BMJ Open

Systematic review of clinician-directed nudges in healthcare contexts

Briana S Last ,[1] Alison M Buttenheim,[2,3,4] Carter E Timon,[5] Nandita Mitra,[6] Rinad S Beidas[3,4,7,8,9]

To cite: Last BS, Buttenheim AM, Timon CE, *et al*. Systematic review of clinician-directed nudges in healthcare contexts. *BMJ Open* 2021;**11**:e048801. doi:10.1136/bmjopen-2021-048801

► Prepublication history and supplemental material for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-048801).

Received 08 January 2021
Accepted 12 May 2021

For numbered affiliations see end of article.

Correspondence to
Ms Briana S Last;
brishiri@sas.upenn.edu

ABSTRACT

Objective Nudges are interventions that alter the way options are presented, enabling individuals to more easily select the best option. Health systems and researchers have tested nudges to shape clinician decision-making with the aim of improving healthcare service delivery. We aimed to systematically study the use and effectiveness of nudges designed to improve clinicians' decisions in healthcare settings.

Design A systematic review was conducted to collect and consolidate results from studies testing nudges and to determine whether nudges directed at improving clinical decisions in healthcare settings across clinician types were effective. We systematically searched seven databases (EBSCO MegaFILE, EconLit, Embase, PsycINFO, PubMed, Scopus and Web of Science) and used a snowball sampling technique to identify peer-reviewed published studies available between 1 January 1984 and 22 April 2020. Eligible studies were critically appraised and narratively synthesised. We categorised nudges according to a taxonomy derived from the Nuffield Council on Bioethics. Included studies were appraised using the Cochrane Risk of Bias Assessment Tool.

Results We screened 3608 studies and 39 studies met our criteria. The majority of the studies (90%) were conducted in the USA and 36% were randomised controlled trials. The most commonly studied nudge intervention (46%) framed information for clinicians, often through peer comparison feedback. Nudges that guided clinical decisions through default options or by enabling choice were also frequently studied (31%). Information framing, default and enabling choice nudges showed promise, whereas the effectiveness of other nudge types was mixed. Given the inclusion of non-experimental designs, only a small portion of studies were at minimal risk of bias (33%) across all Cochrane criteria.

Conclusions Nudges that frame information, change default options or enable choice are frequently studied and show promise in improving clinical decision-making. Future work should examine how nudges compare to non-nudge interventions (eg, policy interventions) in improving healthcare.

Strengths and limitations of this study

- This systematic review synthesises the growing research applying nudges in healthcare contexts to improve clinical decision-making.
- The review uses both systematic search strategies and a snowball sampling approach, the latter of which is useful for identifying relatively novel literature.
- Meta-analysis was not possible due to heterogeneity in methods and outcomes.
- The systematic review was not designed to synthesise research wherein study authors did not identify the intervention as a nudge.

RATIONALE

Research from economics, cognitive science and social psychology have converged on the finding that human rationality is 'bounded'.[1] The intractability of certain decision problems, constraints on human cognition and scarcity of time and resources lead individuals to employ mental shortcuts to make decisions. These mental shortcuts, often called heuristics, are strategies that overlook certain information in a problem with the goal of making decisions more quickly than more deliberative methods.[2] While heuristics can often be more accurate than more complex mental strategies, they can also lead to errors and suboptimal decisions.[2 3] Researchers have discovered interventions to harness the predictable ways in which human judgement is biased to improve decisions. These interventions, known as 'nudges,' reshape the 'choice architecture,' or the way options are presented to decision-makers, to optimise choices.[4] Nudges have been applied to retirement savings, organ donation, consumer health and wellness, and climate catastrophe mitigation demonstrating robust effects.[5–8]

As with retirement savings and dietary choices, clinical decision-making—clinicians' process of determining the best strategy to prevent and intervene on clinical matters—is complex and error-prone. Clinicians often use heuristics when making diagnostic and treatment decisions.[9–11] For example, clinicians are influenced by whether treatment outcomes are framed as losses or gains (eg, doctors prefer a riskier treatment when the outcome is framed in terms of lives lost rather

than lives saved).[12] Heuristics can lead to medical errors.[13] In the face of complex medical decisions, clinicians tend to choose the default treatment option (despite clinical guidelines) or conduct clinical examinations that confirm their prior beliefs.[14 15]

Choice architecture influences clinicians' behaviour regardless of whether clinicians are conscious of it, creating opportunities for nudges.[16] Clinical decisions are increasingly made within digital environments such as electronic health record (EHR) systems.[17] More than 90% of US hospitals now use an EHR.[18 19] Researchers have explored the potential to use these ubiquitous electronic support systems to shape clinical decisions through nudges. They have subtly modified the EHR choice architecture by changing the default options for opioid prescription quantities or by requiring physicians to provide free-text justifications for antibiotic prescriptions.[16] Even when nudges are not implemented in the EHR, researchers extract aggregate data from the EHR, suggesting its increasing role in the study of clinical decision-making.[20]

As health systems and researchers have embraced nudges in recent years, there is growing interest in understanding which nudges are most effective to improve clinical decision-making. Taxonomising nudges is advantageous because many nudges explicitly target heuristics, revealing the mechanism of behaviour change.[21] If nudges that leverage people's tendency to adhere to social norms are consistently more effective than nudges that harness clinicians' default bias, then future nudges can be designed with this insight. Two systematic reviews were recently conducted to evaluate the effectiveness of healthcare nudges. Though both reviews demonstrate promise for the effectiveness of nudges, they offer somewhat conflicting evidence on the most studied and most effective nudge types, suggesting that an additional review may be useful.[22 23] Our review offers complementary and non-overlapping insights on the study of nudges in healthcare settings for the following reasons: (1) we do not exclusively study physicians as our target population,[23] instead we include all healthcare workers; and (2) we do not restrict our research to randomised controlled trials (RCTs) reported in the Cochrane Library of systematic reviews.[22]

Our review also makes use of a nudge taxonomy derived from the widely cited Nuffield Council on Bioethics intervention ladder wherein interventions increase in potency and constrain choice with each new rung.[24 25] Interventions on the bottom of the ladder tend to be more passive, offering decision-makers information and reminders. Interventions in the middle of the ladder leverage psychological insights to motivate decision-makers either through social influence or by encouraging planning. At the top of the ladder, interventions are more assertive and reduce decisions to a limited set of choices or by creating default options. The nudge ladder categorises nudges by the psychological mechanisms by which they operate, the degree to which they maintain autonomy and have the additional advantage of aligning with existing public health and quality improvement literature that make use of the Nuffield Council ladder.[4 26] The nudge ladder offers insights on the heuristics most relevant to the clinical decision-making process and can support health systems in selecting and applying nudges to improve clinical decision-making.

Objective

We systematically evaluated nudge interventions directed at clinicians in healthcare settings to determine the types of nudges that are most studied and most effective in improving clinical decision-making compared with other nudges, non-nudge interventions or usual care. All quantitative study designs were included in our review.

METHODS

Protocol and registration

Before initiating this review, we searched the international database PROSPERO to avoid duplication. After establishing that no such review was underway, we prospectively registered our review (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=123349).

Eligibility criteria

Types of participants

We included only empirical studies published in peer-reviewed journals studying nudges directed at clinicians working in healthcare settings. Clinicians were defined as workers who provide healthcare to patients in a hospital, skilled nursing facility or clinic. Examples of clinicians include physicians, nurses, medical assistants, physician assistants, clinical psychologists, clinical social workers and lay health workers. Studies that exclusively nudged patients were not included.

Types of intervention

Nudges were defined as 'any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives'.[4] Alterations to choice architecture included changes to the information provided to the clinician (eg, translating information, displaying information, presenting social benchmarks), altering the decision structure of the provider (eg, modifying default options, changing choice-related effort, changing the number or types of options or changing decision consequences) and providing decision aids (eg, offering reminders or commitment devices).[27] The study authors did not need to identify the intervention as a nudge to be considered for study inclusion, however given the systematic search string, which includes several behavioural economics terms (see online supplemental appendix 1), studies that did not self-identify as behavioural economic interventions were unlikely to be included.

Interventions that required sustained education or training were not considered nudges. No options could

Table 1 Eligibility criteria

Inclusion criteria	Full-text empirical journal articles.
	English language.
	Published in a peer-reviewed journal.
	The studies in the paper empirically investigated one or more behavioural intervention techniques that were considered nudges or were connected to the choice architecture literature by the original authors. These interventions are all clinician-directed (eg, nurses, doctors, residents, medical assistants), not patient-directed.
	The studies in the paper had behavioural outcome variables, not preferences or attitudes (eg, prescribing behaviour).
Exclusion criteria	Abstracts unavailable in the first-pass screen.
	Review articles, conference abstracts, textbooks, chapters and conference papers.
	Studies without a control group or baseline comparator.
	The studies in the paper applied interventions that restrict the freedom of choice of the target population, included significant economic incentives, ongoing education, complex decision support systems or consultation.

be forbidden and there could be no financial incentives.[28] Though some financial incentives for clinicians may be considered nudges, most studies on financial incentives for clinicians involve significant compensation or 'pay for performance'—of which there is already an existing literature.[29]

Nudges guided clinicians to make improved clinical decisions, including (but not limited to) increasing the uptake of evidence-based practices (EBPs), adherence to health system or policy guidelines and reducing healthcare service costs. EBPs refer to clinical techniques and interventions that integrate the best available research evidence, clinical expertise and patient preferences and characteristics.[30] Study authors had to provide the evidentiary rationale for the nudge.

We did not include studies that analysed the sustainability of nudges in the same study setting and/or sample of providers. In order to analyse studies with independent samples, we included the primary paper and not follow-up papers.

Types of studies

All study designs were included that had a control or baseline comparator—the control or baseline could be usual care or another intervention (nudge or non-nudge). For studies with parallel intervention groups, we did not require that allocation of interventions be randomised (ie, quasi-experimental studies were included). Exclusively qualitative studies were not included. See table 1 for eligibility criteria.

Search

Snowball sampling

The initial search strategy was based on a snowball sampling method[31] using the references from a published commentary on the uses of nudges in healthcare contexts.[16] Reviews identified during the preliminary stage of the systematic search process were also used to snowball articles, though these largely resulted in duplicates. Articles were reviewed at the title level to immediately identify those to be excluded. Those tentatively included were reviewed at the abstract level, followed by the full text for those meeting criteria. Following completion of screening of records retrieved via snowball, a systematic search of several databases was completed.

The methodology for the search was designed based on standards for systematic reviews,[32] in consultation with a medical librarian, as well as with two experts from the field of healthcare behavioural economics. The databases used were: EconLit, Embase, EBSCO MegaFILE, PsycINFO, PubMed, Scopus and Web of Science.

Search terms included combinations, plurals and various conjugations of the words relating to identified nudge interventions. The search string and strategy from[6] was used as a basis for search terms, but adjusted to reflect our research question (see table 1). All peer-reviewed empirical studies published prior to the completion of our search phase (ie, April 2020) were eligible for this review. See online supplemental appendix 1 for the search strings.

