PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Nudge interventions to reduce unnecessary antibiotic prescribing in primary care: a systematic review
AUTHORS	Raban, Magdalena; Gonzalez, Gabriela; Nguyen, A; Newell, Ben; Li, Ling; Seaman, Karla; Westbrook, Johanna

VERSION 1 – REVIEW

REVIEWER	Saha, Sajal Monash University, Australia, General Practice
REVIEW RETURNED	13-Apr-2022

GENERAL COMMENTS	This is an interesting systematic review assessing the type and effect of nudge interventions on antibiotic prescribing in primary care. Overall, the study has been well designed and explicitly reported. The are few areas which could be improved for clarity and readership Abstract • Naming risk of bias tool in the abstract would be good • Giving analysis details could improve abstract reporting • Study outcomes are not clear in the abstract methods though it is clearly stated in the last sentence of the background in the text • What was the reason behind including “Regression discontinuity studies”? state somewhere. • “Vote counting was applied to synthesise effects on overall antibiotic prescribing”-not clear what does it mean in the abstract? This has been clarified and described in the text. Suggest delete from abstract • It is unclear if high risk of bias studies were excluded while analysed? Though stated in the result section “Removing studies with a high risk of bias, the percentage of studies showing a reduction in overall antibiotic prescribing was 76.5% (n=12, 95% CI: 52.7, 90.4) • Distinction between nudge interventions and social norm nudges can be explained in the background as it is being analysed and reported-This would help readers to understand the terms Introduction: • Introduction has been well written. Fourth paragraph line 10 to 33, Nudges have been explained with examples which is great but apart from audit feedback, few more examples can be explained and what component we exactly call nudge intervention, need a bit more clarification. Methods

	• Eligibility criteria can be divided into inclusion and exclusion criteria using separate paragraph and dot points for readers. • The criteria used to determine whether an intervention was a nudge can be explained as dot points though authors use a definition. • When a study trialled more than one nudge intervention how did author analyse and report intervention effect? • Social norm feedback can be clearly defined to improve understanding this nudge by readers • Line 57, commonly used Results • I would suggest a summary table demonstrating effect sizes against type and categories of nudge interventions. Effect sustainability if assessed or remained. Inclusion of the name of nudge and definition would be great.

REVIEWER	Fox, C R University of California Los Angeles, Anderson School of Management
REVIEW RETURNED	19-Apr-2022

GENERAL COMMENTS	Review of “Nudge interventions to reduce unnecessary antibiotic prescribing in primary care: A systematic review” This paper presents an early systematic review of nudge interventions to reduce inappropriate antibiotic prescribing for respiratory and urinary tract infections in primary care practices. The authors identify 17 studies involving 23 different interventions, the large majority of which involve the use of social norms. I write my review not as a health scientist but as a behavioral scientist with some experience in applying nudges in this domain. SUMMARY EVALUATION. I appreciate the instinct to take stock of this promising and growing area of interest. Moreover, a strength of this paper is the steps that the authors take to make their review and analysis systematic, given that meta-analysis is not practical in this domain given the heterogeneity of interventions. On the other hand, the tools used to take stock of the literature in a systematic manner are of limited inferential value, to my reading. Because this literature is relatively sparse (only 17 papers using diverse methods, including at least one published pilot study {Persell 2016} of another study included in the review {Meeker 2016}), I’m not sure how much we can hope to learn about antibiotic nudges from this review above and beyond the conclusion that “many of these kinds of interventions look promising.” What’s more, there already exists at least one very recent and fairly comprehensive narrative review of this same topic (though it was, arguably, less systematic): Richards, A. R., & Linder, J. A. (2021). Behavioral economics and ambulatory antibiotic stewardship: a narrative review. Clinical therapeutics, 43(10), 1654-1667. CONSTRUCTIVE SUGGESTION. Where does this leave us? I’m thinking that given that a very high proportion of studies identified

