saxinstitute

Evidence Check

Addressing unwarranted variation in healthcare

An **Evidence Check** rapid review brokered by the Sax Institute for Cancer Institute NSW. March 2019

An **Evidence Check** rapid review brokered by the Sax Institute for Cancer Institute NSW.
March 2019.

This report was prepared by: Reema Harrison, Elizabeth Manias, Steven Mears, Reece Hinchcliff, David Heslop.
UNSW Sydney, March 2019

Enquiries regarding this report may be directed to the:
Principal Analyst
Knowledge Exchange Program
Sax Institute
www.saxinstitute.org.au
knowledge.exchange@saxinstitute.org.au
Phone: +61 2 91889500

Suggested Citation:

Harrison R, Manias E, Mears S, Hinchcliff R, Heslop D. Addressing unwarranted variation in healthcare: an Evidence Check rapid review brokered by the Sax Institute (www.saxinstitute.org.au) for Cancer Institute NSW, 2019

Addressing unwarranted variation in healthcare

An **Evidence Check** rapid review brokered by the Sax Institute for Cancer Institute NSW.
March 2019

This report was prepared by Reema Harrison, Elizabeth Manias, Steven Mears, Reece Hinchcliff, David Heslop.

Contents

Executive summary

Background

Clinical variation in healthcare describes differences in healthcare practice, processes or outcomes. Such variations are evident throughout healthcare systems and services internationally and reflect natural differences between the individuals, population groups that receive care and service provision. Atlases of Variation in Australia demonstrate the degree of variation that occurs across and within the health system and services. While methods for determining variation utilising statistical and framework approaches are well established, there is less clarity regarding how to define the variation that warrants action. Contemporary approaches to understanding variation have conceptualised clinical variation as a catalyst for action, with continuous quality improvement approaches utilising facilitated feedback adopted as a strategy to respond to variation and determine the changes required. To date, a diverse range of facilitated feedback approaches have been adopted, with lack of synthesis of the methods employed or evidence of their effectiveness. As such this review first seeks to establish the current evidence of methods and frameworks that have been used to determine when variation in healthcare is considered unwarranted and change is justified, through a rapid review update. It then seeks to explore the models of facilitated feedback that have been utilised to respond to variation and evidence of the effectiveness of approaches used to date.

Review questions

This review aimed to address the following questions:

Question 1: What methods and frameworks have been used to determine when variation in healthcare is unwarranted and change is justified?

Question 2: What methods and models of facilitated feedback and facilitated continuous quality improvement have been shown to be effective in addressing unwarranted variation in healthcare?

Summary of methods

Rapid evidence assessment (REA) methodology was used. Two separate searches were undertaken to address the two review questions. For Question 1, a range of text words, synonyms and subject headings were developed for the major concepts of unwarranted clinical variation, standards (and deviation from these standards) and healthcare environment. For Question 2, a range of text words, synonyms and subject headings were developed for the major concepts of clinical variation, quality improvement and (facilitated) feedback. Two electronic databases (Medline and PubMed) were searched from January 2000 to August 2018, in addition to hand-searching of relevant journals, reference lists and grey literature. Results were merged using reference-management software (Covidence) and duplicates removed. The inclusion criteria were independently applied to potentially relevant articles by three reviewers. Findings were presented in a narrative synthesis to highlight key concepts addressed in the published literature.

Key findings

A total of eight publications were included in this review to address Question 1, and 32 studies were included to address question 2. For Question 1, all articles were identified from the database search. For Question 2, 31 articles were identified as eligible from the database search, and one further study was identified from hand-searching published work. Three case studies were extracted from the grey literature

as examples of the approaches identified in the peer reviewed work, in addition to a range of relevant online resources.

The findings provide evidence that facilitated feedback methods and models are used internationally to respond to variation, and that these largely focus on changing clinician behaviour as opposed to systems changes. Only a small number of studies include patient engagement or education within feedback approaches. The body of evidence identified indicates that a range of feedback approaches can reduce clinical practice variation that arises from clinical decision-making and behaviours. Facilitated approaches to providing feedback are widely used in health services internationally to provide a nuanced and continuous improvement approach to respond to variation. However, there is no evidence to suggest that facilitated approaches as a group, or a particular facilitated feedback model or method, is more effective in responding to variations appropriately than simply providing feedback to individuals, teams or networks of health providers. Evaluation of the effectiveness of approaches utilising facilitated feedback are needed to provide evidence to help answer two questions:, firstly regarding whether facilitated feedback offers advantages over feedback without facilitation in the context of addressing variation; and secondly, to determine if there is an optimal model and/or method of facilitation that is more likely to create change where needed.

The evidence in this review identified a lack of recognition of the contribution that patient preferences and factors make to clinical variation in healthcare. While shared decision-making and patient-centred care approaches are identified as important, these have not been sufficiently explored or evaluated in terms of identifying and responding to variation.

Background

Clinical variation has attracted increasing interest in the Australian health system as a mechanism for understanding the quality and appropriateness of care provided to patients, highlighting features such as efficient, effective and timely care.[1] Along with countries including the United States, Canada, Spain, England, Germany, the Netherlands, Norway and New Zealand, Australia has produced atlases of variation in health care to guide service improvements.[2]

The second Australian Atlas of Healthcare Variation (2017) demonstrates substantial variations in the medications, interventions and procedures provided to patients across Australia, with implications for patient outcomes.[3] Variations were reported in a range of care areas including surgery for hysterectomy, cataract surgery, knee replacement and potentially preventable hospitalisations for selected conditions, including diabetes complications. The report also includes a specific section on variations in women's healthcare.

It is widely acknowledged that not all variation is unwarranted and that some variation may in fact be a marker of effective patient-centred care.[4] Unwarranted clinical variation describes "variation that cannot be explained by the condition or the preference of the patient; it is variation that can only be explained by differences in health system performance", for example, the effectiveness of the structures, processes and services that form any given health system.[3] Reducing unwarranted clinical variation is critical in the context of value-based healthcare that comprises two dimensions; allocative value (the degree to which population resources are allocated to different groups within that population) and the optimising of the value of resources through their utilisation for each patient sub-group, which is determined by clinicians.[5] In some healthcare systems such as in the US, healthcare providers are also transitioning from volume-based to value-based payments for care.

Well-established statistical and framework approaches for gathering evidence of variation in the processes and treatments undertaken across health systems internationally have led to a substantial body of literature.[6-9] While methods for detecting variation (e.g. exploring statistically significant deviation from acceptable parameters) are widely acknowledged, methods for determining the variation that warrants action or is considered problematic are strongly debated.[7] A 2017 review exploring approaches to address unwarranted clinical variation demonstrated the challenge of setting parameters for variation that can be considered unwarranted when looking beyond the category of treatments or procedures that are deemed to be 'effective care'; that is, those agreed to be the optimum care for all patients.[9] When no single optimal approach indicates effective care, operationalising current frameworks used for categorising types of variation in order to identify and address instances of unwarranted clinical varation is challenging.[9]

Given the conceptual and operational obstacles for systematic identification of unwarranted clinical varation, contemporary literature has conceptualised clinical variation data as a catalyst for exploring the appropriateness of care in a given location or service.[10] With variation data, it is possible to stimulate discussion regarding the quality and appropriateness of care provision and identify areas in which better value care can be obtained for systems, services and patients.[2] Continuous quality improvement activities have been identified as a strategy to respond to and explore variation. This identification enables health services to determine whether change is required and the change that is warranted, with a range of approaches evident in health systems internationally.

Feedback regarding benchmark data has been utilised in several countries as a basis for continuous quality improvement to address clinical care variation and enhance guideline-adherent care.[11, 12]

In Australia, the Australian Commission on Safety and Quality in Health Care (ACSQHC) has developed the Framework for Australian Clinical Quality Registries as a mechanism for governments and health services to capture the appropriateness and effectiveness of care within their jurisdiction.[13] Similarly, in the UK, clinical registries have been adopted and also linked with financial incentives for appropriate care.[14] Mechanisms for providing rapid feedback to individual clinicians are also identified in the context of responding to clinical variation, with training and checklists developed to accompany feedback data.[15, 16]Furthermore, the provision of facilitated feedback using these clinical registry data has been shown to contribute to improved patients outcomes. For instance, in the United Kingdom, since June 2013, through the National Cancer Registration and Analysis Service (2018), the National Health Service (NHS) has published clinician-level cancer surgery outcome data to address variations and improve outcomes. Results indicate positive patient benefits including reducing morbidity and mortality, and improving patient survival.[17] Similarly, in some US states, report cards are published for individual cardiac surgeons and hospitals. This has been evaluated to be a valid approach in improving patient outcomes. For instance, in a survey of 317 cardiologists in New York State, Brown et al. (2013) found almost all cardiologists (94%) were aware of these report cards and one in four had reported a moderate or substantial influence of these report cards on their referral decision-making.[18]

To date, the widespread use of facilitated feedback approaches to change practice across health systems and services internationally has not been subject to evidence synthesis to determine the effectiveness of methods used in responding to clinical variation. This evidence is critical to guide policy development regarding optimal approaches to respond to clinical variation, hence it is the focus of this present review.

Aim: This review aimed to identify frameworks and approaches to identify variation and establish whether change is justified, and to explore the methods of facilitated feedback used and their effectiveness for addressing clinical variation.

Methods

This literature review utilised a rapid evidence assessment (REA) methodology. A REA is a research methodology that uses the same methods and principles as a systematic review, but makes concessions to the breadth or depth of the process in order to suit a shorter timeframe and address key issues in relation to the topic under investigation.[19]

The purpose of a REA is to provide a balanced assessment of what is already known about a specific problem or issue. The shorter timeframe, lower cost (relative to full systematic reviews), and evaluation of the strength of the evidence make REAs particularly helpful in informing policy and decision-makers, program managers and researchers.

REAs utilise a number of strategies to assist in facilitating rapid synthesis of information. These strategies include: having a narrow question; limiting the timeframe in which studies are published; limiting the scope to English language articles,; and making concessions on how the published studies are synthesised. Often REAs make use of existing high-quality guidelines or systematic reviews/meta-analyses to assist the rapid process. Thus, undertaking a REA maximises information in the existing synthesised literature in order to minimise time and cost.[20]

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA statement) was used to guide the reporting of this rapid review.[21] The PRISMA statement is an evidence-based approach for reporting systematic reviews and meta-analyses.

Question 1 inclusion criteria

Types of publication: Publications were eligible if they were available in English and reported original primary empirical or theoretical work published from 2000 – 2018.

Types of settings: Public or private hospitals, day procedure centres, general practice or other primary/community care facilities that include an adult population (18 years and above).

Types of study design: Conceptual, theoretical, quantitative or qualitative studies of any research design.

Outcomes: Conceptual or theoretical frameworks that are used to identify and/or understand warranted or unwarranted clinical variation in relation to any healthcare outcome and/or data regarding approaches to provide feedback related to clinical variation in relation to any healthcare outcome.