Data collection process

Following retrieval of all records, duplicates were removed using Zotero (www.zotero.org) and via manual inspection. Article screening involved two stages. First, all records were screened at the title and abstract level by a team of four coders (BSL, CT and two research assistants) using the web-based application for systematic reviews, Rayyan (https://rayyan.qcri.org). Criteria in this first-pass screening were inclusive—that is, all interventions directed at clinicians were included. To establish reliability, the coders screened the same 20 articles and then reviewed their screening decisions together. Any disagreements were resolved by consensus. This process was repeated three additional times until 80 articles were screened by all four coders and sufficient reliability was established. Reliability was excellent (Fleiss' κ=0.96). For the remainder of the screening process, screening was done independently by all four coders; the team met weekly to discuss edge cases. This screening process was followed by a full-text examination to determine eligibility

according to more stringent inclusion and exclusion criteria (see table 1). This screening process was done as a team and determinations of article inclusion were decided by consensus.

Patient and public involvement

Patients and the public were not involved in the design, conduct or reporting of this research.

Data items

Study characteristics and outcomes were extracted and tabulated systematically per recommendations for systematic reviews.[32] These data included: (1) study characteristics—author names, healthcare setting, study design, country, date of publication, details of the intervention, justification for the nudge, sample size, primary outcomes, main findings and whether the effect was statistically significant; (2) nudge type; and (3) risk of bias assessment.

BSL and RSB trained the coding team (four Master's students in a Behavioural and Decision Sciences programme) in data extraction. The team coded articles (n=16) together to ensure consensus. RSB reviewed a random sample (n=5) of the final articles to ensure reliability with systematic review reporting standards. BSL subsequently coded the remaining articles (n=18).

Outcomes

We only included studies that included objective measures of clinician behaviour in real healthcare contexts. Studies that measured clinicians' choices in vignette or simulation studies were not included. Results could be presented as either continuous (eg, number of opioid pills prescribed) or binary (eg, whether physicians ordered influenza vaccinations). Outcomes were measured either directly (eg, antibiotic prescribing rates) or indirectly (eg, using cost to estimate changes in antibiotic prescriptions). Participants could not report on their own behaviour because clinicians' self-report can be inaccurate.[33] Both absolute measurements and change relative to baseline were accepted.

Risk of bias in individual studies

We evaluated whether the studies included in the systematic review were at risk for bias, using the Cochrane Risk of Bias Tool.[32 34] BSL trained CT and they assessed articles (n=2) together to ensure consensus. CT independently coded (n=12) articles and BSL coded the remaining articles (n=27). The team met weekly to discuss articles that they were uncertain about and resolved discrepancies by consensus.

Data synthesis

In order to examine which types of nudges were most studied and most effective, we calculated the number

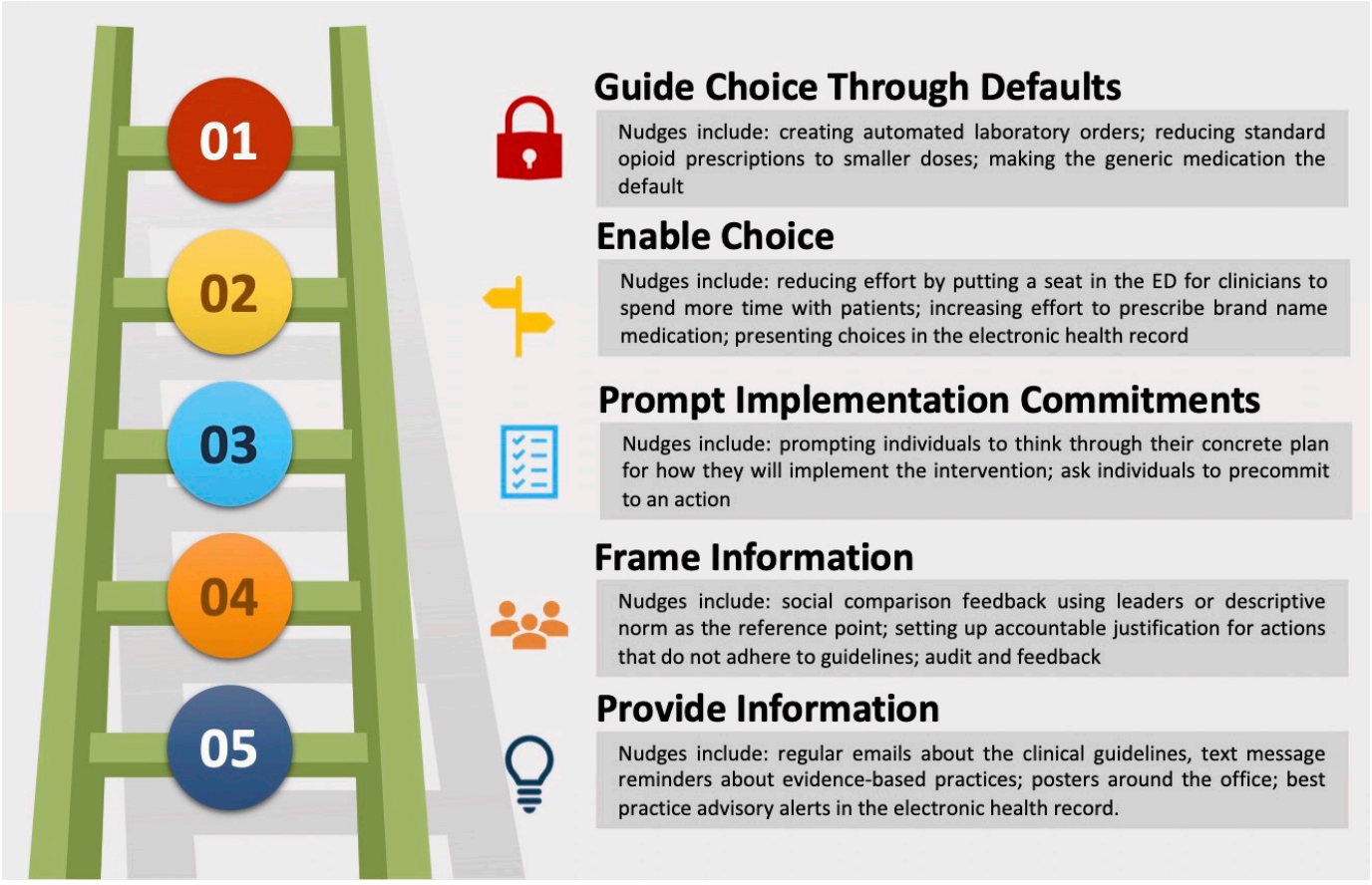

Figure 1 Ladder of nudge interventions. Note, ladder adapted from 24 25. ED, emergency department.

and percentage of studies using each nudge intervention according to the nudge ladder (see figure 1). We reported the effect and statistical significance of the effect when a primary outcome was clearly identified in the study. If no primary outcome was identified by study authors, we determined a primary outcome based on the main research question. For studies that reported multicomponent nudges—ie, interventions that combine several nudges together—we reported the total effect of the intervention. For multicomponent nudge interventions, we coded them according to the nudge ladder with all of the nudge types that apply. For studies with multiple nudge treatment groups, we reported the effect of each treatment arm separately. Only nudge interventions were compared with the control arms.

Due to the differences in the exposure, behavioural outcomes and study designs interventions could not be directly compared with one another quantitively using effect sizes.[35] Hence, meta-analysis of nudge effects was infeasible. To synthesise the results, we used a vote counting method based on the direction and significance of the effect for each study; caution when interpreting results based on statistical significance is warranted.[32] If a simple majority of nudges were significant in a nudge category, the category was deemed effective.

RESULTS

Study selection

The systematic database search identified 3586 entries, which were combined with another 22 articles of interest identified by the snowball sampling method, totaling 3608 articles (see online supplemental appendix 1 for yield). After deduplication of records from the respective databases and snowball sampling techniques, 2486 article records remained. Of the 2486 articles, 2486 articles from the systematic search and snowball method were retrievable and screened in the first stage of title and abstract screening, which reduced the total number of full-text screens to 133 unique articles. Of the 133 articles that were full-text screened, 39 articles[20 36–73] met inclusion criteria and the data from these were extracted and evaluated in this review (see Preferred Reporting Items for Systematic Reviews and Meta-Analyses diagram in figure 2).

Study characteristics

The characteristics of the included studies are summarised in table 2. The majority (n=35, 90%) of studies were conducted in the USA; two (5%) were conducted in the UK, one (3%) in Belgium and one (3%) in Switzerland. Studies were set in a variety of healthcare contexts (eg, outpatient clinics, primary care practices, emergency

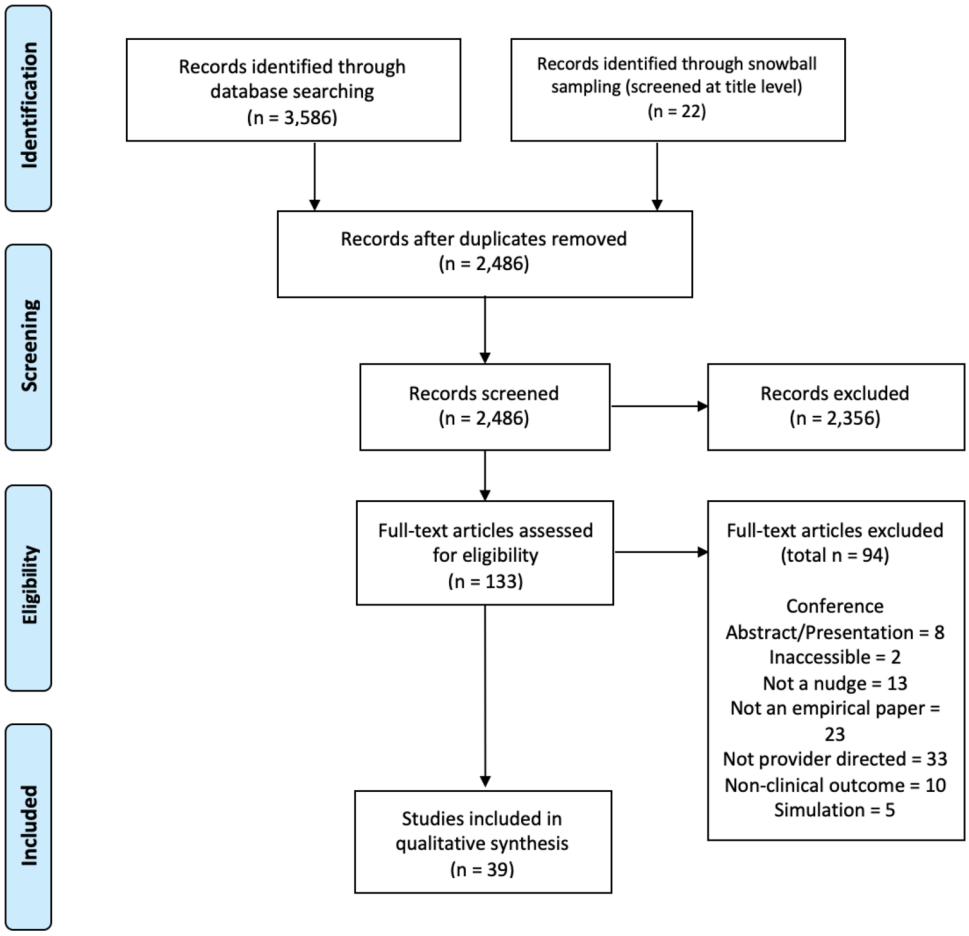

Figure 2 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram.