	in the present paper relied on various social norm interventions, a possibly more useful and unique contribution might more explicitly focus attention on social norm interventions, mentioning other approaches only in passing. This article might further aspire to identify key implementation details that predict greatest success for social norm interventions in this context. Personally, I'd like to see a more detailed account of variations in implementation of social norms across studies, with some speculation about which features promote effectiveness of such nudges. There are so many possible variations of social norm nudges beyond what has been coded for here. For example, descriptive versus injunctive norms, static versus dynamic norms, public versus private sharing, aspirational group performance versus average performance, all antibiotics vs. diagnosis-inappropriate prescriptions, highlighting absolute performance versus relative performance of providers, etc. Norms can be communicated via various channels and capture attention in various ways. Incentives for participation and salience of monitoring can vary. And so forth. If the review examined and interpreted (or speculated about) the impact of such details across studies it would be less systematic, but possibly more useful. I appreciate that the review does currently code for a few of these variables, qualify studies by potential for bias, and count results. On the other hand, there are so many procedural and contextual differences between studies that are scored as similar on the coded dimensions that it is hard to know what to make of this analysis in light of those confounds. Thus, I'd rather see a deeper (if more speculative) analysis of how various implementation details affect results when it comes to social norms that aspire to reduce (diagnosis inappropriate) antibiotic prescriptions. MINOR COMMENTS 1) p.5, line 47-49: in connection with the point that "the effect of nudges can vary depending on the context in which they are applied," I'll offer a shameless plug for a paper that appeared in the BMJ and mentions antibiotic examples: Fox, C. R., Doctor, J. N., Goldstein, N. J., Meeker, D., Persell, S. D., & Linder, J. A. (2020). Details matter: predicting when nudging clinicians will succeed or fail. bmj, 370. 2) p.9 paragraph 2: why were these features of social norms selected for recording and analysis? 3) p.19, lines 24-34: The Persell et al. (2016) pilot study found a potentially interesting Hawthorne-like effect in the control condition but no evidence that any of the interventions were significantly more effective at reducing diagnosis-inappopropriate prescribing (the primary outcome) that merely enrolling patients into a study, providing a bland education module, and paying them for participation. The reduction in the suggested alternative and peer comparison conditions was only significant for the secondary outcome of all Abx prescriptions combined. So it is arguably misleading to present the Persell et al. pilot as a success in those two conditions. Of course as a pilot the study was under-powered. 4) P.22, paragraph 2: I appreciate that the authors mention nudges applied elsewhere but not used in studies in this literature. If the authors decide to focus on social norm interventions as I suggest,

	there are many variations of social norm interventions that have not yet been tried in this literature. It is becoming a rich literature in social psychology and organizational behavior as well as other domains of health care, and there have been many alternative approaches to implementing social norms that might be suggested for Abx prescribing. 5) P.22, paragraph 3: Of course the boomerang effect can be mitigated by supplementing descriptive social norms with injunctive norms, as the authors later observe. However, there have been studies recently in other domains that find that targeting very poor performers with feedback that they are falling very far behind a norm can be demotivating (compared to those whose performance is closer to the identified norm). 6) P.23, paragraph 1: I think the virtue of the aspirational norm identifying top performers (with 0% inappropriate prescribing rates) as in Meeker et al. (2006) is that there are an unlimited number of providers who can become "top performers.". In other studies, where the cutoff for defining the aspirational group keeps on changing, the moving target and competition may be counterproductive. In fact, there is a paper by Dai et al., in press at PNAS that finds publicly identifying 25 best performers each month may be ineffective and contribute to lower morale and higher burnout among physicians. So even aspirational norms can backfire depending on how they are implemented.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

	Comment	**Response**
7	This is an interesting systematic review assessing the type and effect of nudge interventions on antibiotic prescribing in primary care. Overall, the study has been well designed and explicitly reported. The are few areas which could be improved for clarity and readership	We thank the reviewer for their comment.
8	Abstract • Naming risk of bias tool in the abstract would be good	The risk of bias tool has been named in the Abstract
9	• Giving analysis details could improve abstract reporting	We have added a sentence to describe the method of synthesis.
10	• Study outcomes are not clear in the abstract methods though it is clearly stated in the last sentence of the background in the text	We have revised the objectives of the study in the Abstract to be consistent with the text and include detail of the main outcome of interest.