Clinical variation was defined as circumstances in which "patients with similar diagnoses, prognoses and demographic status receive different levels of care depending on when, where and by whom they are treated, despite agreed and documented evidence of best practice".[20] Due to the volume of clinical variation literature, unwarranted clinical variation was used explicitly to maintain focus on the review objectives. Unwarranted clinical variation was defined as "variation that cannot be explained by the condition or the preference of the patient; it is variation that can only be explained by differences in health system performance".[3]

Question 1 exclusion criteria

Articles were excluded if they did not meet the inclusion criteria. Non-empirical literature such as opinion pieces, letters and editorials were excluded, along with studies that employed hypothetical vignettes. Studies from developing countries were excluded because the review aimed to identify approaches that would be applicable to the Australian healthcare context.

Question 2 inclusion criteria

Types of publication: Publications were eligible if they were available in English and reported original primary empirical or theoretical work published from 2000-2018.

Types of settings: Health systems or studies in any of the following settings within health services were included: public or private hospitals, day procedure centres, general practice or other primary/community care facilities.

Types of study design: Conceptual, theoretical, quantitative or qualitative studies of any research design.

Interventions:

- Studies that reported the use of any mode of facilitated feedback to respond to clinical variation.
- Studies reporting feedback processes provided by health system agencies directly to health services providers, health districts, or clinicians were eligible.
- Studies reporting feedback in the context of continuous quality improvement (defined as the use of quality "indicators") to initiate and drive practice changes in an ongoing cycle of continuous improvement were eligible.
- Facilitated feedback defined as the reporting of outcomes directly to key stakeholders with ongoing dialogue geared toward change or any other activities to support change.

Outcomes: Changes in clinical practice variation – perceived or actual.

Question 2 exclusion criteria

Articles were excluded if they did not meet the above inclusion criteria. Non-empirical or primary literature such as opinion pieces, letters and editorials were excluded. Articles that assessed or identified clinical practice variation that did not either focus on feedback and/or quality improvement mechanisms were excluded for Question 2.

Peer reviewed literature

Two separate searches were undertaken to address the two review questions. For Question 1, a range of text words, synonyms and subject headings were developed for the major concepts of unwarranted clinical variation, standards (and deviation from these standards) and healthcare environment. For Question 2, a range of text words, synonyms and subject headings were developed for the major concepts of clinical variation, quality improvement and (facilitated) feedback. These text words, synonyms and subject headings were used to undertake a systematic search of two electronic databases that index journals of particular relevance to the review topic (Medline and PubMed) from January 2000 to August 2018, in order to focus the search on findings that were relevant for contemporary policy development. Two different strategies were employed to address research Question 1 (see Appendix 1) and research Question 2 (Appendix 2). Hand-searching of reference lists of published papers ensured that all relevant published material was captured. Results were merged using reference-management software (Endnote, version X8) and duplicates removed.

Grey literature

Relevant work reported in the grey literature (e.g. reports and papers published by government departments, intragovernmental agencies, public or private health service providers, non-government agencies, consumer organisations, professional bodies, advocacy groups etc.) were identified by searching the websites of relevant organisations (see Appendix 3 for a list of the relevant organisations and websites included). Literature identified was assessed along with the papers from the database searches.

Study selection and data extraction

Database search

Three reviewers (EM, DH, RH) independently screened the titles and abstracts. Copies of the full articles were obtained for those that were potentially relevant. Inclusion criteria were then independently applied to the full text articles by each of the members of the reviewer team (all authors). Disagreements were resolved through final discussion between two members of the review team (RH, EM). The following data were extracted from eligible literature: author(s), publication year, sample, setting, objective, framework used (Question 1) or feedback approach (Question 2) and main findings.

Grey literature

For Question 2, a list of websites and other potential sources of relevant grey literature were determined collaboratively amongst the research team and funding body, based on their prior research into unwarranted clinical variation.[6] The websites listed in Appendix 3 were searched from 10 - 13 September 2018 to identify reports and other sources of information related to the study questions. There were few examples identified of strong empirical research, limiting the additional contribution of the grey literature to findings elicited from the peer reviewed literature summarised above. However, a number of case studies were identified that exemplified some of the major quality improvement approaches used in response to identified variation, and learnings from these were extracted by two authors. A review of approaches to reducing unwarranted clinical variation also identified some relevant material regarding Question 2, despite not focusing primarily on facilitated feedback approaches.

Data synthesis

Findings were analysed using a narrative empirical synthesis in stages, based on the study objectives.[22] A narrative approach was necessary to synthesise the qualitative and quantitative findings. We did not feel a quantitative analytic approach would be appropriate due to the heterogeneity of study designs, contexts, and types of literature included. Initial descriptions of the eligible studies and results were tabulated (Appendices 4 and 5). Patterns in the data were explored to identify consistent findings in relation to the study objectives. Interrogation of the findings explored relationships between study characteristics and their findings; the findings of different studies; and the influence of the use of different outcome measures, methods and settings on the resulting data. The peer reviewed literature was then subjected to an appraisal process before a narrative synthesis of the findings was produced.

Included studies

After removing duplications, 70 records were identified for Question 1 and 342 records for Question 2. Title and abstract screening review resulted in 46 references that fulfilled the inclusion criteria for Question 1 and 53 references that fulfilled the inclusion criteria for Question 2, for which full text of the publications were obtained. A total of eight publications was included in the review for Question 1, all identified from the database search. For Question 2, 32 publications were included in the review based on the inclusion and exclusion criteria – 31 articles identified as eligible from database full text review and one further study that was identified via hand-searching.. A further three studies were identified from the grey literature as case study examples. Appendix 6 shows a flowchart of the literature selection process for Question 1 (Figure 1a) and Question 2 (Figure 1b).. Summary tables of the included studies from the database search are attached as Appendices 4 and 5.

Data appraisal

An assessment of study quality was undertaken using the Quality Assessment Tool of Studies of Diverse Design (QATSDD) for assessing heterogeneous groups of studies.[23] This tool is suitable for assessing the quality of evidence in reviews that synthesise qualitative, quantitative and/or mixed-methods research. The

tool has been used widely in health services research. Publications identified in the database search were scored against each criterion on a four-point scale (0-3) to indicate the quality of each publication and the overall body of evidence. The criteria are shown in Appendix 7. Given the heterogeneity in the nature of the relevant publications, and the grey literature was also included in the review, we did not exclude studies based on the quality assessment. Quality assessment data was used only to explore the strength of the available evidence.

Excluded studies

Question 1: Nineteen studies were excluded during title and abstract screening because they did not meet the inclusion criteria relating to publication type (18) or study population (1). From the 46 remaining studies a further 38 were excluded at the full-text review stage.

Question 2: Title and abstract screening excluded 286 studies. A further 22 studies were excluded at the full-text review stage because they did not meet the inclusion criteria relating to publication type (17) or did not include a feedback approach (2).

Study quality

The data appraisal identified that the included papers retrieved from the database searches were generally of good quality with particular strengths in the application of evidence to inform the quality improvement strategies and in the selection of appropriate study designs and analytic strategies. Many studies included a steering-group to inform design of the study. A key limitation across the body of evidence for Question 2 was the use of small samples which were not necessarily representative of the wider organisation or network. Such studies were often also single-site.

Findings

Question 1: What methods and frameworks have been used to determine when variation in healthcare is unwarranted and change is justified?

Eight studies were identified in this rapid review update emerging from seven countries: Australia (2), UK (1), The Netherlands (1), Germany (1), Canada (1), Sweden (1) and the US (1).[24-31] The included studies comprised of two data linkage studies, two systematic reviews incorporating meta-analysis or prospective survey data, one database analysis, one case study, one observational study and one narrative review. Aligning with the existing review data, two groups of studies reported approaches to identify or determine unwarranted clinical variation. The first group of three studies applied or discussed a framework such as those described in the introduction to determine unwarranted clinical variation in population and/or hospital level data.[26, 28, 29]The second group of five studies applied statistical models, most commonly regression analyses, to identify variation that was considered to be deviating at a statistically significant level from standard, appropriate or expected levels of variability.[24, 25, 27, 30, 31] Figure 2 provides a graphical presentation of the processes identified in the literature (Appendix 8).

I. Framework-based identification of unwarranted clinical variation

The Wennberg Classification System dominated frameworks for identifying unwarranted variation in the review update, with unwarranted clinical variation conceptualised in three categories: effective care, preference-sensitive care and supply-sensitive care.[8] Effective care denotes those services and procedures that have been proven effective in the research literature for all patients. Circumstances in which more than one 'medically acceptable' option exists are described as preference-sensitive care because the choice regarding the best treatment or care option is based on patient preferences, such as in the treatment of early stage prostate cancer. Appropriate decision support tools are critical for these situations to support informed patient decision-making. Supply-sensitive care relates to the capacity of the surrounding healthcare system to provide a given treatment or service. If a service or treatment is readily available then utilisation is likely to be greater than if it is not, resulting in variations between different services and localities.[6]

Three articles were included in this review update that were grounded in the Wennberg approach but focused specifically on the role of physicians and surgeons in creating unwarranted variations.[26, 28, 29] In the first article, Mercuri and Gafni (2018)[29] build a case for the contribution of physicians as the source of unwarranted variations. The narrative review details the evidence that physician-related factors contribute only a small proportion of care variance (<10%) in studies such as those of hospital admission rates, laboratory tests and length of stay. These findings are presented in contrast to the more substantial contribution of patient characteristics and preferences. As such, approaches to identify problematic variation that focus only on physician behaviour are likely to be insufficient. Mercuri and Gafni 2018 extend this argument when detailing the degree of uncertainty in medical management. Defining unwarranted variation by deviation from evidence-based guidelines assumes that all management decisions are based on science, but Mercuri and Gafni argue that in patient-centred practice, a number of factors contribute to decisions about medical management and what is optimal for each patient.

A second study by Mayer et al (2017)[28] reiterates this when demonstrating the extent of deviation from effective care identifiable in total knee or hip arthroplasty, with patient factors identified as one source. The narrative review also builds an argument to challenge the extent of variation explained by supply-sensitive care, highlighting the challenges of determining the point at which this might indicate care is unwarranted.[29]

Feufel (2018)[26] examines emergency department practice against the Wennberg categories and, similar to Mercuri and Gafni, demonstrates the substantial contribution of patient preferences to variation data. Together, these publications indicate that approaches to identify and address variation that do not incorporate the patient contribution may be missing vital detail.