Table 2 Study characteristics

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Allen *et al*[36] (2019, USA)	Health system (16 community hospitals across 8 counties).	Prospective, pre–post design.	Quarterly peer comparison reports were sent to eligible prescribers (by email, fax or in-person). Eligible prescribers (who accounted for 75%–80% of total prescribed 'antibiotic days') were unaware they were high-volume antibiotic prescribers.	Reduce antibiotic prescriptions of fluoroquinolones due to their broad spectrum of activity, known adverse event profile and availability of other less toxic therapeutic options.	Internal medicine; hospitalists; family medicine (n=189). Critical care; pulmonology (n=67). Infectious diseases (n=60).	Primary study outcome was fluoroquinolone days of therapy/1000 patient days (DOT/1000 PD). A day of therapy was defined as at least one dose of a fluoroquinolone in a 24-hour period, per each facility's medication administration records.	Antibiotic use declined 29% (baseline: 83.9 DOT/1000 PD, range: 59.3–118.7; intervention: 58.3 DOT/1000 PD, range: 37.1–76.7). Primary outcome (fluoroquinolone DOT/1000 PD) declined for all facilities included in the study.	p<0.001
Andereck *et al*[37] (2019, USA)	Large urban academic emergency department (ED).	Prospective pre–post design (quality improvement initiative).	Quarterly feedback by email. Prescribers could compare their rates to peers on a de-identified chart of their peers. Formal education and training complimented the peer intervention (eg, a brief 'pharmacy fact ' with each email and a pharmacist lecture).	Unnecessary prescribing patterns have contributed to the opioid epidemic.	Pre-intervention period, 35636 ED visits were discharged. M=44 attending physicians, 30 senior resident physicians and 33 junior resident physicians and advanced practice providers per block met inclusion. Post-intervention period, a total of 18830 ED visits were discharged. M=40 attending physicians, 30 senior residents, and 35 junior residents and advanced practice providers per block met inclusion threshold.	The primary outcome of this evaluation was the overall ED discharge opioid prescribing rate. Prescribing rate was defined as the proportion of discharged patient encounters with an opioid prescription for the department in a specific scheduling block.	Departmental opioid prescribing rates during the evaluation period declined; pre-intervention period rate: 8.6% (95% CI: 8.3% to 8.9%) versus post-intervention period rate: 5.8% (95% CI: 5.5% to 6.1%).	p<0.01
Arora *et al*[38] (2019, USA)	Two general medicine inpatient units.	Prospective, cross-sectional pre–post design.	Changing the electronic health record (EHR), creating a default to monitor patient's vital signs; customised office signs for nurses educating them about best 'sleep-friendly' vitals monitoring practices; pocket-cards with information; 20 min education session.	Sleep is important for patient recovery but patients struggle to sleep in hospitals, which is related to poor outcomes.	n=? providers. 1083 general-medicine patients, 1669 EHR general medicine orders.	Changes in the mean percentage of 'sleep-friendly' (ie, non-nocturnal) orders for checking vital signs and venous thromboembolism prophylaxis compared with baseline.	Increases in the mean percentage of sleep-friendly orders rose for both: no vital sign: 3% to 22%, sleep-promoting venous thromboembolism prophylaxis: 12% to 28%.	p<0.001

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Bourdeaux *et al*[39] (2014, UK)	Inpatient intensive care unit.	Retrospective pre–post design.	Prescription template with preprescribed drugs and fluids doctors choose to use the template on admission.	Chlorhexidine mouthwash reduces ventilator associated pneumonia in critically ill patients. It is cheap and acceptable. Hydroxyethyl starch (HES) is an intravenous fluid that helps circulation.	n=? providers. 2231 ventilated patients were eligible for chlorhexidine, 591 pre-intervention and 1640 post-intervention. 6199 patients were eligible for HES intervention, 2177, pre and 4022 post.	Changes in the delivery of chlorhexidine mouthwash and HES to patients in the intensive care unit.	Percentage of patients prescribed chlorhexidine increased (from 55.3% to 90.4%). The mean volume of HES infused per patient fell and the percentage of patients receiving HES fell (from 54.1% to 3.1%).	1. p<0.001 2. p<0.001
Buntinx *et al*[40] (1993, Belgium)	Department of pathology.	Randomised controlled trial (RCT).	Interventions, four groups. Some arms had feedback and then advice. One arm had peer. comparison.	Cervical screening can help prevent cancer.	183 doctors.	Percentage of smears lacking endocervical cells.	Smears lacking endocervical cells decreased in the groups receiving monthly peer comparison overviews compared with groups not receiving this type of feedback. OR=0.75, 95% CI (0.58 to 0.96).	p<0.05
Chiu *et al*[41] (2018, USA)	Health system (five hospitals).	Prospective pre–post design.	Changing the EHR, lowered the default number of pills on electronic opioid prescriptions from 30 to 12 after procedure.	Postprocedural analgesia prescriptions have contributed to the opioid epidemic.	n=? providers. 1447 procedures before default change and 1463 procedures after the default change.	Changes in the number of opioid pills prescribed per operation.	Decreases in the number of opioid pills prescribed −5.22 (95 CI −6.12 to −4.32).	p<0.01
Delgado *et al*[42] (2018, USA)	Two emergency departments.	Prospective pre–post design.	Changing the EHR, lowered the default number of pills on electronic opioid prescriptions to 10 pills.	Reliance on prescription opioids for postprocedural analgesia has contributed to the opioid epidemic.	n=? providers. 3264 prescriptions were written across the two EDs.	Increase in 10 pill prescriptions relative to control 4 weeks after implementation; changes in the mean number of oxy/APAP tablets prescribed per week.	Increase in proportion of prescriptions for 10 tablets 27.8%, 95% CI 17.4% to 37.5%. No change in the mean number of oxy/APAP tablets prescribed per week.	p<0.001
Hemkens *et al*[43] (2017, Switzerland)	Nationwide.	Pragmatic RCT.	Personalised antibiotic prescription feedback by mail and an online dashboard and a letter on antibiotic prescribing guidelines.	Clinicians often inappropriately prescribe antibiotics for acute respiratory tract infections.	2900 primary care physicians.	Changes in defined daily doses of any antibiotic to any patient per 100 consultations in first year, intention-to-treat, relative to control.	No change in prescribing behaviour: between-group difference, 0.81%; 95% CI −2.56% to 4.30%.	N.S.
Hempel *et al*[44] (2014, USA)	Emergency department.	Prospective pre–post design.	Peer comparison feedback on emergency medicine resident ultrasound scan numbers.	Clinician-performed ultrasounds are part of emergency medicine residency curricula; there is a need for effective teaching.	44 emergency medicine residents.	Changes in number of scans done per shift in the 3 months after intervention (relative to baseline).	Increase in number of scans performed (number of ultrasound exams per shift increased from 0.39 scans/shift to 0.61 scans/shift).	p<0.05

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Hsiang *et al*[45] (2019, USA)	Health system (25 primary care practices).	Retrospective difference-in-differences approach (intervention vs control practices during post-intervention year compared with the two pre-intervention years).	Active choice of a best-practice alert for medical assistants. During vitals check, the EHR) prompted medical assistants to accept/cancel a cancer screening order. If accepted, a pending order was made for the clinician to review and sign during the patient visit.	US Preventive Services Task Force guidelines for breast and colorectal cancer screening.	n=? providers. 26269 women eligible for breast cancer screening, 43647 men eligible for colorectal cancer screening.	Primary outcome was ordering of the screening test during a visit (primary care) compared with control groups relative to two pre-intervention years.	Breast cancer screening tests (22.2% point increase, 95% CI 17.2% to 27.6%) and colorectal cancer screening test increased (13.7% point increase, 95% CI 8% to 18.9%).	p<0.001
Kim *et al*[46] (2018, USA)	11 primary care practices.	Prospective, cross-sectional pre–post design (differences-in-differences).	Changing the EHR, an 'active choice' intervention using a best practice alert directed to medical assistants—prompt to accept or cancel a influenza vaccine order. If accepted, the order was made for the physician to review and sign during the patient visit.	Center for Disease Control recommends universal influenza vaccination.	n=? providers 96, 291 patients.	Changes in influenza vaccination rates compared with control practices over time.	Increase in influenza vaccination rates (9.5% point increase in vaccination rates (95% CI 4.1% to 14.3%).	p<0.001
Kullgren *et al*[47] (2018, USA)	Six adult primary care practices.	12-month stepped wedge cluster RCT, randomisation by clinic.	Clinicians pre-commited to 'Choosing Wisely' choices against low-value orders. They received 1–6 months of point-of-care pre-commitment reminders, patient education handouts and weekly emails.	Clinicians often order costly and inappropriate tests as well as inappropriately prescribe antibiotics for acute respiratory tract infections.	45 primary care physicians and advanced practice providers.	Primary outcome was the difference between control and intervention period percentages of visits with potentially low-value orders.	No change in in the percentage of visits with potentially low-value orders overall, for headaches or for acute sinusitis (–1.4%, 95% CI –2.9% to 0.1%).	N.S.
Lewis *et al*[48] (2019, UK)	Acute medical hospital.	Controlled interrupted time series design.	Message at the bottom of all inpatient and outpatient paper and electronic CT reports, highlighting patients at risk after exposure to ionising radiation and asks the provider if they informed the patient.	CT scans are known to expose individuals to radiation, which can increase cancer risk.	n=? providers.	Immediate change in level or a gradual trend change in CT counts in electronic reports compared with control hospital.	Significant reduction in CT scans (–4.6%, 95% CI (–7.4% to –1.7%).	p=0.002

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Meeker *et al*[49] (2014, USA)	5 primary care clinics.	RCT, randomisation by clinician.	Poster-sized commitment letters in clinicians' personal examination rooms for 12 weeks. These letters displayed clinician photographs, signatures and commitment to not inappropriately prescribe antibiotics for acute respiratory infections.	Clinicians often inappropriately prescribe antibiotics for acute respiratory tract infections despite guidelines and several clinical interventions.	14 clinicians (11 physicians and 3 nurse practitioners) 954 eligible adult patients.	Differences in antibiotic prescribing rates for antibiotic-inappropriate acute respiratory infection diagnoses at baseline and during intervention periods.	Decrease in inappropriate antibiotic prescribing rate compared with control (difference in difference −19.7%, 95% CI (−5.8% to −33.4%).	p<0.05
Meeker *et al*[50] (2016, USA)	47 primary care practices in two different health systems.	2×2×2 factorial RCT (practices received 0, 1, 2 or 3 interventions).	1. Changes in EHR, 'suggested alternatives' presented electronic order sets with non-antibiotic treatments. 2. Changes in EHR, 'accountable justification' clinicians enter free-text justifications for prescribing antibiotics. 3. Peer comparison emails about how clinicians' antibiotic prescribing rates compare to lowest inappropriate prescribers.	Clinicians often inappropriately prescribe antibiotics for acute respiratory tract infections.	248 clinicians (14, 753 visits at baseline and 16, 959 during intervention period).	Changes in rates of inappropriate antibiotic prescribing behaviour compared with baseline.	1. No significant change in inappropriate antibiotic prescriptions; difference in difference: −5%, 95% CI (−7.8% to 0.1%). 2. Decrease in inappropriate antibiotic prescriptions; difference in difference: −7%, 95% CI(−9.1% to −2.9%). 3. Decrease in inappropriate antibiotic prescriptions; difference-in-difference: −5.2%, 95 CI (−6.9% to −1.6%).	1. NS; 2. p<0.001; 3. p<0.001. No statistically significant interactions between interventions.
Nguyen and Davis[51] (2019, USA)	One multispecialty academic medical centre.	Single centre, prospective, quasi-experimental pre–post design.	Peer comparison reports of the percentage of appropriately verified vancomycin orders for each pharmacist. In phase I, reports were blinded. In phase II, reports were unblinded. Intervention phases were compared with a pre-intervention control.	Pharmacist 'order verification' prevents medical errors, which are harmful to patients. Vancomycin is a commonly prescribed drug for hospitalised patients.	n=? providers. 1625 vancomycin orders were included for evaluation (537 orders in the control group, 549 orders in intervention phase I and 539 orders in intervention phase II).	Appropriate vancomycin dose order verification, appropriate dose was determined by the institution's guidelines.	Appropriately verified vancomycin orders significantly increased in the phase II (unblinded) compared with the control group (OR=1.79; 95% CI (1.36 to 2.34).	p<0.001