11	• What was the reason behind including “Regression discontinuity studies”? state somewhere.	We have provided an explanation for this in the Methods section as it is too much detail for the abstract: “Regression discontinuity studies allow assessment of causality in studies where a cut-off point is used to allocate an intervention. This is of particular relevance to social norm nudges, where, e.g. the bottom 10% performers are target by an intervention. Studies have shown that regression discontinuity studies have similar effect estimates to randomised trials, though they require a large sample size.[26, 27]” (p. 8)
12	• “Vote counting was applied to synthesise effects on overall antibiotic prescribing”-not clear what does it mean in the abstract? This has been clarified and described in the text. Suggest delete from abstract	We have removed this sentence from the abstract, and have replaced it with detail of how results were synthesised.
13	• It is unclear if high risk of bias studies were excluded while analysed? Though stated in the result section “Removing studies with a high risk of bias, the percentage of studies showing a reduction in overall antibiotic prescribing was 76.5% (n=12, 95% CI: 52.7, 90.4)	Due to the limited word count, the Abstract presents the results overall and does not exclude studies with a high risk of bias: “Overall, 78.3% (n=23, 95% CI: 58.1, 90.3) of the nudges evaluated reported a reduction in overall antibiotic prescribing rates.” However, the total number of nudges tested, and the number of studies with a high risk of bias are clearly reported. The reader can deduce that since 23 nudge interventions were evaluated in the studies, the results reported are for all interventions. The extra detail excluding studies with a high risk of bias is provided in the main text of the paper.
14	• Distinction between nudge interventions and social norm nudges can be explained in the background as it is being analysed and	While we agree this would be a nice inclusion, there is limited word count in the abstract to allow further elaboration on this detail. However, we have added text to the

	reported-This would help readers to understand the terms	Abstract Methods to state results were examined by the type of nudge.
915	Introduction has been well written. Fourth paragraph line 10 to 33, Nudges have been explained with examples which is great but apart from audit feedback, few more examples can be explained and what component we exactly call nudge intervention, need a bit more clarification.	We have added the following sentence with examples of nudges to the end of the fourth paragraph: "Examples of nudge interventions include changing the default options, changing option consequences, and providing reminders during the decision-making process." (p. 6)
16	Methods • Eligibility criteria can be divided into inclusion and exclusion criteria using separate paragraph and dot points for readers.	While we agree that this is one approach to presenting eligibility criteria, we have chosen a different approach which is consistent with the PRISMA guidelines. The current format allows us to further discuss and justify our criteria.
17	• The criteria used to determine whether an intervention was a nudge can be explained as dot points though authors use a definition.	Please see our response to above comment re the format of our 'Eligibility criteria' section. However, we have taken the reviewer's suggestion and applied it to the comment below (comment 19) about social norm nudges by presenting the features extracted in a bullet point list in Box 1.
18	• When a study trialled more than one nudge intervention how did author analyse and report intervention effect?	Study data were extracted for the impact of each individual nudge, rather than the combined effect of multiple nudges together. We have clarified this is in the Methods as follows (new text in italics): "When a study trialled more than one nudge intervention, *we extracted data on the impact of each nudge individually*." (p. 9)

19	• Social norm feedback can be clearly defined to improve understanding this nudge by readers	We have elaborated on social norm nudges in the Methods as follows: "Social norm feedback nudge interventions are a frequent behaviour change technique in healthcare, often termed audit and feedback. Social norm feedback involves providing people with feedback on their performance relative to their peers. However, this can be implemented in a variety of way. For example, the comparison can be descriptive or injunctive, i.e. associating a judgement to the performance. Psychology and health research has shown that certain features may enhance social norm feedback interventions,[15-17, 28] and thus, we extracted details of how social norm nudges were implemented with the aim that this may further elucidate the important features of effective social norm nudges to reduce antibiotic prescribing in primary care for (Box 1)." (p.11)
20	• Line 57, commonly used	This comment is unclear.
21	I would suggest a summary table demonstrating effect sizes against type and categories of nudge interventions. Effect sustainability if assessed or remained. Inclusion of the name of nudge and definition would be great.	Supplementary file 4 contains all the effect sizes for each intervention grouped by nudge category. However, this table is too expansive to simplify and present within the main paper, as studies have used different outcomes.

Reviewer 2

	Comment	**Response**
22	I appreciate the instinct to take stock of this promising and growing area of interest. Moreover, a strength of this paper is the steps that the authors take to make their review and analysis systematic, given that meta-analysis is not practical in this domain given the heterogeneity of interventions. On the other hand, the tools used to take stock of the literature in a systematic manner are of limited inferential value, to my reading.	We thank the reviewer for drawing our attention to the published narrative review. We believe our systematic review builds on this work. We have also taken up the reviewer's suggestion to further examine social norm nudges (see response to comment 23 below).