II. Statistically-defined identification of unwarranted clinical variation

This review update identified five studies that applied statistical methods to determine variation that is unwarranted in a range of settings. Three studies applied regression analyses to explore patient-related and non-patient related factors in thoracic aortic disease, predictors of variation in outpatient physician visits, and variation in readmission, mortality, costs and multiple process indicators in acute coronary syndrome patients.[24, 27, 28] Through combining a process of systematic review, process and structure questionnaires of cardiac surgery units and an analysis of hospital episode statistics and cardiac surgery audit data, Bottle et al (2017)[24] determined substantial regional variation after controlling for patient and disease factors. The authors concluded from the hospital and clinical audit data that units with higher case volumes were those that treated more-complex patients and had significantly lower risk-adjusted mortality relative to low-volume units. The systematic review indicated that the delivery of care by multidisciplinary teams in high-volume units resulted in better outcomes. But the survey of units suggested this was not the way in which services were currently organised.[24] Johannsen et al (2018)[27] utilised data from 21 Swedish county councils to understand the degree of regional variation in outpatient physician visits that was explained by demand factors such as health, demography and socioeconomic indicators. Regional mortality, as a proxy for population health, and demography explained around 50% of regional variation in visits to outpatient specialists, but did not explain variation in visits to primary care physicians. Variation in primary care visits was explained to some extent (33%) when adjusting for socioeconomic and supply factors but 50–67% of regional variation remained unexplained in their model.[27] Partington et al (2017)[30] used routinely collected hospital data to capture variation in re-admissions and mortality at 30 days and 12 months, in addition to patient costs and data regarding multiple process indicators. One of the four public hospitals was identified as an outlier with statistically significant variations in readmission and mortality data. Variations at the outlier hospital were linked to a range of process indicators including admission patterns, use of invasive diagnostic procedures, length of stay, patients' presenting characteristics and time of presentation. However, the regression analysis could not provide evidence of the nature of the relationship between these variables and outcome measures.[30]

One further study employed forest plots to demonstrate variability in open bypass or endovascular intervention.[31] The authors reported significant variation indicating potential unwarranted variation in the proportion of prosthetic conduit for infrapopliteal bypass in claudication, isolated tibial endovascular intervention for claudication, discharge on antiplatelet and statin, and ultrasound guidance for percutaneous access.[31] The four articles were reliant on retrospective data, with the potential inaccuracies of this impacting the analyses. The methods used in each case were only able to identify activities outside the pre-determined parameters of acceptable variation and were not able to explore nuanced aspects of the context or service/s that influenced these. In addition to these studies that focused on the principles of deviation from effective-care, one other separate study included in the review conducted an econometric analysis of spending, supply and demand variables for patients with diabetes, depression and the general population.[25] In each group, patient-level demand variables explained 62–63% of the total variance and self-reported health status was also a significant predictor, explaining 28% of healthcare spending. Demand variables explained almost all regional variation in spending for depression and 88% for diabetes. Only 12% of the regional variation remained unexplained, with the authors suggesting that this indicated differences between regions due to inefficiencies.[25]

Question 2: What methods and models of facilitated feedback and facilitated continuous quality improvement have been shown to be effective in addressing unwarranted variation in healthcare?

I. Database literature

The database search retrieved 32 articles that were eligible for inclusion from nine countries: US (16), UK (4), Australia (4), The Netherlands (2), Canada (2), Sweden (1), Norway (1), Egypt (1), and New Zealand (1). While this review sought to identify approaches using facilitated feedback to respond to variation, three levels of action in response to variation became apparent through the search and selection process and these were used to frame the findings. The first category of evidence reported approaches to determine effective and appropriate care to address variation, primarily through guideline development. The second category of evidence were approaches that involved providing feedback at an individual, local or organisational level to bring variation to the attention of clinicians and instigate corrective responses where necessary. The third and final category reported methods and models for providing facilitated feedback to respond to variations in clinical practice. While the included studies predominantly identified the overarching goal of minimising variations that were unwarranted, the projects reported continuous processes to respond to clinical variation in general and did not seek to identify unwarranted clinical varation and then to address this specifically. None of the included studies explicitly included an aim to address unwarranted clinical variation.

Category 1 – Setting out optimal (effective) care

I. Guideline or pathway

While a broad literature is available around the development and use of guidelines, two studies were identified in the current review that explicitly discussed utilising guidelines in the context of quality improvement to address problematic variation.[32] Cammisa et al (2011)[33] reported on a study exploring variation in data from a health plan in the US. Using a clinical practice guideline to determine effective practice, the project sought to identify and address overuse of chronic and acute back pain practices in five areas. The intervention involved outreach to practices to bring the clinical practice guidelines – and deviation from these – to their attention through discussion over a six-month period. The intervention process led to significant reductions in many of the overused practices, however, the study did not control for possible factors that may have influenced this behaviour sufficiently to determine causation.[33] In a case study of a quality improvement initiative, Davies (2015)[32] reported a person-centred approach to enhance quality in a Community Options Program in Australia. The program captured survey data from users exiting the service to map their journey, and a working group including consumer representatives assessed this against the risk register to develop new guidelines for best practice. Staff were provided educational sessions to raise awareness of the resulting 25 guidelines and case management tools and their practice was audited against these using key quality indicators and measures.[32]

II. Reporting of quality data

Four studies outlined approaches for benchmarking care nationally or contributing to publicly-reported datasets as strategies to identify variation that may be problematic and incite change.[11, 34-36] Such approaches were included in this review when they incorporated steps to address variation by providing feedback to service providers about the variations arising in their care compared to benchmarks. Eagar et al (2010) described a national system called the Palliative Care Outcomes Collaboration (PCOC) to measure outcomes and quality of palliative care services and benchmark services across Australia. A PCOC quality improvement facilitator met with the services in the collaboration to embed the collection of standardised clinical assessment into practice to improve care quality, in addition to national benchmarking meetings being held. Data on whether the approach was successful in reducing variation or addressing unwanted variation was not reported.[35]

The role of national quality registries in quality improvement was explored in one study.[11] The authors explored the use of quality registry data by heads of clinics and physicians in quality improvement activities as a strategy to address variation. The findings indicated that national quality registries can provide data that, when used in feedback to staff, can provide the basis for identifying and discussing variations and appropriate responses. Use of national quality registries varies widely and these are not routinely incorporated in efforts to address variation.[11] Similarly, Grey et al (2014)[36] explored the presentation and interpretation of the Atlas of Healthcare Variation in New Zealand for frontline quality improvement to understand and target variation. Stakeholders reported using funnel plots to enable practitioners to benchmark against peers and identify areas of variation for scrutiny. This benchmarking provides the basis for quality improvement activities to address variation.[36]

A US study by Abdul-Baki et al (2014)[34] reported that public reporting as an intervention in itself was associated with an increase in adenoma detection rates in a private endoscopy practice. The study investigators suggested that even at the broadest level, providing feedback data may improve care quality and reduce variations. However, the mechanism by which this mode of feedback may work is not established and the pre- and post-study design was not sufficiently sensitive or controlled to determine causation. On a smaller scale, in a secondary analysis, Das et al (2008)[37] reported that involvement in a trial to capture data on the quality and improve the management of Barrett's oesophagus through surveillance also led to reduced variation in practice.

Category 2 - Feedback

I. Local or individual quality data feedback

Four studies included in this review involved data about the practice of individuals or teams being captured and reported back at local level within an organisation, organisational unit or individuals. In these studies, feedback was provided without facilitation. A more substantial body of literature reported below details studies that went beyond simply providing variation data at a local level. Individual provider reports were explored in two studies.[38, 39] In a study by Stafford (2003)[39], primary care providers were provided with data over a nine-month period comparing their use of ECG compared to their peers. The findings showed a reduction in variation in the ordering of ECGs and their use after the intervention period. McFadyen et al (2015)[38] reported on two indicators that were supported by evidence-based guidelines to encourage behaviour change and improve quality through reduced unwarranted variation. Individual feedback increased appropriate treatment on one indicator but did not impact the other over the study period. A key finding was that the physician group (urological surgeons) within the hospitals that did not show improvement on one of the indicators also had the poorest attendance at the engagement sessions held before and during the project.[38]

Chart review was used in a study by Kelly et al (2016)[40] to establish adherence to the local treatment pathway for the management of atrial fibrillation with rapid ventricular response (AFRVR). Local teams made emergency departments aware of their adherence levels and best practice guidelines leading to a substantial increase in adherence to the pathway from 8–68%. Local monitoring was also used in the study by Smith et al (2013)[41] to review variation data in cardiac surgical procedures and identify where change was required. Regular monitoring of quality data enabled early detection of variations and action to be taken as required. In primary care, Gaumer et al (2008)[42] developed an information system 'Feedback and Analytic Comparison Tool' to enable clinicians to monitor their own quality data and act accordingly. This system purely provided feedback to allow clinicians to identify practice variations but did not utilise health information technology to identify the feedback that warranted action.[42] Another study explored provision of data across a network.[43] A cancer primary care network in the UK identified clinical audit and the provision of risk assessment tools as two of four QI approaches for reducing variation. While the impact of clinical audit feedback alone was not established in isolation to the other quality improvement activities, a

significant increase of 29% in referral rates was reported across the participating general practices.[43] In the context of cancer networks, clinicians felt better supported to sustain improvement efforts when there was effective leadership marked by organisational stability and consistent messaging.[43]

Category 3 – Facilitated feedback methods and models

I. Quality improvement to address process variation

Quality improvement projects were the largest group of studies identified in the database search. Twelve quality improvement projects were retrieved from the search, most of which identified process variation and then utilised educational approaches to change clinician behaviour.[15, 16, 44-53] A range of methods was used to inform the facilitated feedback in such studies including clinical algorithms, the theoretical domains framework for behaviour change and health information technology.[44, 46, 47, 52] In their narrative review, Tomson and van der Veer (2013)[53] detailed a range of local and national projects that utilise evidence-based guidelines to support QI initiatives to address unwarranted variation. They reported that the projects that saw reductions in problematic variation and enhanced quality were local level QI projects that engaged a package of clinical actions to achieve the improvement aim. The authors highlighted the inefficiency of a multitude of local level projects and the potential value, but also discussed the challenges of national or collaborative approaches. A central difficulty identified in this review is the completion of such QI initiatives as an additional activity to routine clinical work[53] These findings are reflected in several studies that presented QI approaches including feedback to address variation, as detailed here.

At the simplest level, a process such as that reported by Lee at al (2016) was used, in which a random selection of medical records was audited against 15 quality measures for inflammatory bowel disease, and then reaudited after an educational session in which the quality measures and performance against these was reviewed. Lee et al identified a positive correlation between the intervention and compliance with the quality measures, with compliance increasing by 16%.[48] Two studies progressed this approach by developing algorithms for a range of evidence-based practices as the basis for determining compliance. ALMohiza et al (2016)[44] reported a 16-week quality improvement project among physical therapists working in rehabilitation services in the US. A clinical treatment algorithm was developed to determine evidence-based effective practices and deviation from these was considered 'non-compliant', indicating problematic variation. Following a behavioural intervention program including a webinar, test and competency training, adherence to the processes identified as effective by the clinical algorithm was assessed and improved by 5–10%. Over-utilised treatments reduced by 16% post-intervention.[44] Similarly, Caterson et al 2015[46] reported the development and use of a Standardised Clinical Assessment and Management Plan (SCAMP) in plastic surgery with a decision-tree algorithm. Adherence to the SCAMP algorithm was used to identify variation and direct quality improvement efforts to address this.[46] Key performance indicators were used by Griffths and Gillibrand (2017) to identify variations in individual practice and report this back alongside a quality improvement project.[16] The project included implementing four checklists based on evidence-based guidelines along with a weekly training event to try to reduce variations in pathology practices. The project isolated the effect of the intervention from the training component and established that utilising a checklist alone was associated with conforming to the evidence-based approach, rather than the addition of the training component.[16] Having the checklist available at the point of dissection was critical.