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
O'Reilly-Shah *et al*[52] (2018, USA)	Department of anaesthesiology in a large health system (two academic hospitals, two private practice hospitals and two academic surgery centres).	Retrospective pre–post design (stepwise cluster implementation in five facilities).	1. Audit and feedback on provider level and department-level compliance with lung-protective ventilation for attending physicians. 2. Audit and feedback for advance practice providers and residents. 3. Changes to the EHR, default setting on anaesthesia machines for tidal volume was decreased from 700 mL to 400 mL.	There is a need to improve compliance with anaesthesiology surgical quality metrics.	n=? providers, Total surgical case count (n)=14 793 unique patients (n)=12 785. Five facilities.	Rates of compliance with low tidal wave ventilation compared with baseline.	Attending physician dashboards increased compliance odds 41% (OR 1.41, 95% CI 1.17 to 1.69). Adding advanced practice provider and resident dashboards increased compliance odds 93% (OR 1.93, 95% CI 1.52 to 2.46). Changing ventilator defaults led to 376% increase in compliance odds OR 3.76, 95% CI 3.1 to 4.57.	1. p=0.002 2. p<0.001 3. p<0.001
Olson *et al*[53] (2015, USA)	Clinical pathology, haematology, and oncology departments in a health system.	Prospective pre–post design (multiple baseline).	Changes in the EHR default order sets for post-transfusion haematocrits and platelet counts changed from 'optional' to 'preselected.' Platelet count default settings later changed back to 'optional'.	Need to improve the monitoring of post-transfusion outcomes.	>500 residents and fellows. 7578 orders for red blood cell transfusion, 3285 total orders for platelet transfusion.	Rates of laboratory test ordering for post-transfusion counts after default change and post default change.	Increase in haematocrit and platelet post-transfusion count orders after default for order was set to 'pre-selected' (8.3% to 57.5% change). After switch back to 'optional', significant decrease in orders.	p<0.001
Orloski *et al*[54] (2019, USA)	2 urban, academic emergency departments.	Prospective, controlled pre–post trial.	Placed institution-branded folding seats in the ED and an educational campaign on good communication. Only the intervention ED received folding seats.	Patient satisfaction is important.	n=? providers. 2827 patients were surveyed.	Primary outcome was the impact of provider sitting on patient satisfaction. Secondary outcome was provider sitting frequency.	Sitting at any point during an ED encounter increased patient satisfaction across all measures (polite: 67% vs 59%, cared: 64% vs 54%, listened: 60% vs 52%, informed: 57% vs 47%, time: 56% vs 45%. Odds of provider sitting increased 30% when a seat was in the room, OR=1.3, 95% CI (1.1 to 1.5).	p<0.0001

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Parrino[55] (1989, USA)	One tertiary referral hospital.	Prospective pre–post design.	Monthly peer comparison letters sent to two groups (surgical and nonsurgical physicians) who were in the top 50 percentiles of prescribers for antibiotic expenditures.	Antibiotics are often inappropriately prescribed and can be expensive.	202 physicians, surgical (n=83) and non-surgical (n=119).	Changes in expenditures (total dollars) on antibiotics per physician (mean difference from quarter 3 to quarter 4 compared with control group before and after feedback).	No significant change in total dollars spent on antibiotics (mean difference: $797.50 vs $1355.33).	N.S.
Patel *et al*[58] (2014, USA)	One general internal medicine and one family medicine practice.	Retrospective cross-sectional pre–post design.	Modify EHR default from showing brand and generic medications to displaying only generics at first, with the ability to opt out.	Generic medications are less expensive than brand-name medications and are of comparable quality.	Internal medicine (IM) and family medicine (FM) attending physicians (IM, n=38; FM, n=17) and residents (IM, n=166; FM, n=34).	Monthly prescriptions of brand-name and generic equivalent for: beta-blockers, statins and proton-pump inhibitors compared with control.	Increase in generic prescribing behaviour for all three medications; 5.4% points, 95% CI, (2.2% to 8.7%).	p<0.001
Patel *et al*[57] (2016, USA)	Three internal medicine practices.	Prospective cross-sectional pre–post design (difference in differences).	Changing the EHR through 'active choice' using a best practice alert for medical assistants and physicians, prompting them to accept/cancel an order for a colonoscopy, mammography or both. Physician needed to review and sign order during visit.	Guidelines suggest that increasing early cancer detection can be done through regular screening practices.	n=? providers, One intervention practice, two controls. 7560 patients eligible for colonoscopy with 14546 clinic visits and 8337 patients eligible for mammography with 14410 clinic visits.	Percentage of patients eligible for screening who received a cancer screening order.	Increase in mammography (12.4% points, 95% CI 8.7% to 16.2%) and colonoscopy orders (11.8% points, 95% CI 8% to 15.6%).	p<0.001
Patel *et al*[73] (2016, USA)	All specialties across a health system.	Pre–post design, difference-in-differences approach.	'Active choice' in the EHR. An opt-out 'checkbox' that said 'dispense as written' was added to the prescription EHR screen, and if unchecked the drug's generic version was prescribed.	Generic medications are linked to higher adherence to medication regimens and better clinical outcomes.	n=? providers. Pre-intervention data: 811561 eligible prescription sets during 10-month pre-intervention period to 655011 prescriptions during 7-month post-intervention period.	Generic prescribing rates for 10 medical conditions, ie, 10 drugs.	The overall generic prescribing rate increased significantly (75.3% to 98.4%).	p<0.001

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Patel *et al*[56] (2017, USA)	Three internal medicine practices.	Prospective cross-sectional, pre–post design (difference-in-differences).	Changing the EHR through 'active choice' using a best practice alert directed to medical assistants and physicians—prompting to accept/cancel an order for the influenza vaccine. Physician needed to review and sign during the patient visit.	The Center for Disease Control recommends universal influenza vaccination.	n=? providers, One intervention practice, two control practices. 45 926 patients.	Changes in influenza vaccination rates.	Increase in vaccination rates (adjusted difference-in-difference: 6.6% points; 95% CI 5.1% to 8.1%).	p<0.001
Patel *et al*[59] (2018, USA)	One health system, 32 primary care practices (PCPs).	Three-arm cluster randomised clinical trial.	1. 'Active choice' and 'accountable justification'. Physicians received an email with number of eligible patients for statin therapy who had not been prescribed a statin and were asked to actively choose to prescribe atorvastatin, 20 mg, once daily, atorvastatin at another dose or another statin or not prescribe a statin and describe a reason. 2. Active choice and accountable justification and peer comparison emails describing how physicians compared with peers.	50% of eligible patients do not receive statins despite evidence of their efficacy.	96 PCPs 4774 patients eligible but not receiving statin therapy.	Percentage of eligible patients receiving statin prescription orders compared with usual care.	1. No significant increase in statin prescription rates versus usual care (adjusted difference: 4.1%, 95% CI −0.8% to 13.1%). 2. Increase in statin prescription compared with usual care (adjusted difference, 5.8%; 95% CI 0.9% to 13.5%).	1. N.S. 2. p<0.01

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Persell *et al*[60] (2016, USA)	General internal medicine clinic.	2×2×2 factorial RCT with three interventions.	1. 'Accountable justification' in EHR. Physicians received an alert when inappropriately prescribing an antibiotic and provided free-text justification. 2. 'Suggested alternatives' in EHR when physicians inappropriately prescribe antibiotics. 3. Peer comparison monthly performance feedback compared with lowest 10% of inappropriate prescribers.	Clinicians frequently prescribe antibiotics inappropriately for acute respiratory infections.	n=? providers. 3276 visits in the pre-intervention year and 3099 visits in the intervention year.	Rate of oral inappropriate antibiotic prescriptions for acute respiratory infection diagnoses compared with control group and baseline.	No significant decrease in inappropriate prescribing rates compared with control group. Significant decrease in inappropriate prescribing across all groups (including controls) compared with baseline: 1. OR=0.98, 95% CI (0.42 to 2.29). 2. OR=0.68, 95% CI (0.29 to 1.58). 3. OR=0.45, 95% CI (0.18 to 1.11).	N.S.
Ryskina *et al*[61] (2018, USA)	Six general medicine teams in one health system.	Single-blinded cluster RCT, Randomisation by 2 week service. block.	Peer comparison emails sent to physicians on general medicine teams, summarising their routine lab test orders versus the service average that week.	Routine laboratory tests for hospitalised patients can be wasteful and are overused.	Six attending physicians, 114 interns and residents.	Number of routine laboratory orders placed by each physician per patient day.	No significant changes in number of laboratory orders by each physician (–0.14 tests per patient-day vs control group, 95% CI –0.56 to 0.27).	N.S.
Sacarny *et al*[20] (2018, USA)	Highest volume primary care prescribers of quetiapine in 2013 and 2014, whose patients have Medicare.	RCT (intent to treat) placebo-control parallel-group design, balanced randomisation (1:1) to control group (placebo letter) and treatment group (peer comparison letter).	Mailed peer comparison letters saying that prescriber's quetiapine prescribing was under review and was high relative to same-state peers, which was concerning and could be medically unjustified.	Antipsychotic agents like quetiapine fumarate are often overprescribed when not clinically indicated/supported with the potential to cause patient harm.	5055 PCPs, 231 general practitioners, 2428 were in family medicine, and 2396 were in internal medicine.	Total quetiapine days prescribed by physicians from the intervention start to 9 months in intervention versus control.	Decrease in quetiapine days per prescriber in treatment versus control arm; –11.1%, 95% CI (–13.1% to –9.2%).	p<0.001