<table>
<tr>
<td></td>
<td>Because this literature is relatively sparse (only 17 papers using diverse methods, including at least one published pilot study {Persell 2016} of another study included in the review {Meeker 2016}), I'm not sure how much we can hope to learn about antibiotic nudges from this review above and beyond the conclusion that "many of these kinds of interventions look promising." What's more, there already exists at least one very recent and fairly comprehensive narrative review of this same topic (though it was, arguably, less systematic):

Richards, A. R., & Linder, J. A. (2021). Behavioral economics and ambulatory antibiotic stewardship: a narrative review. Clinical therapeutics, 43(10), 1654-1667.</td>
<td>Furthermore, we have again carefully considered whether to include the Meeker 2016 and Persell 2016 papers as separate studies, given Persell was a pilot study. We note both studies are listed as separate protocols on ClinicalTrials.gov and though have overlapping dates, were held in different locations (summary table below). Thus, we believe there is enough justification to include these as separate studies.

<table>
<tr><td>Persell 2016

(NCT01454960)</td><td>Jul 2011-Sep 2014</td><td>Chicago</td></tr>
<tr><td>Meeker 2016 (NCT01454947)</td><td>Aug 2011-Sep 2014</td><td>Massachusetts and Southern California</td></tr>
</table>

We have, however, clarified in the text and that the Persell study was a pilot as follows:

"One study was a pilot study[45] of a larger trial [43], but was included as a separate study as it was conducted in a different population." (Results, first paragraph, second sentence)</td>
</tr>
<tr>
<td>23</td>
<td>Where does this leave us? I'm thinking that given that a very high proportion of studies identified in the present paper relied on various social norm interventions, a possibly more useful and unique contribution might more explicitly focus attention on social norm interventions, mentioning other approaches only in passing. This article might further aspire to identify key implementation details that predict greatest success for social norm interventions in this context. Personally, I'd like to see a more detailed account of variations in implementation of social norms across studies, with some speculation about which features promote effectiveness of such nudges. There are so many possible</td>
<td>We thank the reviewer for their suggestion. We have expanded our examination of the implementation of social norm nudges as suggested – we now examine 12 implementation features. To accommodate this, we have replaced the table with a figure (Figure 2).

Additionally, we have updated the harvest plot of social norm nudge features (previously figure 3). We have focused the harvest plot on three features: which prescribers were targeted, the comparison</td>
</tr>
</table>

	variations of social norm nudges beyond what has been coded for here. For example, descriptive versus injunctive norms, static versus dynamic norms, public versus private sharing, aspirational group performance versus average performance, all antibiotics vs. diagnosis-inappropriate prescriptions, highlighting absolute performance versus relative performance of providers, etc. Norms can be communicated via various channels and capture attention in various ways. Incentives for participation and salience of monitoring can vary. And so forth. If the review examined and interpreted (or speculated about) the impact of such details across studies it would be less systematic, but possibly more useful. I appreciate that the review does currently code for a few of these variables, qualify studies by potential for bias, and count results. On the other hand, there are so many procedural and contextual differences between studies that are scored as similar on the coded dimensions that it is hard to know what to make of this analysis in light of those confounds. Thus, I'd rather see a deeper (if more speculative) analysis of how various implementation details affect results when it comes to social norms that aspire to reduce (diagnosis inappropriate) antibiotic prescriptions.	group and whether an injunctive norm was used. The text in the Methods and Results sections of the paper have been updated in the following sections: • Methods, Data collection and data items • Methods, Synthesis of results • Results, Description of nudge interventions • Results, Effect of nudge interventions on overall antibiotic prescribing rates
24	p.5, line 47-49: in connection with the point that "the effect of nudges can vary depending on the context in which they are applied," I'll offer a shameless plug for a paper that appeared in the BMJ and mentions antibiotic examples: Fox, C. R., Doctor, J. N., Goldstein, N. J., Meeker, D., Persell, S. D., & Linder, J. A. (2020). Details matter: predicting when nudging clinicians will succeed or fail. bmj, 370.	We thank the reviewer for this suggestion and have added this citation.
25	p.9 paragraph 2: why were these features of social norms selected for recording and analysis?	We have expanded on this section in the Methods in line with the changes made to the comment 23. This section now reads:

		"Social norm feedback nudge interventions are a frequent behaviour change technique in healthcare, often termed audit and feedback. Social norm feedback involves providing people with feedback on their performance relative to their peers. However, this can be implemented in a variety of way. For example, the comparison can be descriptive or injunctive, i.e. associating a judgement to the performance. Psychology and health research has shown that certain features may enhance social norm feedback interventions,[15-17, 28] and thus, we extracted details of how social norm nudges were implemented with the aim that this may further elucidate the important features of effective social norm nudges to reduce antibiotic prescribing in primary care for (Box 1)." (p. 11)
26	p.19, lines 24-34: The Persell et al. (2016) pilot study found a potentially interesting Hawthorne-like effect in the control condition but no evidence that any of the interventions were significantly more effective at reducing diagnosis-inappropriate prescribing (the primary outcome) that merely enrolling patients into a study, providing a bland education module, and paying them for participation. The reduction in the suggested alternative and peer comparison conditions was only significant for the secondary outcome of all Abx prescriptions combined. So it is arguably misleading to present the Persell et al. pilot as a success in those two conditions. Of course as a pilot the study was under-powered.	Overall antibiotic use was our primary outcome of interest and hence that is the outcome we reported on in the harvest plots (Methods, Synthesis of results section). Supplementary file 4 shows the results for both outcomes. We acknowledge the point re the Hawthorne effect in the Persell study and note there could be intervention spill over to controls, since randomisation was done at the physician level and physicians were in the same health centre. However, given the additions already made in this revision, we have not commented on this due to the limited space.
27	P.22, paragraph 2: I appreciate that the authors mention nudges applied elsewhere but not used in studies in this literature. If the authors decide to focus on social norm interventions as I suggest, there are many variations of social norm interventions that have not yet been tried in this literature. It is	We agree that further research is required to identify the features of social norm nudges that would enhance intervention effects. We have modified this paragraph to reflect the expanded results from the social norm nudge comparisons, as follows:

	becoming a rich literature in social psychology and organizational behavior as well as other domains of health care, and there have been many alternative approaches to implementing social norms that might be suggested for Abx prescribing.	"Social norm feedback was the most frequently evaluated nudge, and the evidence suggests that comparisons should include an aspirational target, injunctive norm or target high prescribers to enhance intervention effects. However, future research should explore the types of features that will further enhance social norm feedback nudges in this context. Only four studies examined nudge strategies other than social norm nudges, such as changing option consequences, providing reminders and facilitating commitment, thus further research is also needed to evaluate other nudge strategies despite promising results thus far of their effectiveness." (Discussion, first paragraph)
28	P.22, paragraph 3: Of course the boomerang effect can be mitigated by supplementing descriptive social norms with injunctive norms, as the authors later observe. However, there have been studies recently in other domains that find that targeting very poor performers with feedback that they are falling very far behind a norm can be demotivating (compared to those whose performance is closer to the identified norm).	We have included this point in the fourth paragraph of the Discussion. "However, care should be taken when deciding on the comparison group, as if becoming a 'top performer' is perceived as unattainable, this can be demotivating. This can occur when the comparison norm is dynamic, i.e. changes according the group's behaviour, which was the case in all our studies that provided feedback more than once (Figure 2). For example, if the comparison group is consistently the top 10%, 90% of people will never reach the target. One study included in our review reported that the top 10% of prescribers did not change their prescribing behaviour following the social norm nudge, despite an overall reduction following the intervention.[46] The authors speculate this may be due to the message not motivating behaviour change. Furthermore, individuals need to trust the data being presented is an accurate representation of their performance, and in the case of antibiotic prescribing, adequately accounts for the clinical features of the populations they treat."

29	P.23, paragraph 1: I think the virtue of the aspirational norm identifying top performers (with 0% inappropriate prescribing rates) as in Meeker et al. (2006) is that there are an unlimited number of providers who can become “top performers.”. In other studies, where the cutoff for defining the aspirational group keeps on changing, the moving target and competition may be counterproductive. In fact, there is a paper by Dai et al., in press at PNAS that finds publicly identifying 25 best performers each month may be ineffective and contribute to lower morale and higher burnout among physicians. So even aspirational norms can backfire depending on how they are implemented.	We have included this point in the fourth paragraph of the Discussion – see response to comment 28.