A measurement and education project was reported by Deyo et al (2000)[47] with the US Institute of Healthcare Improvement to address variations in care for lower back pain across 22 participating organisations including health plans and medical centres. Those organisations and services with "outlier" rates of imaging or referral (identified as statistical outliers from the normal range of imaging or referral in each organisation) were used to identify clinics or physicians for targeted intervention.[47] The intervention program including three learning sessions, focusing on areas of practice variation identified by the

participating organisations from their own data, in addition to a final national congress. Participants worked within their own teams to problem-solve and then across teams from other organisations. A key component of the process was to for services to present their clinical variation data and perform continuous repeated measurements to track change in variations. Findings suggest that the approach was effective in reducing unwarranted variations, although outcome measures used to assess variation were different across the participating sites based on their clinical goals and data sources. Reduced variations were identified in outcomes such as levels of X-rays ordered, prescribed bed-rest and also 100% increase in the use of patient education materials that may also work to address unwarranted variations.[47]

Dorfsman et al (2018) utilised variations from guideline-based care in the organisation's emergency medicine departments to develop monthly educational sessions for residents working in that department. The sessions explored the evidence base for a particular practice and variation, expert discussions on areas in which the evidence base was not conclusive regarding effective care and encouraged debates among residents attending.[54] Findings did not establish whether the training addressed unwarranted variations or changed behaviour, but 77% of the 31 residents surveyed indicated that the sessions aided their understanding of why clinical practice variations may occur.[48] A network education model was reported by Nguyen et al (2007)[50] as a strategy to reduce unwarranted variation in dialysis using arteriovenous fistula (AVF). Forty-six facilities contributed to four targeted regional workshops that explored the root causes of low AVF rates by interviews with vascular surgeons, nephrologists, dialysis staff, and interventional radiologists. The analysis identified three key barriers to a higher AVF rate: 1) Failure of nephrologists to act as vascular access team leaders; 2) Lack of AVF training for vascular access surgeons, including vessel assessment skills, vein mapping, and complex surgical techniques and 3) Late referral of chronic kidney failure patients to nephrology. A literature review was then conducted to identify best demonstrated practice regionally and the strategies successfully used by this team were included in the quality improvement project. Four intervention workshop meetings were held and intervention site participants took away follow-up materials to address the content locally. Of the 35 attending physicians, 91% reported that they had changed their practice to address variations based on the intervention in consistent areas relating to AVF use over the five-year period in which outcome data were collected.[50] Similarly, Nordstrom et al (2016)[51] reported on the impacts of a learning collaborative among 28 physician practices that collected and reported on their quality improvement data through four sessions, in addition to didactic lectures, case presentations and discussion of practice-improvement strategies to reduce variation in the provision of buprenorphine. Findings indicated that there was a substantial reduction of up to 50% in variations across all seven quality measures.[51] A collaborative in urological surgery adopted a facilitated feedback approach with performance feedback and review in relation to clinical guidelines.[49] The authors reported that the urological collaborative demonstrated substantial reductions in variations in practice patterns and guideline adherence following the feedback intervention.[49]

Two studies explicitly reported on the use of health information technology (HIT) to facilitate feedback approaches.[15, 45] Baker and Newland (2008)[45] reported findings of a project to reduce variation in the care process for cardiac surgical patients that compared no QI data with automated QI data alone, and automated QI data with implementation of a continuous quality improvement project. This study pulled together the use of health information technology, quality reporting and improvement interventions. Adherence to protocol and reduction in practice variation was enhanced in the automated feedback program but optimised by the use of a CQI approach.[45] Dykes et al (2011)[15] incorporated an automated care pathway in the electronic medical record into an intervention to enhance care for stroke patients that included providing evidence to clinicians and patients, a self-management tool and discipline-specific feedback regarding guideline adherence. The study reported that point-of-care evidence enhanced adherence to guidelines including those around patient self-management education in stroke care.[15]

II. Health information technology (HIT)

Whle health information technology was identified in several studies as part of the approach to identifying variation, the review identified seven studies that focused on HIT methods for identifying variation that warrants action.[42, 55-58] Brattheim et al (2011)[56] sought to develop a process support mechanism harnessing HIT to identify variation that requires action, and reported a case study of surveillance and qualitative work in vascular surgery. While much of the observed variation was intended, they indicated that HIT systems may be utilised to undertake risk analyses and mitigate risks associated with planned activities by operating a schema of a healthcare system rather than healthcare professionals as actors within the system.[56] Ghaffarzadegan et al. (2013)[59] progress the concept of predicting risk of variation using health IT in developing a system dynamics simulation model. The learning-based simulation model utilises behavioural theory and applies this to physicians' past practice to make predictions about their likelihood of adopting particular practices, in this case caesarean delivery. Using the simulation approach, the model can then predict practice variation across obstetricians, assuming that variations are based on physician behaviour rather than patient factors.[59]

Three of the studies include in the review outlined the use of HIT clinical decision support tools explicitly as tailored feedback approaches to reduce unwarranted variations, progressing a thread seen within many of the reported quality improvement projects. Two studies reported on the use of clinical decision support tools to optimise the appropriate use of imaging for lower back pain.[55, 59] Ip et al (2014)[58] reported on use of a clinical decision support intervention targeting magnetic resonance imaging (MRI) for low back pain, which incorporated two accountability tools. The first tool was mandatory peer-to-peer consultation when test utility was uncertain, and the second was providing quarterly practice variation reports to providers, an approach that links back to those identified in category two described above. The multi-faceted intervention demonstrated a 32–33% decrease in the use of MRI for any body part, indicating that this approach could address unwarranted variation relating to overutilisation.[58] Min et al (2017)[55] embedded a point-of-care checklist in the computerised entry form for image ordering, in addition to a patient education program involving summary document explaining when medical imaging is necessary being included in the lower back pain pamphlet. Post-intervention, the median proportion of lower back pain patients who received an imaging order reduced by 5% and the median decrease in image ordering among the 43 emergency department physicians in the study was 13%.[55]

Cook et al (2014)[57] utilised HIT to develop a mechanism for determining pre-operatively those patients for whom a standardised care pathway would be appropriate for their cardiac surgical care. Post-operatively the patients on the standardised pathway continued this pathway in ICU and then within the Progressive Care Unit. For those remaining on the pathway, an electronic protocol triggered the removal of the bladder catheter; therefore, practice variation in the time to remove a catheter for those on the pathway should be minimal. The electronic decision tool was complemented by quality improvement methods including educational reinforcement and procedural training around catheter removal, and performance reports provided back to staff at one, three and six-month intervals. Findings indicated that 97% compliance with guidelines was achieved in relation to timing of removal of the catheter, suggesting that the decision support tool contributed to reducing unwarranted variation.[57]

III. Shared decision-making

Shared decision-making was identified in many articles as important for reducing problematic variation but this phenomenon was only studied explicitly in one article.[60] The study by Brabers et al [60] explored whether shared decision-making reduced medical practice variations in the choice of either single or double embryo transfer after in vitro fertilisation (IVF). The secondary analysis of a randomised controlled trial reported on the impact of a shared decision-making intervention that comprised a decision aid, support of an IVF nurse and offer of reimbursement for an extra treatment cycle among 222 couples waiting for IVF. The findings

revealed lower variation in the choice of single vs double embryo transfer after IVF in the intervention hospitals compared to the control hospitals. However, variation within hospitals increased among the intervention hospitals ,and the role that shared decision-making played in influencing the level of variation identified was not distinguished.[60]

IV. Grey literature

Several relevant case studies were identified across a range of websites detailed in Appendix 3, including the International Consortium for Health Outcomes Measurement (ICHOM) and The King's Fund website, which highlight critical issues with relevance to this project. The case studies identified do not present empirical evidence with the rigour required for a peernreviewed journal publication. In many cases, the project identified aimed to address variation, but the approach or role of facilitated feedback or continuous quality improvement in achieving this were not detailed. However, the learnings highlighted through the cae study descriptions offer considerable practical insights into how uwarranted clinical varation, as both a principle and process, can promote quality improvement at local and macro levels of health systems. Two review authors reviewed and discussed the case studies.

The key learnings offered from consulting the grey literature were distilled down to the three interlinked principles as outlined below, and are exemplified through the three selected case studies that follow.

The three principles in addressing variation derived from the grey literature are:

- Definitions of, and data regarding, variation should be developed collaboratively to foster a shared language of the issue. All healthcare stakeholders, including clinicians and consumers, should be involved in this process to ensure that datasets of variation, whether focused on local or national levels, include meaningful metrics that are predominantly focused on clinical outcomes.
- The value of variation data is largely determined by how effectively it is framed and ultimately used to drive behaviour change. Ideally, data demonstrating potential unwarranted variation should provide a tool that facilitates clinician and consumer reflection and empowerment.
- Authentic, collective ownership over the process of identifying and using data on uwarranted clinical variation for quality improvement purposes presents a significant opportunity for health services and health systems internationally, despite the human and financial resources required to undertake these tasks effectively.

The database material predominantly identified feedback approaches occurring at an organisational level. This included a report by the Health Quality and Safety Commission New Zealand that reviewed approaches to addressing unwarranted clinical variation, including forms of feedback and facilitated feedback relevant to this review.[61] The report highlighted the use of benchmarking, clinical practice guidelines, patient engagement and concurrent rounds with immediate intervention by a quality improvement analyst among the common strategies utilised in a case study of five US hospitals. The use of HIT to enhance connectivity between information sources regarding patients, but also to analyse and feed back performance data in order to enhance quality, was also featured in the report. The included articles highlighted the notion of shared decision-making (that includes empowering patients to make informed decisions) as a critical strategy to address clinical variation in the context of preference-sensitive and supply-sensitive care, using the Dutch health system as one example regarding supply-sensitive care. Supply sensitive care includes everyday care used by people with acute and chronic conditions (e.g. referrals, prescription of drugs, tests). Primary care providers and patients play a significant role in the frequency of utilisation of such care. In the Dutch system, general practitioners play a central role in operating a bundled payment system for chronic diseases as a process of managed competition. In this process, insurers pay a single fee to a 'care group' to cover a full range of care for a fixed period, based on national guidelines. While it is not a feedback-based approach, this model is significant as a strategy to address the supply-sensitive component. As with much of the grey material identified, no evidence of its impact on reducing clinical care variation was available.[61]

Case study 1: The Santeon Hospitals: Collaborating for value

The first case study – Collaborating for Value: the Santeon Hospitals in the Netherlands – is derived from the ICHOM website and exemplifies the three principles outlined above, and their holistic linkage.[62] It describes a large-scale program that evolved from a system using retrospective data to assist clinical and management decisions, into one that instead utilises real-time and highly relevant data to promote improvement cyclically at a local level. Key lessons exemplified by and explicitly noted within this case study are:

- It is more effective and efficient to commence such programs with existing or easily accessible data.
- Multidisciplinary teamwork and trust are necessary to foster agreement upon how unwarranted clinical variation data will be stored, reported and used.
- Regular multidisciplinary meetings are needed to collectively interpret the data and agree upon the improvement actions it motivates.
- Consumer involvement is required in every step of the process to correctly prioritise the issues examined and improvement activities subsequently undertaken.
- A pragmatic approach to analysis of data on unwarranted clinical variation encourages more effective and efficient quality improvement programs.
- Mutual accountability retains tangible involvement in and respect for the program over time and across different organisational units and services.