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Sedrak *et al*[62] (2017, USA)	Three hospitals in one health system.	RCT comparing a 1-year nudge to a 1-year pre-nudge period, accounting for time and patient features. Randomisation at test-level.	Intervention laboratory tests showed Medicare allowable fees at the time of order in the EHR and control laboratory tests did not show prices.	A significant number (30%) of laboratory tests in the USA may be wasteful. Increasing price transparency at the time of laboratory order entry may influence provider decisions and decrease wasteful tests.	n=? providers. 60 diagnostic laboratory tests, 30 most frequently ordered and 30 most expensive. 142, 921 hospital admissions, 98529 patients.	Frequency of tests ordered per patient-day. Secondary outcome was the number of tests done per patient-day and the Medicare fees.	No significant changes in number of tests ordered between intervention and control group (0.05 tests ordered per patient-day; 95% CI −0.002 to 0.09).	N.S.
Sharma *et al*[63] (2019, USA)	One health system, 5 radiation oncology practices.	Stepped-wedge cluster randomised clinical trial.	Change EHR through a default imaging order for no daily imaging during palliative radiotherapy, which physicians could opt-out from by specifying another imaging frequency.	Guidelines suggest that imaging using radiography or CT on a daily basis is unnecessary for patients undergoing palliative radiotherapy. Daily imaging can be costly and increase treatment duration for patients.	21 radiation oncologists 1019 patients who received 1188 palliative radiotherapy courses (n=747 at university practice; n=441 at community practices) to bone, soft tissue, brain or various sites.	Primary outcome was binary outcome (whether radiotherapy courses with daily imaging were ordered). Daily imaging course was defined as imaging during ≥80% of palliative therapy treatments.	Default led to a significant reduction in daily imaging adjusted OR=0.43; 95% CI 0.24 to 0.77; adjusted difference in % points, −18.6; 95% CI, −34.1 to −2.1.	p=0.004
Shively *et al*[64] (2020, USA)	Veterans' Affairs Health System (seven primary care practices).	Prospective pre–post design.	Peer comparison feedback—an educational session for all primary care providers and monthly emails with their antibiotic prescribing rate, their colleague's rates and the system's goal rates.	Clinicians frequently inappropriately prescribe antibiotics despite guidelines.	Baseline=65 primary care professionals (PCPs) serving 40734 patients, 28402 office visits. Intervention=73 PCPs serving 41191 patients, 32982 office visits.	Monthly mean rate of antibiotic prescribing rates. Secondary outcomes were inappropriate antibiotic prescribing rates and appropriate antibiotic prescribing rates.	Mean rate of monthly antibiotic prescriptions significantly reduced 35.6%. Unnecessary antibiotic prescribing decreased 33.9% and the appropriate antibiotic rates increased 50.8%.	p<0.001
Srinivasan *et al*[65] (2020, USA)	Inpatient units in a 350-bed children's hospital.	Prospective pre–post design.	EHR reminders, provider education (including a quiz), and peer comparison feedback (how unit rates compared with other units in the hospital, shown on posters and sent by email).	American Academy of Pediatrics guidelines for universal, yearly influenza vaccination for all children 6 months and older.	n=? providers. Baseline=6089 admitted children (6 months and older) to the medical and surgical units. Intervention=6206 children admitted.	Primary outcome was percentage of children discharged with one dose (or greater) of the influenza vaccine (from the hospital or before admission).	Significant increase in the percentage of discharged children with at least 1 dose of the influenza vaccine (4.7-fold increase, from 10% to 46%).	p<0.001

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Suffoletto and Landau[66] (2019, USA)	Emergency departments in one hospital system, 16 hospitals.	A pilot RCT (randomisation by provider).	Audit and feedback (A&F) emails versus peer norm comparison (PC) emails to other emergency medicine providers at their hospital.	Opioid epidemic is still a persistent problem; need to reduce opioid prescriptions.	37 emergency medicine providers.	Mean monthly opioid prescriptions by provider.	Opioid prescriptions reduced non-significantly in both conditions (audit and feedback, and peer norm comparison). Mean reduction (SD) was 3.3. (9.6) for controls, 3.9 (10.5) in A&F and 7.3 (7.8) for A&F+PC.	N.S.
Szilagyi *et al*[67] (2014, USA)	Practices in two large research networks.	RCT, randomisation unit by practices in two practice-based research networks.	EHR prompts/ alerts at all office visits with vaccine recommendations. Reminder sheet on the provider's desk in the exam room with indicated vaccines.	Guidelines recommend adolescent immunisation for a host of diseases; yet vaccination rates are not in line with guidelines.	n=? providers. Two practice networks: one network: 5 interventions, 5 control practices; one network: 6 interventions, 6 control practices.	Changes in adolescent immunisation rates, by practice.	No significant difference in immunisation rates between intervention and control practices for any vaccine or combination of vaccines (eg, adjusted OR for Human Papillomavirus vaccine at one site: 0.96; 95% CI 0.64 to 1.34), at another: adjusted OR=1.06; 95% CI 0.68 to 1.88.	N.S.
Trent *et al*[68] (2018, USA)	One medical centre, an urban, safety net, level one trauma centre.	Stepped wedge design and cluster randomisation.	Monthly audit and feedback emails with blinded peer comparison feedback adherence to guidelines for pneumonia and severe sepsis. Physicians also received emails about patients that got non-adherent service to review.	Adherence to guidelines for pneumonia and sepsis treatment are low in emergency departments.	n=? providers. 469 patients during entire study period.	Primary outcome was guideline-adherent antibiotic choices (guidelines determined by the institution).	Adherence to antibiotic guidelines significantly increased after audit and feedback with peer comparison was introduced (adjusted OR=1.8, 95% CI: 1.01 to 3.2).	p<0.05
Wigder *et al*[69] (1999, USA)	Emergency department in a 600-bed hospital, with a level 1 trauma centre.	Prospective, pre–post design.	1. Education campaign of 'Ottawa rule'. 2. Physicians shown baseline data. 3. Audit and feedback. Knee injury patient charts put in physician mailboxes praising them for 'Ottawa rule' adherence or informing of non-adherence	Physicians over order X-rays when guidelines (ie, the 'Ottowa rule') recommend less invasive and cheaper ways for evaluating knee problems/ injuries.	27 physicians.	Primary outcome was changes in patients with knee injuries who received an X-ray study. Secondary outcome was percentage of X-ray orders with abnormal results.	Significant decrease (23%) in number of X-ray studies, increase (58.4%) in percentage of abnormal X-rays compared with baseline.	p<0.001

Continued

Table 2 Continued

Authors (year, country)	Setting	Design	Intervention	Justification	Sample size	Outcomes measured	Main findings	Significance
Winickoff *et al*[70] (1984, USA)	Department of internal medicine at one group practice.	Three interventions: Pre–post design for first two. Third intervention: RCT with crossover design (over a 1 year, crossover at 6 months).	1. Educational meeting for clinical standard. 2. Peer comparison, meeting presenting group standard adherence pre and post the educational meeting. 3. Peer comparison feedback, monthly feedback about how physicians compare to peers at practice.	Many clinicians do not follow guidelines for colorectal screening.	n=? for first two interventions 16 physicians for RCT (third intervention).	Number of stool tests completed for colorectal cancer screening across groups who received peer comparison intervention.	1. Baseline compliance = 66.8%, intervention compliance = 69.2% 2. Intervention compliance = 68.5%. 3. Increase in number of stool tests done (66.7% to 82.2% across groups).	1. N.S. 2. N.S. 3. p<0.001
Zivin *et al*[71] (2019, USA)	Two health systems.	Prospective, pre–post design.	Modify EHR default for all schedule II opioid prescriptions to 15 pills (many EHRs had 30-day defaults previously, others had no default).	The opioid epidemic; overprescription of opioids for postprocedural pain management is a problem and out of step with guidelines.	448 prescribers. 6390 opioid prescriptions.	Primary outcome was changes in the proportion of opioid prescriptions for 15 pills for high frequency prescribers.	Percentage of 15-pill prescriptions by high prescribers increased from 2.3% to 8.1% (χ^2=6.72), 15-pill opioid prescription rates increased at both sites (4.1% to 7.2% at one site, 15.9% to 37.2% at other site).	p<0.04
Zwank *et al*[72] (2017, USA)	Emergency department of a level one trauma centre.	Retrospective pre–post design.	Changing the EHR default number of pills for opioid prescriptions from 15 tablets to a number the physician had to enter themselves.	The opioid epidemic; overdose deaths due to prescriptions from opioids as analgesics.	n=? providers. 7019 eligible prescriptions.	Changes in the total opioid pill quantity per prescription.	No significant change in mean number of opioid tablets per prescription. Mean tablets dispensed increased from 15.31 (SD=5.30) tablets to 15.77 (SD=7.30).	N.S.

departments) and targeted a variety of clinical decisions (eg, opioid prescriptions, preventative cancer screening, checking vital signs of hospitalised patients). Nudges were directed at a variety of medical professionals (including physicians, nurses, medical assistants and providers with a license to prescribe medication). Many (n=20, 51%) of the studies did not report the sample size of clinicians interacting with the nudges. Instead, the studies tended to report the sample size in terms of how many patients were affected by the nudge or the number of prescription or laboratory orders under study. Fourteen (36%) studies were RCTs; 23 studies (59%) were pre–post designs; one study (3%) was a controlled interrupted time series design; and one study (3%) was a quasi-experimental randomised design. In terms of cluster RCTs, four studies (10%) were parallel cluster RCTs and three studies were stepped wedge cluster RCTs (8%). Most studies (n=32, 82%) employed a control group/comparator that consisted of usual care or no intervention. One study (3%) used a minimal educational intervention, another study (3%) examining peer comparison letters used a placebo letter and five studies (13%) employed a factorial design in which multiple combined interventions were tested against individual interventions separately.

Of the 39 studies included in the review, 48 nudges were tested. Some studies contained multiple substudies, study arms or treatment groups, which were coded and analysed separately (see table 3). Given that some interventions (n=5) were multicomponent (ie, combinations of multiple nudges) these studies were analysed separately using the nudge ladder (see table 4).

Analysing the single component nudges using the nudge ladder, 6 nudges involved guiding choice through default options (eg, changing the default opioid prescription quantity in the EHR); 9 nudges involved enabling choice (eg, electronic prompts to accept or cancel orders for influenza vaccination); 22 nudges involved framing information (eg, peer comparison letters to the clinicians in the top 50th percentile of antipsychotic prescriptions); two nudges involved prompting implementation commitments (eg, displaying clinicians' pre-commitment letters in their own examination rooms) and four nudges involved providing information (eg, an EHR reminder to clinicians when their patients were due for immunisations). Five studies involved multicomponent nudges, with four studies involving a combination of two nudges and one study involving a combination of three nudges (see table 4).

Risk of bias of included studies

Most studies were at high risk for selection bias including random sequence generation (n=25) and allocation concealment (n=25). Attrition bias was low risk based on incomplete outcome data (n=31). A large number of trials were judged as unclear for selective reporting (n=21). In terms of blinding of participants, most studies were high risk (n=25) and in terms of blinding outcome assessment, 25 studies were judged as having unclear risk of bias. Overall, 13 studies (33%) were considered low risk of bias across all criteria (see table 5).

