended), the urgency of the clinical situation, the degree to which patients are involved in decision-making, and so on. The insights that such a research programme can provide into nudge principles will benefit the design of future clinical decision aids, such as Electronic Health Record systems, electronic order sets and hand-held apps.

CONCLUSION

Clinical mismanagement of low back pain is an international problem. Traditional education-based approaches to quality improvement have had limited effectiveness in changing clinician behaviour. Our study suggests that embedding a default option nudge into clinical decision support tools can reduce the choice of low-value care options for low back pain. This has the potential to significantly improve patient care while preserving autonomous clinical decision-making.

REFERENCES

1. Scott IA, Duckett SJ. In search of professional consensus in defining and reducing low-value care. *The Medical journal of Australia* 2015;203(4):179-81.
2. Levinson W, Kallewaard M, Bhatia RS, et al. 'Choosing Wisely': a growing international campaign. *BMJ quality & safety* 2015;24(2):167-74. doi: 10.1136/bmjqs-2014-003821.
3. Mafi JN, Parchman M. Low-value care: an intractable global problem with no quick fix. *BMJ quality & safety* 2018;27(5):333-36. doi: 10.1136/bmjqs-2017-007477
4. Foster NE, Anema JR, Cherkin D, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. *Lancet (London, England)* 2018;391(10137):2368-83. doi: 10.1016/s0140-6736(18)30489-6
5. Webster BS, Bauer AZ, Choi Y, et al. Iatrogenic consequences of early magnetic resonance imaging in acute, work-related, disabling low back pain. *Spine* 2013;38(22):1939-46. doi: 10.1097/BRS.0b013e3182a42eb6
6. Jarvik JG, Hollingworth W, Martin B, et al. Rapid magnetic resonance imaging vs radiographs for patients with low back pain: a randomized controlled trial. *Jama* 2003;289(21):2810-8. doi: 10.1001/jama.289.21.2810

7. Williams CM, Maher CG, Hancock MJ, et al. Low back pain and best practice care: A survey of general practice physicians. *Archives of internal medicine* 2010;170(3):271-7. doi: 10.1001/archinternmed.2009.507

8. Mathieson S, Valenti L, Maher CG, et al. Worsening trends in analgesics recommended for spinal pain in primary care. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society* 2018;27(5):1136-45. doi: 10.1007/s00586-017-5178-4

9. Grimshaw J, Eccles M, Tetroe J. Implementing clinical guidelines: current evidence and future implications. *The Journal of continuing education in the health professions* 2004;24 Suppl 1:S31-7. doi: 10.1002/chp.1340240506

10. Eccles M, Steen N, Grimshaw J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. *Lancet (London, England)* 2001;357(9266):1406-9. doi: 10.1016/s0140-6736(00)04564-5

11. Scott IA, Soon J, Elshaug AG, et al. Countering cognitive biases in minimising low value care. *The Medical journal of Australia* 2017;206(9):407-11.

12. Tannenbaum D, Doctor JN, Persell SD, et al. Nudging physician prescription decisions by partitioning the order set: results of a vignette-based study. *Journal of general internal medicine* 2015;30(3):298-304. doi: 10.1007/s11606-014-3051-2

13. Abadie A, Gay S. The impact of presumed consent legislation on cadaveric organ donation: a cross-country study. *Journal of health economics* 2006;25(4):599-620. doi: 10.1016/j.jhealeco.2006.01.003

14. Patel MS, Day SC, Halpern SD, et al. Generic Medication Prescription Rates After Health System-Wide Redesign of Default Options Within the Electronic Health Record. *JAMA internal medicine* 2016;176(6):847-8. doi: 10.1001/jamainternmed.2016.1691

15. Patel MS, Day S, Small DS, et al. Using default options within the electronic health record to increase the prescribing of generic-equivalent medications: a quasi-experimental study. *Annals of internal medicine* 2014;161(10 Suppl):S44-52. doi: 10.7326/m13-3001