Case study 2: National Health Service: Getting It Right First Time (GIRFT)

The second case study was derived from the website of The King's Fund and details a program called 'Getting It Right First Time' (GIRFT) within the English NHS.[63] GIRFT is a national review program led by frontline clinicians in a range of disciplinary areas. It was piloted in orthopaedic surgery and has progressed to 30 medical specialities. The review process utilises a range of data including outcomes and costs of care in a process of peer review. A peer-to-peer review process is conducted with each set of quality data to identify the variations that warrant change and the change required. This process provides a nuanced strategy for determining unwarranted variation that is sufficiently adaptive to enable appropriate response in light of the latest available evidence. The initial review of the program by The King's Fund reported that:

- Peer-to-peer review programs such as GIRFT have substantial potential to reduce unwarranted variation and the associated costs but require an environment that facilitates success.
- Key features of an environment that facilitates successful programs are clinician engagement to be open and collegiate in responding to variation data, but also managerial engagement and action.
- Managerial support is critical to enable sufficient time to be dedicated to reviewing and identifying action areas from variation data. Some organisations are therefore likely to be more successful than others in realising the potential value of a program such as GIRFT based on clinician engagement and managerial support.

Case study 3: National Health Service Atlas of Variation

The third case study was drawn from a series of explorations of the response of English Primary Care Trusts to the NHS Atlas of Variation, as detailed in a number of sources, including Schang et al. (2014).[61, 64] The importance of this case study is that it is focused specifically on providing insights into whether and how unwarranted clinical variation engenders action at a broad, health systems level. A framework is provided to illustrate the prerequisites and pathways for using geographic data on unwarranted clinical variation to drive health system improvement.[61] At a macro-level, both the NHS Atlas of Variation and those of other countries (e.g. Spain, Australia) can be appreciated for their capacity to focus quality improvement attention, at both a local and systems level, on specific clinical areas and processes.

Despite the NHS Atlas of Variation being perceived as one of the more rigorous such atlases internationally, the case study shows that its utility for promoting quality improvement at either the local or system level

remains somewhat limited. This is predominantly due to a lack of awareness of its existence, and the commonplace view that its usefulness for revising local policies and clinical process decisions is questionable.[64] The same variables responsible for 'effective' use of unwarranted clinical variation data to drive quality improvement in the above two case studies were also largely present in this case study i.e. the need to engage senior clinicians; mutual agreement on responsibilities to address unwarranted clinical variation where identified; and the importance of understanding its underlying causes, rather than merely being aware of its existence, in order address it. However, an additional issue identified in this case study was the value of the NHS Atlas of Variation in providing a simple visual tool to drive cross-stakeholder engagement. This aligns with the increasing use of quality benchmarking dashboards in health services and systems, further emphasising the need for data on unwarranted clinical variation to be presented appealingly and simply, in order to render it meaningful and spur emotional commitment to change existing practices.

Discussion and synthesis of findings

This rapid review sought to determine the methods and frameworks that have been used to determine when variation in healthcare is unwarranted and change is justified, and the methods and models of facilitated feedback to respond to such variation. A review update was conducted to build on an existing report of the methods and frameworks that have been used to determine when variation in healthcare is unwarranted and change is justified.[6] The update identified eight additional studies in the peer-reviewed published literature that provided further evidence for the predominant use of the Wennberg Framework to classify variation in three ways (effective-care; patient-preference and supply-sensitive), in addition to the use of statistical approaches to identify variations, generally applied to areas in which effective-care is established. When exploring methods and models of facilitated feedback, the review identified 32 studies in the peer-reviewed published literature and a range of grey literature on this matter, as well extracting and detailing three relevant case studies. Here, we discuss the findings in relation to each review question.

Question 1: What methods and frameworks have been used to determine when variation in healthcare is unwarranted and change is justified?

Our findings reiterate that capturing variation data across a health system is a critical first step to identifying problematic variations arising in health care. Atlases of variation have been published internationally in recent years to address this issue, demonstrating the variations arising geographically across each country.[2, 65-67] Although a range of frameworks exist, work to determine the variation that requires action and change is largely based on categorising variation into the three groups established by Wennberg and colleagues. Statistical methods are used to distinguish deviations from effective care. Where optimal (effective) care can be determined, there is evidence to suggest that deviation from guidelines and care pathways can be used to determine unwarranted clinical variation.[24, 27, 28, 31] In the context of supply-sensitive variations in care, econometric analyses have been undertaken that demonstrate variations in utilisation rates in different settings, but do not provide an understanding of where variations are problematic.[25] Two of the studies discussing framework approaches highlighted the substantial contribution of patient preferences to variation data. These indicated that approaches to identify and address variation that solely focus on effective care are not sufficient to determine the care that is problematic and warrants change, unless the role of patient preferences is fully understood.[26, 29]

Question 2: What methods and models of facilitated feedback and facilitated continuous quality improvement have been shown to be effective in addressing unwarranted variation in healthcare?

Responses to clinical variation operate at a number of levels, from presenting evidence to facilitated tailored feedback or quality improvement approaches. Twenty studies in the peer-reviewed literature detailed models of CQI or feedback that incorporated facilitation, many more examples are apparent in the wider grey literature. Extending the arguments made by Mercuri and Gafni (2018), feedback to care providers (individually or in aggregated data) about the nature of variation arising in their care provision may provide a nuanced approach to care variation, and where appropriate, can enable exploration of deviations from effective care in the context of patient-preferences.[29] This review identified evidence to suggest that simply providing feedback about performance against quality indicators or evidence-based practice, operationalised in guidelines or pathways, was associated with changing clinician behaviours and reduced

variations in practice.[34, 35, 37] While providing feedback alone can, in some circumstances, encourage reflections and improvement actions, clinical variation data that is tailored to particular health professionals, services or systems, and provided to these audiences via facilitated feedback processes may have greater capacity to drive large-scale change. Many examples of teams, health facilities or networks of health facilities that provided facilitated feedback were identified in the peer-reviewed literature, with facilitated feedback defined broadly as feedback that is focused on bringing about change.[16, 44-53] Several models of facilitated feedback were identified that linked to two broad categories: local or organisation-wide quality improvement feedback and shared decision -making.

Most published literature reported local or organisation-wide quality improvement, often through team training programs. Health information technology was the principal method for capturing and, in some cases, reporting variation data back to facilitate change.[55, 56, 58, 59] HIT was central to continuous quality improvement projects that occurred in teams or organisations, for example through generation of clinical treatment algorithms and automated generation of quality indicators to drive or contribute to the feedback sessions.[15, 45] Outcomes that were assessed in facilitated feedback and enabled continuous quality improvement approaches included the detection rates for the conduct of medical technologies, reduced overuse of technologies or treatments, changes in patient clinical outcomes and adherence to practice protocols.[55, 57, 58]The increasing availability of HIT and real-time analytics in health services internationally makes it likely that the relationship between HIT and clinical variation data and subsequent behaviour change will only continue to strengthen over time.

The peer reviewed literature demonstrates variability in approaches and there does not appear to be a single preferred model for structuring facilitated feedback in the context of clinical variation. Furthermore, there is no evidence to suggest a particular model is more or less effective than others. Our review indicates that facilitated feedback is considered to provide the necessary structure to direct improvement efforts, but the included studies do not provide empirical evidence that this is the case over and above feedback without facilitation, or simply presenting evidence at the point-of -are in the context of clinical care variation. The above approaches were exemplified in the grey literature, with two strong examples selected for inclusion here.[62, 64] The grey literature demonstrates widespread recognition and perceived face validity of facilitated feedback methods and approaches.

Facilitated feedback is widely accepted as an optimal feedback approach to enhance performance in clinical practice.[68, 69] This review however indicates there is a currently a lack of evidence to determine the effectiveness of facilitated feedback over other basic feedback approaches, or any particular facilitated feedback model or method, in the specific context of addressing variation. There is a possibility that some facilitated feedback approaches may have unintended consequences on patients and clinicians. For example, public reporting of clinical-level outcome data may lead to 'data gaming' and risk aversion behaviour by clinicians. It may also lead to socioeconomic variations in pateints accessing well-performing clinicians and hospitals.[70, 71]

Most approaches identified in the peer reviewed and grey literature for responding to variation and reducing unwanted variations focus solely or predominantly on variations in clinicians' practice.[29] Such approaches indicate an assumption that most variation is due to clinician choice rather than patient preferences or patient factors, and are poorly aligned with the move towards patient-centred care.[72] Mercuri and Gafni (2018)[29] highlight a range of evidence that indicates only around 5–10% of variations relate to physician choice. The review also identified a lack of studies that examined the impact of decisions that were based on deviations from guidelines (e.g. limiting MRI ordering rights for GPs) in terms of cost and care improvements. Decreased variations in clinical practice actions may potentially impact on patient care and lead to the establishment of false economies. For example, the ordering of MRIs by specialists, which result in patients paying to see the specialist for two consults; an initial consultation and then follow-up with the MRI report. Fewer costs are incurred if patients attend a single specialist consult, because the MRI has been

completed with the GP. Current literature does not provide data to account for the reasons for implementing deviation from guidelines as a method of controlling unwarranted clinical varation, which may include costs, clinical benefit/care imperative, reduction in risk exposure, moral/ethical reasons, political reasons or all or some of these factors combined.

Shared decision-making (SDM) was discussed in the literature as a model for reducing unwanted clinical variations but was only evaluated in one study relating to decision-making on one component of IVF care.[60] The role of SDM in the level of variation identified was not distinguishable in this study. The concept of SDM as a strategy to respond appropriately to and reduce unwarranted variation is appealing and supported in wider literature.[73] However, the review findings indicate there is a lack of sufficient evidence of the impact of SDM in the context of reducing variations or responding appropriately to these to determine if the model is effective. Further evidence is required before SDM could be recommended in policy guidance for this purpose. Initiatives such as the ACSQHC national patient-reported outcomes program may in future provide an avenue for integrating patient preferences efficiently into clinical variation algorithms to highlight possible unwarranted clinical variation in real-time and provide notification to healthcare providers. Notable across the review findings was the lack of analysis regarding the sustainability of facilitated feedback approaches and the outcomes achieved through individual projects.

Applicability

In developing a framework for determining when variation is unwarranted and how facilitated feedback and facilitated continuous quality improvement can be best utilised to address it, the evidence synthesised in this rapid review suggests Cancer Instititue NSW should consider the following:

- There is evidence that quality indicators, guidelines and optimal care pathways provide a set of parameters from which unwarranted variation can potentially be determined, in circumstances in which there is an optimal approach. Evidence also supports the perceived value of quality indicators, guidelines and optimal care pathways to identify variation overall and support the provision of feedback to care providers when no optimal approach is agreed.
- Providing feedback to clinicians is identified across a range of a settings as being associated with changes in variation such as reducing overuse of tests and treatments, reducing variations in optimal patient clinical outcomes and increasing guideline or protocol adherence.
- Feedback approaches that relate to performance indicators may address variations arising due to clinicians' behaviours, but may not necessarily address variations that relate to patient preferences.
- Facilitated approaches to providing feedback including shared decision-making are widely used in health services internationally to provide a nuanced and continuous improvement approach to respond to variation. However, there is no evidence to suggest that facilitated approaches as a group, or a particular facilitated feedback model or method is more effective in responding to variations appropriately than simply providing feedback to individuals, teams or networks of health providers.
- Evaluation of the effectiveness of approaches utilising facilitated feedback is needed to provide evidence on two questions: firstly, regarding whether facilitated feedback offers advantages over feedback without facilitation in the context of addressing variation; and secondly to determine if there is an optimal model and/or method of facilitation that is more likely to create change where needed.