Synthesis of results

With significance defined as ($p<0.05$), 33 of the 48 nudges (73%) significantly improved clinical decisions, suggesting that nudges are generally effective. According to the nudge ladder, all six (100%) of the nudges that involved changing the default option to guide decision-making were significantly related to clinician behaviour change in the hypothesised direction. Seven of the nine (78%) nudges that enabled choice led to significant change in clinician behaviour. Fourteen of the 22 (64%) nudges that involved framing information changed behaviour significantly, suggesting their effectiveness. One of the two (50%) nudges that prompted implementation commitments was significantly effective and the other was not. None of the four (0%) nudges that provided information to clinicians resulted in statistically significant results. The five studies (100%) that combined nudges in multicomponent interventions all led to statistically significant changes in the hypothesised direction.

Guiding choice through default options or enabling choice through an 'active opt-out' model (ie, active choice) were the most effective interventions in changing clinician behaviour. These nudges also tended to result in the largest effect sizes. Nudges that framed information—the plurality of nudges under study—tended to also change clinician behaviour. The other types of nudges were inconclusive or had more insignificant findings than significant findings. Given that it was infeasible to conduct a meta-analysis to statistically compare the nudge effects and vote-counting is subject to several methodological issues, findings should not be viewed as definitive.

DISCUSSION

Summary of evidence

This systematic review of 39 studies found that a variety of nudge interventions have been tested to improve clinical decisions. Thirty-three of the 48 (73%) clinician-directed nudges significantly improved clinical practice in the hypothesised direction. Nudges that changed default options or enabled choice were the most effective and nudges framing information for clinicians were also largely effective. Conversely, nudges that provided information to the clinician through reminders and prompting implementation commitments did not conclusively lead to significant changes in clinician behaviour.

One strength of the taxonomy organising this review is the ability to explicate why certain nudges are more effective and the mechanism by which they operate. Drawing on the nudge ladder, evidence suggests that less potent healthcare nudges lower on the ladder such as providing information and prompting commitments may be less effective than more potent nudges that are higher on the ladder such as changing the default options. This accords

Table 3 Studies organised according to nudge ladder

Nudge ladder	Study	Significant effect in the hypothesised direction?	Majority in category significant?
Provide information	Meeker *et al*[50] USA—Arm 1	N.S.	No
	Persell *et al*[60] USA— Arm 2	N.S.	
	Sedrak *et al*[62] USA	N.S.	
	Szilagyi *et al*[67] USA	N.S.	
Frame information	Allen *et al*[36] USA	p<0.001	Yes
	Andereck *et al*[37] USA	p<0.01	
	Buntinx *et al*[40] Belgium	p<0.05	
	Hemkens *et al*[43] Switzerland	N.S.	
	Hempel *et al*[44] USA	p<0.05	
	Lewis *et al*[48] UK	p=0.002	
	Meeker *et al*[50] USA– Arm 2	p<0.001	
	Meeker *et al*[50] USA– Arm 3	p<0.001	
	Nguyen and Davis[51] USA	p<0.001	
	O'Reilly-Shah *et al*[52]— Arm 1	p=0.002	
	O'Reilly-Shah *et al*[52]— Arm 2	p<0.001	
	Parrino[55] USA	N.S.	
	Persell *et al*[60] USA— Arm 1	N.S.	
	Persell *et al*[60] USA— Arm 3	N.S.	
	Ryskina *et al*[61] USA	N.S.	
	Sacarny *et al*[20] USA	p<0.001	
	Shively *et al*[64], USA	p<0.001	
	Suffoletto and Landau[66] USA	N.S.	
	Trent *et al*[68], USA	p<0.05	
	Winickoff *et al*[70] USA— Study 1	N.S.	
	Winickoff *et al*[70] USA— Study 2	N.S.	
	Winickoff *et al*[70] USA— Study 3	p<0.001	
Prompt implementation commitments	Kullgren *et al*[47] USA	N.S.	No
	Meeker *et al*[49] USA	p<0.05	
Enable choice	Bourdeaux *et al*[39] UK	p<0.001 for both	Yes
	Hsiang *et al*[45] USA	<0.001	
	Kim *et al*[46] USA	p<0.001	
	Orloski *et al*[54] USA	p<0.0001	
	Patel *et al*[73] USA	p<0.001	
	Patel *et al*[57] USA	p<0.001	
	Patel *et al*[56] USA	p<0.001	
	Patel *et al*[59] USA— Arm 1	N.S.	
	Zwank *et al*[72] USA	N.S.	
Guide choice through defaults	Chiu *et al*[41] USA	p<0.01	Yes
	Delgado *et al*[42] USA	p<0.001	
	Olson *et al*[53] USA	p<0.001	
	Patel *et al*[58] USA	p<0.001	
	Sharma *et al*[63] USA	p=0.004	
	Zivin *et al*[71] USA	p<0.04	

Articles that included multiple intervention treatment groups, studies or study arms are described.

Table 4 Multicomponent intervention studies organised according to nudge ladder

Nudge ladder	Study	Significant effect in the hypothesised direction?
Provide information + guide choice through defaults	Arora *et al*[38] USA	p<0.001
Provide information + frame information	Wigder *et al*[69] USA	p<0.001
Enable choice + frame information	Patel *et al*[59] USA— Arm 2	p<0.001
Frame information + guide choice through defaults	O'Reilly-Shah *et al*[52] USA— Arm 3	p<0.001
Provide information + frame information + enable choice	Srinivasan *et al*[65] USA	p<0.001

with nudge research in other areas outside of healthcare.[74] For example, one study comparing various types of nudges that increase the salience of information (eg, including providing reminders, leveraging social norms and framing information) with defaults found that only default nudges were effective at changing consumer pro-environmental behaviour.[8] One large RCT of calorie labelling in restaurants found that posting caloric benchmarks (an informational nudge) paradoxically increased caloric intake for consumers.[75]

The theoretical reasons for why less potent nudges (ie, nudges at the bottom of nudge ladder) often fail are well established. People have a limited capacity to process information, so providing more data to decision-makers can be distracting or cognitively loading.[76] The timing of information is also essential—information is beneficial if it is top-of-mind during the decision.[77] Some of the social comparison nudges in this review provided information at opportune times, others did not.[43] Additionally, information improves decisions only if existing heuristics encourage errors. Often the information individuals receive may not be new to them. Worse still, informational nudges can have negative unintended consequences. For example, alert fatigue describes when clinicians are so inundated by alerts that they become desensitised and either miss or postpone their responses to them.[78] Finally, often reminders and information frames can be insufficiently descriptive in the course of action they suggest, rendering them futile. Given how much of clinicians' time is spent with the EHR, health system decision supports must be effective and not self-undermining.

More potent nudges (ie, nudges at the top of the nudge ladder) are successful because they act on several key heuristics.[79] Defaults leverage inertia wherein overriding the default requires an active decision.[80] When people are busy and their attention scarce, they tend to rely on the status quo.[81] Moreover, people often see the default option as signaling an injunctive norm.[82] They see the default choice as the recommended choice and do not want to actively override this option unless they are very confident in their private decision. It is not surprising that our study found that defaults were effective. It is also not surprising that nudges leveraging peer comparison tended to also be effective at shaping clinician behaviour—clinicians who received messages that their behaviour was abnormal compared with their peers, received a signal that helped them update their behaviour.

Overall, results align with the conclusions of one[23] of the two recent systematic reviews of nudges tested in healthcare settings.[22 23] Differences in findings may be explained by different search strategies. One of these systematic reviews exclusively searched RCTs included in the Cochrane Library of systematic reviews and found that priming nudges—nudges that provide cues to participants—were the most studied and most effective nudges.[22] In that review, priming encompassed heterogenous interventions that span cues that elude conscious awareness, audit-and-feedback and clinician reminders—to name a few—which may account for why study authors found those nudges to be the most numerous. The findings from our review conform with the results of the more traditional systematic review, conducted using a systematic search of several databases.[23] The latter review, like this one, found that default and social comparison nudges were the most frequently studied and most effective nudges. However, study authors focused their review on physician behaviour, and our review is more expansive by studying all healthcare workers.

Limitations

Many of the studies in this review included at least some education (ie, a non-nudge intervention) such as a reminder of the clinical guidelines. Because many studies (59%) were pre–post designs, they could not use these brief trainings in a control arm to evaluate the independent effect of the nudge. Therefore, we cannot decisively conclude whether nudges alone are responsible for the changes in clinician behaviour. Similarly, many of the studies (51%) did not report the number of clinicians involved in the study (often reporting the sample in terms of how many patients or laboratory orders were affected by the nudge). Though unlikely, many of the effects could presumably be driven by a small portion of clinicians.

There was considerable variability in how researchers operationalised their primary outcome of interest. The effect of nudges may be contingent on the behaviour

Table 5 Cochrane risk of bias assessment tool

Authors (year, country)	Random sequence generation	Allocation concealment	Blinding (participants and personnel)	Blinding outcome assessors	Incomplete outcome data	Selective reporting
Allen *et al*[36] (2019, USA)	−	−	−	?	+	?
Andereck *et al*[37] (2019, USA)	−	−	−	?	+	?
Arora *et al*[38] (2019, USA)	−	−	−	?	−	?
Bourdeaux *et al*[39] (2014, UK)	−	−	−	?	+	?
Buntinx *et al*[40] (1993, Belgium)	+	+	+	+	+	?
Chiu *et al*[41] (2018, USA)	−	−	−	?	+	+
Delgado *et al*[42] (2018, USA)	−	−	−	?	+	?
Hemkens *et al*[43] (2017, Switzerland)	+	+	+	+	+	+
Hempel *et al*[44] (2014, USA)	−	−	−	?	?	?
Hsiang *et al*[45] (2019, USA)	−	−	−	?	+	?
Kim *et al*[46] (2018, USA)	−	−	−	?	+	+
Kullgren *et al*[47] (2018, USA)	+	+	+	+	+	+
Lewis *et al*[48] (2019, UK)	−	−	−	?	?	+
Meeker *et al*[49] (2014, USA)	+	+	+	+	+	+
Meeker *et al*[50] (2016, USA)	+	+	+	+	+	+
Nguyen and Davis[51] (2019, USA)	−	−	−	?	?	?
O'Reilly-Shah *et al*[52] (2018, USA)	−	−	−	?	+	+
Olson *et al*[53] (2015, USA)	−	−	−	?	+	?
Orloski *et al*[54] (2019, USA)	−	−	−	?	?	?
Parrino[55] (1989, USA)	−	−	−	?	+	?
Patel *et al*[58] (2014, USA)	−	−	−	?	+	?
Patel *et al*[73] (2016, USA)	+	+	+	+	+	+
Patel *et al*[57] (2016, USA)	+	+	+	+	+	+
Patel *et al*[56] (2017, USA)	+	+	+	+	+	+

Continued

Table 5 Continued

Authors (year, country)	Random sequence generation	Allocation concealment	Blinding (participants and personnel)	Blinding outcome assessors	Incomplete outcome data	Selective reporting
Patel *et al*[59] (2018, USA)	+	+	+	+	+	+
Persell *et al*[60] (2016, USA)	+	+	+	+	+	+
Ryskina *et al*[61] (2018, USA)	+	+	+	+	+	+
Sacarny *et al*[20] (2018, USA)	+	+	+	+	+	+
Sedrak *et al*[62] (2017, USA)	+	+	+	+	+	+
Sharma *et al*[63] (2019, USA)	+	+	+	+	+	+
Shively *et al*[64] (2020, USA)	−	−	−	?	?	?
Srinivasan *et al*[65] (2020, USA)	−	−	−	?	?	?
Suffoletto and Landau[66] (2019, USA)	+	+	+	+	+	+
Szilagyi *et al*[67] (2014, USA)	+	+	+	+	+	+
Trent *et al*[68] (2018, USA)	+	+	+	+	+	+
Wigder *et al*[69] (1999, USA)	−	−	−	?	?	?
Winickoff *et al*[70] (1984, USA)	First two studies: − Third study: +	First two studies: − Third study: +	First two studies: − Third study: +	?	+	+
Zivin *et al*[71] (2019, USA)	−	−	−	?	+	?
Zwank *et al*[72] (2017, USA)	−	−	−	?	+	?