16. Malhotra S, Cheriff AD, Gossey JT, et al. Effects of an e-Prescribing interface redesign on rates of generic drug prescribing: exploiting default options. *Journal of the American Medical Informatics Association : JAMIA* 2016;23(5):891-8. doi: 10.1093/jamia/ocv192

17. Munigala S, Jackups RR, Jr., Poirier RF, et al. Impact of order set design on urine culturing practices at an academic medical centre emergency department. *BMJ quality & safety* 2018;27(8):587-92. doi: 10.1136/bmjqs-2017-006899

18. Leis B, Frost A, Bryce R, et al. Standard admission order sets promote ordering of unnecessary investigations: a quasi-randomised evaluation in a simulated setting. *BMJ quality & safety* 2017;26(11):938-40. doi: 10.1136/bmjqs-2017-006898 [published Online First: 2017/09/16]

19. Patel MS, Volpp KG, Asch DA. Nudge Units to Improve the Delivery of Health Care. *The New England journal of medicine* 2018;378(3):214-16. doi: 10.1056/NEJMp1712984

20. 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(6):562-70. doi: 10.1001/jama.2016.0275

21. Vaughn VM, Linder JA. Thoughtless design of the electronic health record drives overuse, but purposeful design can nudge improved patient care. *BMJ quality & safety* 2018;27(8):583-86. doi: 10.1136/bmjqs-2017-007578

22. Traeger A, Buchbinder R, Harris I, et al. Diagnosis and management of low-back pain in primary care. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* 2017;189(45):E1386-e95. doi: 10.1503/cmaj.170527

23. Chou R, Qaseem A, Owens DK, et al. Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians. *Annals of internal medicine* 2011;154(3):181-9. doi: 10.7326/0003-4819-154-3-201102010-00008

24. Pennycook G, Cheyne JA, Koehler DJ, et al. Is the cognitive reflection test a measure of both reflection and intuition? *Behavior research methods* 2016;48(1):341-8. doi: 10.3758/s13428-015-0576-1

25. Frederick S. Cognitive reflection and decision making. *Journal of Economic perspectives* 2005;19(4):25-42.

26. Lee EW, Dubin N. Estimation and sample size considerations for clustered binary responses. *Statistics in medicine* 1994;13(12):1241-52. [published Online First: 1994/06/30]

27. Mid North Coast and Northern NSW Local Health District. Health Pathways Mid and North Coast of NSW 2018 [Available from: https://manc.healthpathways.org.au/index.htm accessed 28 July 2018 2018.

28. Jiwa M, Gordon M, Arnet H, et al. Referring patients to specialists: a structured vignette survey of Australian and British GPs. *BMC family practice* 2008;9:2. doi: 10.1186/1471-2296-9-2

29. Peabody JW, Luck J, Glassman P, et al. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *Jama* 2000;283(13):1715-22.

30. Britt H MG, Bayram C, Henderson J, Valenti L, Harrison C, Pan Y, Charles J, Pollack AJ, Chambers T, Gordon J, Wong C. A decade of Australian general practice activity 2006–07 to 2015–16. General practice series no. 41. Sydney: Sydney University Press. 2016.

31. Mongin P, Cozic M. Rethinking nudge: not one but three concepts. *Behavioural Public Policy* 2018;2(1):107-24.

32. Hertwig R. When to consider boosting: some rules for policy-makers. *Behavioural Public Policy* 2017;1(2):143-61.

LEGENDS

Figure 1. Visual appearance of the partition display nudge (i) and default option nudge (ii) on a computer screen.

The partition display nudge (i) displayed the three guideline concordant options first in a vertical list, and then displayed the low value options in a horizontal list. The default option nudge (ii) displayed the three guideline concordant options first in a vertical list, followed by a message stating 'click to see more options' below third option. Upon one click, the low-value options were made fully visible in a vertical list.

Figure 2. Participant study flow

Figure 3. Proportion of scenarios where at least one low-value option was selected