References

1. Corrigan J, Donaldson M, Kohn L, Maguire S, Pike K. Crossing the quality chasm. A new health system for the 21st century. Washington, DC: National Academy of Sciences, National Academies Press; 2001.
2. DaSilva P, Gray JA. English lessons: can publishing an atlas of variation stimulate the discussion on appropriateness of care? Med J Aust. 2016;205(10):S5-s7.
3. Australian Commission on Safety and Quality in Health Care. The second Australian atlas of healthcare variation. ACSQHC; 2017 9781925224825. Available from: https://www.safetyandquality.gov.au/atlas/atlas-2017
4. Buchan H. Gaps between best evidence and practice: causes for concern. Med J Aust. 2004;180(6 Suppl):S48-9.
5. Gray M. Value based healthcare. BMJ. 2017;356:j437.
6. Harrison R, Manias E, Mears S, Heslop D, Hinchcliff R, et al. Addressing unwarranted clinical variation: A rapid review of current evidence. J Eval Clin Pract. 2019;25(1):53-65.
7. Mercuri M, Gafni A. Medical practice variations: what the literature tells us (or does not) about what are warranted and unwarranted variations. J Eval Clin Pract. 2011;17(4):671-7.
8. Wennberg J, Gittelsohn A. Variations in medical care among small areas. Sci Am. 1982;246(4):120-34.
9. Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. BMJ. 2002;325(7370):961-4.
10. Duggan A, Koff E, Marshall V. Clinical variation: why it matters. Med J Aust. 2016;205(10):S3-s4.
11. Fredriksson M, Halford C, Eldh AC, Dahlstrom T, Vengberg S, et al. Are data from national quality registries used in quality improvement at Swedish hospital clinics? Int J Qual Health Care. 2017;29(7):909-15.
12. Wilcox N, McNeil JJ. Clinical quality registries have the potential to drive improvements in the appropriateness of care. Med J Aust. 2016;205(10):S27-s29.
13. Australian Commission on Safety and Quality in Health Care. Framework for Australian clinical quality registries. Sydney Australia: ACSQHC; 2014, March. Available from: https://www.safetyandquality.gov.au/wp-content/uploads/2014/09/Framework-for-Australian-Clinical-Quality-Registries.pdf
14. Department of Health Payment by Results Team. A simple guide to Payment by Results. United Kingdom: National Health Service; 2012. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/213150/PbR-Simple-Guide-FINAL.pdf
15. Dykes PC, Acevedo K, Boldrighini J, Boucher C, Frumento K, et al. Clinical practice guideline adherence before and after implementation of the HEARTFELT (HEART Failure Effectiveness & Leadership Team) intervention. J Cardiovasc Nurs. 2005;20(5):306-14.
16. Griffiths M, Gillibrand R. Use of key performance indicators in histological dissection. J Clin Pathol. 2017;70(12):1019-23.
17. Public Health England. National Cancer Registration and Analysis Service. 2018. Available from: http://www.ncin.org.uk/publications/reports/
18. Brown DL, Epstein AM, Schneider EC. Influence of cardiac surgeon report cards on patient referral by cardiologists in New York state after 20 years of public reporting. Circ Cardiovasc Qual Outcomes. 2013;6(6):643-8.
19. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J. 2009;26(2):91-108.
20. Varker T, Forbes D, Dell L, Weston A, Merlin T, et al. Rapid evidence assessment: increasing the transparency of an emerging methodology. J Eval Clin Pract. 2015;21(6):1199-204.
21. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-9, w64.
22. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, et al. Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme. Lancaster: Institute of Health Research, University of Lancaster; 2006. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.3100&rep=rep1&type=pdf
23. Sirriyeh R, Lawton R, Gardner P, Armitage G. Reviewing studies with diverse designs: the development and evaluation of a new tool. J Eval Clin Pract. 2012;18(4):746-52.
24. Bottle A, Mariscalco G, Shaw MA, Benedetto U, Saratzis A, et al. Unwarranted Variation in the Quality of Care for Patients With Diseases of the Thoracic Aorta. J Am Heart Assoc. 2017;6(3)
25. de Vries EF, Heijink R, Struijs JN, Baan CA. Unraveling the drivers of regional variation in healthcare spending by analyzing prevalent chronic diseases. BMC Health Serv Res. 2018;18(1):323.
26. Feufel MA. How to Uncover Sources of Unwarranted Practice Variation: A Case Study in Emergency Medicine. Qual Health Res. 2018;28(9):1486-98.
27. Johansson N, Jakobsson N, Svensson M. Regional variation in health care utilization in Sweden - the importance of demand-side factors. BMC Health Serv Res. 2018;18(1):403.

28. Mayer M, Naylor J, Harris I, Badge H, Adie S, et al. Evidence base and practice variation in acute care processes for knee and hip arthroplasty surgeries. PLoS One. 2017;12(7):e0180090.
29. Mercuri M, Gafni A. Examining the role of the physician as a source of variation: Are physician-related variations necessarily unwarranted? J Eval Clin Pract. 2018;24(1):145-51.
30. Partington A, Chew DP, Ben-Tovim D, Horsfall M, Hakendorf P, et al. Screening for important unwarranted variation in clinical practice: a triple-test of processes of care, costs and patient outcomes. Aust Health Rev. 2017;41(1):104-10.
31. Soden PA, Zettervall SL, Curran T, Vouyouka AG, Goodney PP, et al. Regional variation in patient selection and treatment for lower extremity vascular disease in the Vascular Quality Initiative. J Vasc Surg. 2017;65(1):108-18.
32. Davies DJ. Quality improvement initiatives in a case management service: case study. Aust J Prim Health. 2015;21(1):14-21.
33. Cammisa C, Partridge G, Ardans C, Buehrer K, Chapman B, et al. Engaging physicians in change: results of a safety net quality improvement program to reduce overuse. Am J Med Qual. 2011;26(1):26-33.
34. Abdul-Baki H, Schoen RE, Dean K, Rose S, Leffler DA, et al. Public reporting of colonoscopy quality is associated with an increase in endoscopist adenoma detection rate. Gastrointest Endosc. 2015;82(4):676-82.
35. Eagar K, Watters P, Currow DC, Aoun SM, Yates P. The Australian Palliative Care Outcomes Collaboration (PCOC)--measuring the quality and outcomes of palliative care on a routine basis. Aust Health Rev. 2010;34(2):186-92.
36. Grey C, Wells S, Exeter DJ, Hanham G, Zhao J, et al. Stakeholder engagement for the New Zealand Atlas of Healthcare Variation: cardiovascular disease secondary prevention: VIEW-3. N Z Med J. 2014;127(1400):81-91.
37. Das D, Ishaq S, Harrison R, Kosuri K, Harper E, et al. Management of Barrett's esophagus in the UK: overtreated and underbiopsied but improved by the introduction of a national randomized trial. Am J Gastroenterol. 2008;103(5):1079-89.
38. McFadyen C, Lankshear S, Divaris D, Berry M, Hunter A, et al. Physician level reporting of surgical and pathology performance indicators: a regional study to assess feasibility and impact on quality. Can J Surg. 2015;58(1):31-40.
39. Stafford RS. Feedback intervention to reduce routine electrocardiogram use in primary care. Am Heart J. 2003;145(6):979-85.
40. Kelly AM, Pannifex J. A Clinical Network Project Improves Care of Patients with Atrial Fibrillation with Rapid Ventricular Response in Victorian Emergency Departments. Heart Lung Circ. 2016;25(3):e33-6.
41. Smith IR, Gardner MA, Garlick B, Brighouse RD, Cameron J, et al. Performance monitoring in cardiac surgery: application of statistical process control to a single-site database. Heart Lung Circ. 2013;22(8):634-41.
42. Gaumer G, Hassan N, Murphy M. A simple primary care information system featuring feedback to clinicians. Int J Health Plann Manage. 2008;23(3):185-202.
43. Rubin G, Gildea C, Wild S, Shelton J, Ablett-Spence I. Assessing the impact of an English national initiative for early cancer diagnosis in primary care. Br J Cancer. 2015;112 Suppl 1:S57-64.
44. AlMohiza MA, Sparto PJ, Marchetti GF, Delitto A, Furman JM, et al. A Quality Improvement Project in Balance and Vestibular Rehabilitation and Its Effect on Clinical Outcomes. J Neurol Phys Ther. 2016;40(2):90-9.
45. Baker RA, Newland RF. Continous quality improvement of perfusion practice: the role of electronic data collection and statistical control charts. Perfusion. 2008;23(1):7-16.
46. Caterson SA, Singh M, Orgill D, Ghazinouri R, Han E, et al. Development of Standardized Clinical Assessment and Management Plans (SCAMPs) in Plastic and Reconstructive Surgery. Plast Reconstr Surg Glob Open. 2015;3(9):e510.
47. Deyo RA, Schall M, Berwick DM, Nolan T, Carver P. Continuous quality improvement for patients with back pain. J Gen Intern Med. 2000;15(9):647-55.
48. Lee AJ, Kraemer DF, Smotherman C, Eid E. Providing Our Fellows in Training with Education on Inflammatory Bowel Disease Health Maintenance to Improve the Quality of Care in Our Health Care System. Inflamm Bowel Dis. 2016;22(1):187-93.
49. Miller DC, Murtagh DS, Suh RS, Knapp PM, Schuster TG, et al. Regional collaboration to improve radiographic staging practices among men with early stage prostate cancer. J Urol. 2011;186(3):844-9.
50. Nguyen VD, Lawson L, Ledeen M, Treat L, Buss J, et al. Successful multidisciplinary interventions for arterio-venous fistula creation by the Pacific Northwest Renal Network 16 vascular access quality improvement program. J Vasc Access. 2007;8(1):3-11.
51. Nordstrom BR, Saunders EC, McLeman B, Meier A, Xie H, et al. Using a Learning Collaborative Strategy With Office-based Practices to Increase Access and Improve Quality of Care for Patients With Opioid Use Disorders. J Addict Med. 2016;10(2):117-23.
52. Tavender EJ, Bosch M, Gruen RL, Green SE, Michie S, et al. Developing a targeted, theory-informed implementation intervention using two theoretical frameworks to address health professional and organisational factors: a case study to improve the management of mild traumatic brain injury in the emergency department. Implement Sci. 2015;10:74.
53. Tomson CR, van der Veer SN. Learning from practice variation to improve the quality of care. Clin Med (Lond). 2013;13(1):19-23.
54. Dorfsman ML, Hart DE, Wolfson AB. Implementation of a Novel Conference Series on Clinical Practice Variations Provides an Opportunity for Constructive Discussion of Faculty Practice Patterns: Do as We Say? ... or Do as We Do? AEM Educ Train. 2018;2(1):15-19.
55. Min A, Chan VWY, Aristizabal R, Peramaki ER, Agulnik DB, et al. Clinical Decision Support Decreases Volume of Imaging for Low Back Pain in an Urban Emergency Department. J Am Coll Radiol. 2017;14(7):889-99.