+ indicates low risk of bias, − indicates high risk of bias, and ? indicates unclear risk of bias. See [34] for a full description of the Cochrane Risk of Bias tool.

under study. One study[71] examining changes in opioid prescriptions led to a change in the number of 15-pill prescriptions (ie, the change in 'default' orders) but not in the total quantity of opioid pills prescribed, whereas other studies resulted in changes in the total number of opioid pills ordered after an EHR default change. Establishing common metrics would enable direct comparison across studies and would allow us to conclusively determine if the nudge was effective overall at improving clinical decisions.

The considerable number of included papers reporting a statistically insignificant result decreases the usual concern over publication bias, which would skew the results towards desirable and more statistically significant outcomes. The majority of studies (n=21, 54%) were at unclear risk of selective reporting of outcomes (see table 5). Moving forward, the field would benefit from reporting of all experimentation, whether its results are successful, unsuccessful, significant or insignificant. Though not a majority, a large portion of studies (n=12, 31%) were conducted by the same research team in the same health system. To validate that clinician-directed nudges are effective in other settings, other researchers should conduct nudge studies.

Though the nudge taxonomy used in the current review offered a way to classify the nudges described in the studies included, it was not developed empirically. The nudge ladder was developed based on a theoretical understanding of nudge interventions. It is important to understand whether the conceptual distinctions made between nudge types are scientifically reliable and valid.

Future research

Behavioural economics recognises that nudges are 'implicit social interactions' between the decision-maker and the choice architect.[83] When faced with a nudge, people evaluate the motivations and values of the choice architect as well as how their decision will be understood by the choice architect and others. People tend to adhere to the default option when the choice architect is trusted, well-intentioned and expert. Several non-healthcare default studies backfired when consumers distrusted the choice architect or felt they were nudged to spend more money.[84] Clinicians may reject nudges when they perceive health systems' preferences to conflict with their patients' interests. Research should attend to how engaged clinicians are in the implementation process and how they make inferences about the motivations and values of the choice architect when interacting with nudges using qualitative methods and surveys.

Nudges are also dependent on how decision-makers believe they will be perceived. For example, around 40% of adults seeking care for upper respiratory tract infections want antibiotics and general practitioners report that patient expectations are a major reason for prescribing antibiotics.[85 86] Nudges that attempt to curtail antibiotic prescribing behaviour may shape clinicians' behaviours in unexpected ways given clinicians' desire to demonstrate to their patients that they are taking serious action. Subtle features of how nudges are implemented may also influence clinicians' perceptions of the choice architect, heighten awareness of how their own actions may be perceived and may undermine the nudge. Investigations of clinicians' choice environments and clinicians' perspectives using qualitative and survey methods are crucial to the success of nudges.

Future research should also explore how clinician-directed nudges interact with one another in clinicians' choice environments. In our review, all multicomponent nudge studies (n=5) were effective. However, it is also possible that nudges may crowd each other out when several different clinical decisions are targeted. In addition to alert fatigue, clinicians may experience nudge fatigue and begin to ignore decision support embedded in the EHR. Research should seek to understand how to develop nudges that can work synergistically with one another. Health systems and scientists can work together to understand which guidelines to prioritise and to develop decision support systems within their electronic interfaces that guide providers to make better clinical decisions.

Little work has been done on the sustainability of nudges beyond the study period, with some notable exceptions.[87] Particularly for nudges that require continued intervention on the part of the choice architects (eg, peer comparison interventions), it is necessary to also understand their cost-effectiveness. Finally, understanding how nudges can be implemented across health systems is essential given that many of the studies included in this review were conducted in one health system.

CONCLUSION

This study adds to the growing literature on the study and effectiveness of nudges in healthcare contexts and can guide health systems in their choices of the types of nudges they should implement to improve clinical practice. The review describes how nudges have been employed in healthcare contexts and the evidence for their effectiveness across clinician behaviours, demonstrating potential for nudges, particularly nudges that change default settings, enable choice, or frame information for clinicians. More research is warranted to examine how nudges scale and their global effect on improving clinical decisions in complex healthcare environments.

Author affiliations
[1]Psychology, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[2]Department of Family and Community Health, University of Pennsylvania School of Nursing, Philadelphia, Pennsylvania, USA
[3]Center for Health Incentives and Behavioral Economics (CHIBE), University of Pennsylvania, Philadelphia, Pennsylvania, USA
[4]Penn Implementation Science Center at the Leonard Davis Institute of Health Economics (PISCE@LDI), University of Pennsylvania, Philadelphia, Pennsylvania, USA
[5]College of Liberal and Professional Studies, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[6]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[7]Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[8]Department of Medical Ethics and Health Policy, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA
[9]Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

Acknowledgements The authors would like to thank Mitesh Patel, Anne Larrivee, Melanie Cedrone, Pamela Navrot, and Amarachi Nasa-Okolie for their assistance in the project.

Contributors BSL conceived of and designed the research study; acquired and analysed the data; interpreted the data; drafted the manuscript and substantially revised it. AMB helped design the research study; analysed the data; interpreted the data; and substantially revised the manuscript. CET analysed the data; interpreted the data; and substantially revised the manuscript. NM interpreted the data and substantially revised the manuscript. RSB helped conceive of and design the research study; interpreted the data; and substantially revised the manuscript. All authors approved the submitted version; have agreed to be accountable for the contributions; attest to the accuracy and integrity of the work, even aspects for which the authors were not personally involved.

Funding Funding for this study was provided by grants from the National Institute of Mental Health (P50 MH 113840, Beidas, Buttenheim, Mandell, MPI) and National Cancer Institute (P50 CA 244960, Beidas, Bekelman, Schnoll). BSL also receives funding support from the National Science Foundation Graduate Research Fellowship Program (DGE-1321851).

Competing interests BSL, AMB, CET and NM declare no financial or non-financial competing interests. RSB reports royalties from Oxford University Press, has received consulting fees from the Camden Coalition of Healthcare Providers, currently consults for United Behavioral Health and sits on the scientific advisory committee for Optum Behavioral Health.

Patient consent for publication Not required.

Ethics approval Given the nature of systematic reviews, no human participant research was conducted for this original research contribution. Thus, the systematic review was not deemed subject to ethical approval and no human participants were involved in this study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no data sets generated for this study. Given the nature of systematic reviews, the data set generated and analysed for the current study is already available. All studies analysed for the present review are referenced for readers.

ORCID iD

Briana S Last http://orcid.org/0000-0002-5473-1357

REFERENCES

1 Simon HA. *Models of bounded rationality: empirically grounded economic reason*. Cambridge, MA: MIT press, 1997.
2 Gigerenzer G, Gaissmaier W. Heuristic decision making. *Annu Rev Psychol* 2011;62:451–82.
3 Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. *Science* 1974;185:1124–31.
4 Thaler RH, Sunstein CR. *Nudge: improving decisions about health, wealth, and happiness. Revised and expanded Edition*. New York, NY: Penguin Books, 2009.
5 Szaszi B, Palinkas A, Palfi B, *et al*. A systematic scoping review of the choice architecture movement: toward understanding when and why nudges work. *J Behav Decis Mak* 2018;31:355–66.
6 Arno A, Thomas S. The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis. *BMC Public Health* 2016;16:676.
7 Harbers MC, Beulens JWJ, Rutters F, *et al*. The effects of nudges on purchases, food choice, and energy intake or content of purchases in real-life food purchasing environments: a systematic review and evidence synthesis. *Nutr J* 2020;19:1–27.
8 Momsen K, Stoerk T. From intention to action: can nudges help consumers to choose renewable energy? *Energy Policy* 2014;74:376–82.
9 Blumenthal-Barby JS, Krieger H. Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. *Med Decis Making* 2015;35:539–57.
10 Croskerry P. A universal model of diagnostic reasoning. *Acad Med* 2009;84:1022–8.
11 Graber M. Diagnostic errors in medicine: a case of neglect. *Jt Comm J Qual Patient Saf* 2005;31:106–13.
12 Almashat S, Ayotte B, Edelstein B, *et al*. Framing effect debiasing in medical decision making. *Patient Educ Couns* 2008;71:102–7.
13 Saposnik G, Redelmeier D, Ruff CC, *et al*. Cognitive biases associated with medical decisions: a systematic review. *BMC Med Inform Decis Mak* 2016;16:138.
14 Mendel R, Traut-Mattausch E, Jonas E, *et al*. Confirmation bias: why psychiatrists stick to wrong preliminary diagnoses. *Psychol Med* 2011;41:2651–9.
15 Redelmeier DA, Shafir E. Medical decision making in situations that offer multiple alternatives. *JAMA* 1995;273:302–5.
16 Patel MS, Volpp KG, Asch DA. Nudge units to improve the delivery of health care. *N Engl J Med* 2018;378:214–6.
17 Chaiyachati KH, Shea JA, Asch DA, *et al*. Assessment of inpatient time allocation among first-year internal medicine residents using time-motion observations. *JAMA Intern Med* 2019;179:760.
18 Hsiao C-J, Hing E. Use and characteristics of electronic health record systems among office-based physician practices: United States, 2001-2012. *NCHS Data Brief* 2014;143:1–8.
19 Henry J, Pylypchuk Y, Searcy TONC Data Brief, no.35. Adoption of electronic health record systems among U.S. Non-Federal acute care hospitals: 2008-2015. Washington DC Office of the National Coordinator for Health Information Technology; 2016.
20 Sacarny A, Barnett ML, Le J, *et al*. Effect of peer comparison letters for high-volume primary care prescribers of quetiapine in older and disabled adults: a randomized clinical trial. *JAMA Psychiatry* 2018;75:1003–11.
21 Lewis CC, Klasnja P, Powell BJ, *et al*. From classification to causality: advancing understanding of mechanisms of change in implementation science. *Front Public Health* 2018;6:136.
22 Yoong SL, Hall A, Stacey F, *et al*. Nudge strategies to improve healthcare providers' implementation of evidence-based guidelines, policies and practices: a systematic review of trials included within Cochrane systematic reviews. *Implementation Science* 2020;15:1–30.
23 Wang SY, Groene O. The effectiveness of behavioral economics-informed interventions on physician behavioral change: a systematic literature review. *PLoS One* 2020;15:e0234149.
24 Nuffield Council on Bioethics. *Public health: ethical issues [Internet]*. London, UK: Nuffield Council on Bioethics, 2007: 1–225. https://www.nuffieldbioethics.org/assets/pdfs/Public-health-ethical-issues.pdf
25 Patel MS. Nudges for influenza vaccination. *Nat Hum Behav* 2018;2:720–1.
26 Hillier-Brown FC, Summerbell CD, Moore HJ, *et al*. The impact of interventions to promote healthier ready-to-eat meals (to eat in, to take away or to be delivered) sold by specific food outlets open to the general public: a systematic review. *Obes Rev* 2017;18:227–46.
27 Münscher R, Vetter M, Scheuerle T. A review and taxonomy of choice architecture techniques. *J Behav Decis Mak* 2016;29:511–24.
28 Sunstein CR. Nudges vs. shoves. *Harv L Rev F* 2013;127:210.
29 Flodgren G, Eccles MP, Shepperd S, *et al*. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Syst Rev* 2011:CD009255.
30 Sackett DL, Rosenberg WMC, Gray JAM, *et al*. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312:71–2.
31 Biernacki P, Waldorf D. Snowball sampling: problems and techniques of chain referral sampling. *Sociol Methods Res* 1981;10:141–63.
32 Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*. 1 edn. Chichester (UK): John Wiley & Sons, 2008.
33 Davis DA, Mazmanian PE, Fordis M, *et al*. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* 2006;296:1094–102.
34 Higgins JPT, Altman DG, Gøtzsche PC, *et al*. The Cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
35 Nugent WR. The (non)comparability of the correlation effect size across different measurement procedures: a challenge to meta-analysis as a tool for identifying "evidence based practices". *J Evid Based Soc Work* 2011;8:253–74.
36 Allen JM, Dunn R, Bush J. Effect of prescriber peer comparison reports on fluoroquinolone use across a 16-facility community hospital system. *J Am Coll Clin Pharm* 2019;2:502–8.
37 Andereck JW, Reuter QR, Allen KC, *et al*. A quality improvement initiative featuring Peer-Comparison prescribing feedback reduces emergency department opioid prescribing. *Jt Comm J Qual Patient Saf* 2019;45:669–79.
38 Arora VM, Machado N, Anderson SL, *et al*. Effectiveness of SIESTA on objective and subjective metrics of nighttime Hospital sleep disruptors. *J Hosp Med* 2019;14:38–41.
39 Bourdeaux CP, Davies KJ, Thomas MJC, *et al*. Using 'nudge' principles for order set design: a before and after evaluation of an electronic prescribing template in critical care. *BMJ Qual Saf* 2014;23:382–8.
40 Buntinx F, Knottnerus JA, Essed GG, *et al*. Long-term effect of feedback and peer comparison on the sampling quality of cervical smears--a randomized controlled trial. *Eur J Cancer Prev* 1995;4:153–8.