56. Brattheim B, Faxvaag A, Seim A. Process support for risk mitigation: a case study of variability and resilience in vascular surgery. BMJ Qual Saf. 2011;20(8):672-9.
57. Cook DJ, Thompson JE, Suri R, Prinsen SK. Surgical Process Improvement: Impact of a Standardized Care Model With Electronic Decision Support to Improve Compliance With SCIP Inf-9. Am J Med Qual. 2014;29(4):323-8.
58. Ip IK, Gershanik EF, Schneider LI, Raja AS, Mar W, et al. Impact of IT-enabled intervention on MRI use for back pain. Am J Med. 2014;127(6):512-8.e1.
59. Ghaffarzadegan N, Epstein AJ, Martin EG. Practice variation, bias, and experiential learning in cesarean delivery: a data-based system dynamics approach. Health Serv Res. 2013;48(2 Pt 2):713-34.
60. Brabers AE, van Dijk L, Groenewegen PP, van Peperstraten AM, de Jong JD. Does a strategy to promote shared decision-making reduce medical practice variation in the choice of either single or double embryo transfer after in vitro fertilisation? A secondary analysis of a randomised controlled trial. BMJ Open. 2016;6(5):e010894.
61. Love TE, Ehrenberg N, Sapere Research Group. Addressing unwarranted variation: literature review on methods for influencing practice. Wellington, New Zealand: Health Quality & Safety Commission New Zealand; 2014 25 March 2014 Report No.: 0478385692. Available from: http://www.hqsc.govt.nz/assets/Health-Quality-Evaluation/PR/Variation-literature-review-on-methods-for-influencing-practice-May-2014.pdf
62. Okunade O, Arora J, Haverhals A, Niessen L. Collaborating for Value: the Santeon Hospitals in the Netherlands. Cambridge, MA: International Consortium for Health Outcomes Measurement; 2017. [Access Date. Available from: https://ichom.org/files/case-studies/Santeon_Case_Study_Final.pdf
63. Timmins N. Tackling variations in clinical care: assessing the Getting It Right First Time (GIRFT) programme. King's Fund; 2017. Available from: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/Getting_it_right_Kings_Fund_June_2017.pdf
64. Schang L, Morton A, DaSilva P, Bevan G. From data to decisions? Exploring how healthcare payers respond to the NHS Atlas of Variation in Healthcare in England. Health Policy. 2014;114(1):79-87.
65. Health Quality & Safety Commission New Zealand. Atlas of Healthcare Variation New Zealand: Health Quality & Safety Commission New Zealand; Intelligence HQ; 2019. Available from: https://www.hqsc.govt.nz/our-programmes/health-quality-evaluation/projects/atlas-of-healthcare-variation/
66. Institute for Clinical Evaluative Sciences (ICES). Atlases & Reports. Ontario Canada: 2019. Available from: https://www.ices.on.ca/Publications/Atlases-and-Reports
67. RightCare. The NHS Atlas of Variation in Healthcare: Reducing Unwarranted Variation to Increase Value and Improve Quality. National Health Service; 2011, November.
68. Royal Australian and New Zealand College of Obstetricians and Gynaecologists. Evidence-Informed Facilitated Feedback: A feedback guide for Training Supervisors conducting appraisal and assessment sessions Melbourne: RANZCOG; 2018. Available from: https://www.ranzcog.edu.au/RANZCOG_SITE/media/RANZCOG-MEDIA/Training%20and%20Assessment/Specialist%20Training/Training%20Resources/Evidence-Informed-Facilitated-Feedback.pdf
69. Sargeant J, Lockyer J, Mann K, Holmboe E, Silver I, et al. Facilitated Reflective Performance Feedback: Developing an Evidence- and Theory-Based Model That Builds Relationship, Explores Reactions and Content, and Coaches for Performance Change (R2C2). Acad Med. 2015;90(12):1698-706.
70. Scott IA, Ward M. Public reporting of hospital outcomes based on administrative data: risks and opportunities. Med J Aust. 2006;184(11):571-5.
71. Vallance AE, Fearnhead NS, Kuryba A, Hill J, Maxwell-Armstrong C, et al. Effect of public reporting of surgeons' outcomes on patient selection, "gaming," and mortality in colorectal cancer surgery in England: population based cohort study. BMJ. 2018;361:k1581.
72. Australian Commission on Safety Quality in Health Care. Patient-centred care: Improving quality and safety through partnerships with patients and consumers. Sydney: ACSQHC; 2011 0987061712. Available from: https://www.safetyandquality.gov.au/wp-content/uploads/2012/03/PCC_Paper_August.pdf
73. Australian Commission on Safety Quality in Health Care. Medical Practice Variation: Background Paper. Sydney: ACSQHC; 2013. Available from: https://www.safetyandquality.gov.au/wp-content/uploads/2013/10/SAQ110_Medical_Practice_variation_V10_WEB.pdf

Appendix 1: Search strategy

Question 1

Clinical Variation

Ovid Medline, 2017 to the present (run 28/08/18)

Below is the amended version of the search strategy.

- The updated search on clinical variation for 2017–2018 is line 22
- These results exclude paediatric articles and non-English language articles.

#	Searches	Results
1	Practice Patterns, Physicians'/	52,989
2	exp physicians/ or clinician*.af. or physician*.af. or exp medical staff/	839,680
3	exp hospitals/ or Hospitalization/ or hospitali*.mp.	502,959
4	(variation* adj2 (Clinical care or Medical care or Healthcare or health care or Medical practice or physician* or clinical or practice or clinician* or pattern*)).mp.	10,132
5	Guideline Adherence/ or Practice Guidelines as Topic/ or Healthcare Disparities/ or clinical protocols/ or organizational policy/ or evidence based*.ti,ab,kw,sh. or exp "Quality of Health Care"/	6,335,879
6	1 or 2 or 3 or 5	7,015,351
7	6 and 4	5392
8	6 and ((Regional adj2 variation*) or (geographical adj2 variation*)).mp.	4745
9	7 or 8	9960
10	limit 9 to (english language and yr="2017 -Current")	913
11	remove duplicates from 10	881
12	6 and (small area analysis or small area variation).mp.	1202
13	limit 12 to (english language and yr="2017 -Current")	35
14	remove duplicates from 13	33
15	11 or 14	907
16	4 and (regional or geographical).mp.	773
17	limit 16 to (english language and yr="2017 -Current")	98
18	remove duplicates from 17	95
19	15 or 18	958
20	19 and (unwarranted* or undesirable or inappropriate or warranted or unexplained or explained or unacceptable).mp.	111
21	exp child/ or exp infant/ or (pediatric* or paediatric* or childhood or children).af.	3,005,698
22	20 not 21	88

Appendix 2: Search strategy

Question - Feedback

Ovid Medline, 2000 to the present (run 28/08/18)

Below is the amended version of the search strategy.

- These results exclude paediatric articles and non-English language articles.
- Authors included as many feedback and quality terms as possible, lines 18–35, giving 339 hits in line 36.

#	Searches	Results
1	Practice Patterns, Physicians'/	52,989
2	exp physicians/ or clinician*.af. or physician*.af. or exp medical staff/	839,680
3	exp hospitals/ or Hospitalization/ or hospitali*.mp.	502,959
4	(variation* adj2 (Clinical care or Medical care or Healthcare or health care or Medical practice or physician* or clinical or practice or clinician* or pattern*)).mp.	10,132
5	Guideline Adherence/ or Practice Guidelines as Topic/ or Healthcare Disparities/ or clinical protocols/ or organizational policy/ or evidence based*.ti,ab,kw,sh. or exp "Quality of Health Care"/	6,335879
6	1 or 2 or 3 or 5	7,015,351
7	6 and 4	5392
8	6 and ((Regional adj2 variation*) or (geographical adj2 variation*)).mp.	4745
9	7 or 8	9960
10	6 and (small area analysis or small area variation).mp.	1202
11	4 and (regional or geographical).mp.	773
12	exp child/ or exp infant/ or (pediatric* or paediatric* or childhood or children).af.	3,005,698
13	9 or 10 or 11	11,354
14	limit 13 to yr="2000 -Current"	9076
15	limit 14 to english language	8767
16	15 not 12	6914
17	remove duplicates from 16	4015
18	17 and feedback.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	67
19	17 and facilitated.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	18
20	17 and multifaceted.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	18

21	17 and comparative performance.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	1
22	17 and "controlled before after studies".mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	3
23	17 and ((colleague* or peer*) adj3 assess*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	2
24	17 and (workplace based or work place based or work based).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	4
25	17 and facilitator.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	4
26	17 and quality improvement.af.	205
27	17 and practice improvement.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	8
28	17 and (practice adj2 improvement*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	16
29	17 and evaluation program*.af.	3
30	17 and mentor*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	5
31	17 and continuous quality.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	9
32	17 and continuous improvement.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	0
33	17 and (quality management or TQM).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	41
34	17 and cooperative behavior.af.	24

35	17 and professional development.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	6
36	or/18–35	339

Appendix 3: Search strategy (grey material)

The following organisations working in the field of health care quality and safety were contacted and their websites searched to identify relevant work, publications or programs.

1. Quality and safety in-country units Health Quality and Safety Commission New Zealand Australian Commission for Safety and Quality in Health Care Clinical Excellence Commission (NSW, Australia) Agency for Clinical Innovation (NSW, Australia)	 https://www.hqsc.govt.nz https://www.safetyandquality.gov.au http://www.cec.health.nsw.gov.au http://www.cec.health.nsw.gov.au
2. National Health Service (UK)	www.nhs.uk
3. The King's Fund (UK)	https://www.kingsfund.org.uk
4. Agency for Health Research and Quality (US)	https://www.ahrq.gov
5. Canadian Institute for Health (Canada)	https://www.cihi.ca/en
6. Institute for Healthcare Improvement (US)	http://www.ihi.org
7. Australian Institute for Patient and Family Centred Care (AIPFCC)	http://www.aipfcc.org.au/about.html
8. European Collaboration for Healthcare Optimization	https://cordis.europa.eu/project/rcn/94070_en.html
9. Institute for Clinical Evaluative Sciences (Canada)	https://www.ices.on.ca
10. Atlas of Variations in Medical Practice in the National Health System (Spain)	http://www.atlasvpm.org/en/english-version
11. Wennberg Collaborative (US)	http://wennbergcollaborative.org/
12. Michigan Urological Surgery Improvement Collaborative (MUSIC)	http://musicurology.com/
13. American College of Surgeons National Surgical Quality Improvement Program (NSQIP)	https://www.facs.org/quality-programs/acs-nsqip
14. Australian Health Services Research Institute	http://ahsri.uow.edu.au/index.html (including ePPOC: electronic persistent pain outcomes collaboration) PCOC: palliative care outcomes collaboration, AROC: Australasian rehabilitation outcomes centre)
15. International Consortium for Health Outcomes Measurement	http://www.ichom.org/measure/ (How to measure)