41 Chiu AS, Jean RA, Hoag JR, *et al*. Association of lowering default pill counts in electronic medical record systems with postoperative opioid prescribing. *JAMA Surg* 2018;153:1012–9.
42 Delgado MK, Shofer FS, Patel MS, *et al*. Association between electronic medical record implementation of default opioid prescription quantities and prescribing behavior in two emergency departments. *J Gen Intern Med* 2018;33:409–11.
43 Hemkens LG, Saccilotto R, Reyes SL, *et al*. Personalized prescription feedback using routinely collected data to reduce antibiotic use in primary care: a randomized clinical trial. *JAMA Intern Med* 2017;177:176–83.
44 Hempel D, Pivetta E, Kimberly HH. Personalized peer-comparison feedback and its effect on emergency medicine resident ultrasound scan numbers. *Crit Ultrasound J* 2014;6:1.
45 Hsiang EY, Mehta SJ, Small DS, *et al*. Association of an active choice intervention in the electronic health record directed to medical assistants with clinician ordering and patient completion of breast and colorectal cancer screening tests. *JAMA Netw Open* 2019;2:e1915619.
46 Kim RH, Day SC, Small DS, *et al*. Variations in influenza vaccination by clinic appointment time and an active choice intervention in the electronic health record to increase influenza vaccination. *JAMA Netw Open* 2018;1:e181770.
47 Kullgren JT, Krupka E, Schachter A, *et al*. Precommitting to choose wisely about low-value services: a stepped wedge cluster randomised trial. *BMJ Qual Saf* 2018;27:355–64.
48 Lewis S, Young B, Thurley P, *et al*. Evaluation of a nudge intervention providing simple feedback to clinicians of the consequence of radiation exposure on demand for computed tomography: a controlled study. *Clin Med* 2019;19:290–3.
49 Meeker D, Knight TK, Friedberg MW, *et al*. Nudging guideline-concordant antibiotic prescribing: a randomized clinical trial. *JAMA Intern Med* 2014;174:425–31.
50 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.
51 Nguyen CT, Davis KA. Evaluating the impact of peer comparison on vancomycin dose order verification among pharmacists. *J Am Coll Clin Pharm* 2019;2:137–42.
52 O'Reilly-Shah VN, Easton GS, Jabaley CS, *et al*. Variable effectiveness of stepwise implementation of nudge-type interventions to improve provider compliance with intraoperative low tidal volume ventilation. *BMJ Qual Saf* 2018;27:1008–18.
53 Olson J, Hollenbeak C, Donaldson K, *et al*. Default settings of computerized physician order entry system order sets drive ordering habits. *J Pathol Inform* 2015;6:16.
54 Orloski CJ, Tabakin ER, Shofer FS, *et al*. Grab a seat! Nudging providers to sit improves the patient experience in the emergency department. *J Patient Exp* 2019;6:110–6.
55 Parrino TA. The nonvalue of retrospective peer comparison feedback in containing Hospital antibiotic costs. *Am J Med* 1989;86:442–8.
56 Patel MS, Volpp KG, Small DS, *et al*. Using active choice within the electronic health record to increase influenza vaccination rates. *J Gen Intern Med* 2017;32:790–5.
57 Patel MS, Volpp KG, Small DS, *et al*. Using active choice within the electronic health record to increase physician ordering and patient completion of high-value cancer screening tests. *Healthc* 2016;4:340–5.
58 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. *Ann Intern Med* 2014;161:S44–52.
59 Patel MS, Kurtzman GW, Kannan S, *et al*. Effect of an automated patient Dashboard using active choice and peer comparison performance feedback to physicians on statin prescribing: the prescribe cluster randomized clinical trial. *JAMA Netw Open* 2018;1:e180818.
60 Persell SD, Doctor JN, Friedberg MW, *et al*. Behavioral interventions to reduce inappropriate antibiotic prescribing: a randomized pilot trial. *BMC Infect Dis* 2016;16:373.
61 Ryskina K, Jessica Dine C, Gitelman Y, *et al*. Effect of social comparison feedback on laboratory test ordering for hospitalized patients: a randomized controlled trial. *J Gen Intern Med* 2018;33:1639–45.
62 Sedrak MS, Myers JS, Small DS, *et al*. Effect of a price transparency intervention in the electronic health record on clinician ordering of inpatient laboratory tests: the price randomized clinical trial. *JAMA Intern Med* 2017;177:939–45.
63 Sharma S, Guttmann D, Small DS, *et al*. Effect of introducing a default order in the electronic medical record on unnecessary daily imaging during palliative radiotherapy for adults with cancer: a Stepped-Wedge cluster randomized clinical trial. *JAMA Oncol* 2019;5:1220–2.
64 Shively NR, Buehrle DJ, Wagener MM. Improved antibiotic prescribing within a Veterans Affairs primary care system through a multifaceted intervention centered on peer comparison of overall antibiotic prescribing rates. *Antimicrob Agents Chemother* 2020;64:e00928–19.
65 Srinivasan M, Huntman J, Nelson M, *et al*. Use of peer comparison, provider education, and electronic medical record triggers to increase influenza vaccination rates in hospitalized children. *Hosp Pediatr* 2020;10:76–83.
66 Suffoletto B, Landau A. Nudging emergency care providers to reduce opioid prescribing using peer norm comparison feedback: a pilot randomized trial. *Pain Med* 2020;21:1393–9.
67 Szilagyi PG, Serwint JR, Humiston SG, *et al*. Effect of provider prompts on adolescent immunization rates: a randomized trial. *Acad Pediatr* 2015;15:149–57.
68 Trent SA, Havranek EP, Ginde AA, *et al*. Effect of audit and feedback on physician adherence to clinical practice guidelines for pneumonia and sepsis. *Am J Med Qual* 2019;34:217–25.
69 Wigder HN, Cohan Ballis SF, Lazar L, *et al*. Successful implementation of a guideline by peer comparisons, education, and positive physician feedback. *J Emerg Med* 1999;17:807–10.
70 Winickoff RN, Coltin KL, Morgan MM, *et al*. Improving physician performance through peer comparison feedback. *Med Care* 1984;22:527–34.
71 Zivin K, White JO, Chao S, *et al*. Implementing electronic health record default settings to reduce opioid Overprescribing: a pilot study. *Pain Med* 2019;20:103–12.
72 Zwank MD, Kennedy SM, Stuck LH, *et al*. Removing default dispense quantity from opioid prescriptions in the electronic medical record. *Am J Emerg Med* 2017;35:1567–9.
73 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 Intern Med* 2016;176:847–8.
74 Sunstein CR. Nudges that fail. *Behav Public Policy* 2017;1:4–25.
75 Downs JS, Wisdom J, Wansink B, *et al*. Supplementing menu labeling with calorie recommendations to test for facilitation effects. *Am J Public Health* 2013;103:1604–9.
76 Marois R, Ivanoff J. Capacity limits of information processing in the brain. *Trends Cogn Sci* 2005;9:296–305.
77 Karlan D, McConnell M, Mullainathan S, *et al*. Getting to the top of mind: how reminders increase saving. *Manage Sci* 2016;62:3393–411.
78 Sendelbach S, Funk M. Alarm fatigue: a patient safety concern. *AACN Adv Crit Care* 2013;24:378–86.
79 Jachimowicz JONM, Duncan S, Weber EU, *et al*. When and why defaults influence decisions: a meta-analysis of default effects. *Behav Public Policy* 2019;3:159–86.
80 Madrian BC, Shea DF. The power of suggestion: inertia in 401(k) participation and savings behavior. *Q J Econ* 2001;116:1149–87.
81 Kahneman D, Knetsch JL, Thaler RH. Anomalies: the endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives* 1991;5:193–206.
82 Everett JAC, Caviola L, Kahane G, *et al*. Doing good by doing nothing? the role of social norms in explaining default effects in altruistic contexts. *Eur J Soc Psychol* 2015;45:230–41.
83 Krijnen JMT, Tannenbaum D, Fox CR. Choice architecture 2.0: behavioral policy as an implicit social interaction. *Behavioral Science & Policy* 2017;3:i–18.
84 Brown CL, Krishna A. The skeptical shopper: a metacognitive account for the effects of default options on choice. *J Consum Res* 2004;31:529–39.
85 Fletcher-Lartey S, Yee M, Gaarslev C, *et al*. Why do general practitioners prescribe antibiotics for upper respiratory tract infections to meet patient expectations: a mixed methods study. *BMJ Open* 2016;6:e012244.
86 Linder JA, Singer DE. Desire for antibiotics and antibiotic prescribing for adults with upper respiratory tract infections. *J Gen Intern Med* 2003;18:795–801.
87 Olshan D, Rareshide CAL, Patel MS. Longer-Term durability of using default options in the electronic health record to increase generic prescribing rates. *J Gen Intern Med* 2019;34:349–50.