Appendix 4: Summaries of included studies

Table 1: Summary of included studies for Question 1

Author	Year	Country	Study Design	Setting	Sample	Aim	Approach for identifying UCV
Bottle	2017	UK	Systematic review and meta-analysis	Hospital episode statistics (HES); National Adult Cardiac Surgery Audit (NACSA) - English cardiac surgery units	24,548 patients from HES; 8058 patients from NACSA; 33 studies in the SR with 103,543 patients.	To evaluate the contribution of UCV to regional differences in outcome observed in thoracic aortic disease patients in England and identify areas of structure and process for quality improvement.	Statistical approach to determine unwarranted variation.
deVries	2018	The Netherlands	Data linkage - Health survey, usage, and claims	18 Dutch regions	10,767 patients with diabetes; 3,735 patients with depression; 44,684 general population	To describe the unadjusted regional variation in healthcare spending and explore the extent to which demand and supply factors explain regional variation in healthcare spending.	Statistical approach to determine unwarranted variation.
Feufel	2018	Germany	Case study - secondary ethnographic	Emergency departments at two mid-western hospitals in the US	3 attending physicians; 18 periods of observation	To understand and target the drivers of unwarranted practice variations using a mixed-methods approach by advancing the understanding of mechanisms underlying practice variation and evaluating and expanding the repertoire of interventions to increase the quality, equity and efficiency of practice variations.	Wennberg model

Author	Year	Country	Study Design	Setting	Sample	Aim	Approach for identifying UCV
Johansson	2018	Sweden	Observational longitudinal	21 Swedish regions (county councils) been 2001–2014.	273 region–year observations of visits to primary or to specialist physicians	To establish to what degree regional variation is explained by observed demand factors such as health, demography and socioeconomic factors.	Statistical approach to determine unwarranted variation.
Mayer	2017	Australia	Systematic literature review and prospective medical record audit	Part 2 – 19 arthroplasty hospitals (10 public, 9 private) in Australia.	Part 1 – 48 studies and 1 guideline; Part 2 – 120 surgeons	To identify interventions historically used for knee or hip arthroplasty and establish if routine use is supported by high-level evidence and whether surgeon use aligns with the evidence.	Wennberg model – not explicit
Mercuri	2018	Canada	Narrative review	N/A	No detail of the number of studies included.	Narrative review exploring if physical-related variation is problematic for patient care.	Critique of current models
Partington	2017	Australia	Data linkage – inpatient and mortality	Emergency departments at 4 South Australian hospitals	7950 patients presenting with acute coronary syndromes	To identify and assess the burden of UCV in clinical practice.	Statistical approach to determine unwarranted variation.
Soden	2017	US	Database analysis	National clinical registry	52,373 interventions - infrainguinal open bypass (31%) or endovascular (69%) 2009–2014.	To compare variation across patients undergoing infrainguinal open bypass or endovascular intervention in the Vascular Quality Initiative.	Statistical approach to determine unwarranted variation.

UCV: unwarranted clinical evaluation; N/A: not applicable

Table 2: Summary of included studies for Question 2

Author	Year	Country	Study design	Setting	Sample	Aim	Approach
ALMohiza	2016	US	Cluster randomised trial	15 outpatient neurological speciality clinics	23 physical therapists	To implement and evaluate a quality improvement initiative in neurologic outpatient practice.	Process variation
Abdul-Baki	2015	US	Pre- and post-study	1 metropolitan endoscopy centre.	17,526 colonoscopy reports	To assess whether public reporting of colonoscopy quality was associated with improvement in adenoma detection rate.	Reporting of quality measures
Baker	2008	US	Pre- and post-study	Cardiac Surgery Research database	979 cardio-pulmonary bypass patients	To demonstrate the influence of automated generation of quality indicators for cardiopulmonary bypass and the implementation of a CQI program on the process of care.	Local QI feedback
Brabers	2016	The Netherlands	Pre- and post-study	5 hospitals	222 couples waiting for IVF	To explore whether shared decision-making reduces medical practice variations in IVF.	Shared decision making
Brattheim	2011	Norway	Case study	Vascular surgery units in 3 hospitals	29 patient episode observations	To explore the characteristics and sources of process variability in surgical care.	Process variation
Cammisa	2011	US	Pre- and post-study	Healthplan database 2006-2008	34 high volume practices were visited during the intervention period	To create a guideline intervention to decrease overuse in the management of acute and chronic back pain.	Guideline
Caterson	2015	US	Methodological work	1 tertiary hospital	Impact-based reconstruction	To investigate the standardised clinical assessment and management plan concept for breast reconstruction.	Guideline

Author	Year	Country	Study design	Setting	Sample	Aim	Approach
Cook	2014	US	Pre- and post-study	1 hospital progressive care unit	86 baseline and 187 intervention surgical patients	To improve the quality of care in indwelling catheter use following surgery.	Health information technology - Decision support
Das	2008	UK	Cross sectional survey	British Society of Gastroenterology Membership	228 gastroenterologists	To provide a review of the management of Barrett's oesophagus in the UK and compare to national guidelines.	Reporting of quality measures
Davies	2015	Australia	Case study	1 aged care service	1 Community Options Case Management service in New South Wales	To develop good practice guidelines and tools to support person-centred practice.	Shared decision-making
Deyo	2000	US	Pre- and post-study	22 health organisations including 12 hospitals, insurance plans, multicentred health services and independent services.	3 team members from each of the 22 organisations	To use scientific evidence and behaviour change approaches to improve care for back pain.	Process variation
Dorfsman	2018	US	Pre- and post-study	3 emergency medicine programs in academic health centres	31 residents	To use clinical practice variations as a training tool for residents.	Process variation

Author	Year	Country	Study design	Setting	Sample	Aim	Approach
Dykes	2005	US	Pre- and post-study	1 community hospital	Pre-test sample: 90 heart failure patients over 65 and 55 control stroke patients over 65. Post-test sample: 96 heart failure patients over 65 and 75 control stroke patients over 65	To examine interdisciplinary knowledge and adherence to core recommendations before and after HEART Failure Effectiveness and Leadership Team intervention.	Local QI feedback
Eagar	2010	Australia	Conceptual	Palliative Care Outcomes Collaboration of 111 services	Benchmarking round 1: 51 services; Benchmarking round 2: 94 services	To measure the outcomes and quality of specialist palliative care services and to benchmark services on a national basis through an independent third party.	Reporting of quality measures

Author	Year	Country	Study design	Setting	Sample	Aim	Approach
Fredriksson	2017	Sweden	Cross sectional survey	78 hospitals reporting to The Swedish Registry of Gallstone Surgery and Endoscopic Retrograde Cholangiopancreatography, 71 hospital clinics reporting to the Swedish Stroke Register and 31 hospital clinics reporting to the Swedish Lung Cancer Registry	3–6 respondents from each organisation	To investigate the use of national quality registries in local quality improvement.	Reporting of quality measures
Gaumer	2008	Egypt	Case study	14 primary care clinics	NA	To develop a health information system to support quality improvement approaches to help clinicians understand practice variation.	Local QI feedback
Ghaffarzadegan	2013	US	Conceptual work	Discharge data from non-federal acute hospitals in Florida	Hospital discharges from 300 randomly selected obstetricians between 1992–2008	To develop a system dynamics simulation model of obstetricians' delivery mode decision.	Health information technology

Author	Year	Country	Study design	Setting	Sample	Aim	Approach
Grey	2014	New Zealand	Cross sectional survey	Public and private health sector organisations in New Zealand	28 stakeholders one-on-one feedback and 100+ meeting attendees	To gain feedback about the interpretation and use of Atlas data for frontline quality improvement.	Reporting of quality measures
Griffiths	2017	UK	Pre- and post-study	Royal College of Pathologists	Training event and grand rounds resulted in 50 checklists completed - no data re attendance at these.	To investigate the feasibility of developing key performance indicators to measure adherence to a specified process of histopathological surgical dissection.	Local QI feedback
Ip	2014	US	Pre- and post-study	183 practices within an integrated health system	2240 adult lower back pain patients between 2007-2010	To examine the impact of a multi-faceted clinical decision support intervention on MRI use in patients with lower back pain.	Health information technology - Decision support
Kelly	2016	Australia	Pre- and post-study	48 hospitals	149,888 patients undergoing percutaneous coronary intervention 2002–2004	To demonstrate that meaningful interpretation from funnel plots can be derived from a New York dataset.	Local QI feedback
Lee	2016	US	Pre- and post-study	Community and specialist inflammatory bowel disease clinics in one health service	50 electronic medical charts of 6 gastroenterology fellows	To incorporate an in-service educational session on IBD health maintenance to increase trainees' knowledge and awareness.	Process variation

Author	Year	Country	Study design	Setting	Sample	Aim	Approach
McFadyen	2015	Canada	Pre- and post-study	One provincial health region	56 clinicians - general surgeons, surgical oncologists, urologists and pathologists.	To provide clinicians with an individualised feedback report to improve quality.	Local QI feedback
Miller	2011	US	Pre and post-study	Three urology practices.	858 urology presentations	To improve patterns of care for radiological staging of newly diagnosed prostate cancer.	Local QI feedback
Min	2017	Canada	Pre and post-study	One major acute care centre.	43 emergency physicians	To determine whether point-of-care clinical decision support can effectively reduce inappropriate medical imaging of patients who present to the emergency room with low back pain.	Health information technology - Decision support
Nguyen	2007	US	Pre- and post-study	44 facilities in the Northwest Renal Network	4 workshops attended by - 36 nephrologists, 16 VA surgeons and 1 radiologist; 35 physicians responded to the follow-up survey	To use educational interventions to promote arteriovenous fistula creation.	Process variation
Nordstrom	2016	US	Pre- and post-study	Cohorts of physician practices across Vermont	28 physician practices in 4 cohorts	To examine physician engagement and change in buprenorphine practice	Process variation

Author	Year	Country	Study design	Setting	Sample	Aim	Approach
Rubin	2015	UK	Mixed methods	8179 primary care practices	92 interviewees - GP, GP cancer leads, public health staff and cancer network staff.	To explore whether quality improvement activities were associated with a change in referral practice.	Local QI feedback
Smith	2013	Australia	Pre- and post-study	Cardiac surgical unit at one hospital	5265 consecutive cardiac procedures 2003–2012	To explore the application of graphical statistical process techniques to inform routine cardiac surgical mortality and morbidity review processes.	Local QI feedback
Stafford	2003	US	Pre- and post-study	117 primary care providers associated with one hospital	105,682 patients and 511328 patient visits	To evaluate the impact of a feedback intervention on reducing rate and variation of ECG orders.	Local QI feedback
Tavender	2015	UK	Conceptual	One emergency department	N/A	To develop a targeted theory-based intervention that improves the management of mild traumatic brain injury.	Local QI feedback
Tomson	2013	The Netherlands	Narrative review	N/A	N/A	To describe quality improvement techniques that maintain clinical quality.	Local QI feedback

N/A: not applicable; QI: Quality improvement; CQI: Continuous quality improvement; IVF: in vitro fertilisation; IBD: inflammatory bowel disease.

Appendix 6: PRISMA diagrams

Figure 1a. PRISMA diagram: Question 1

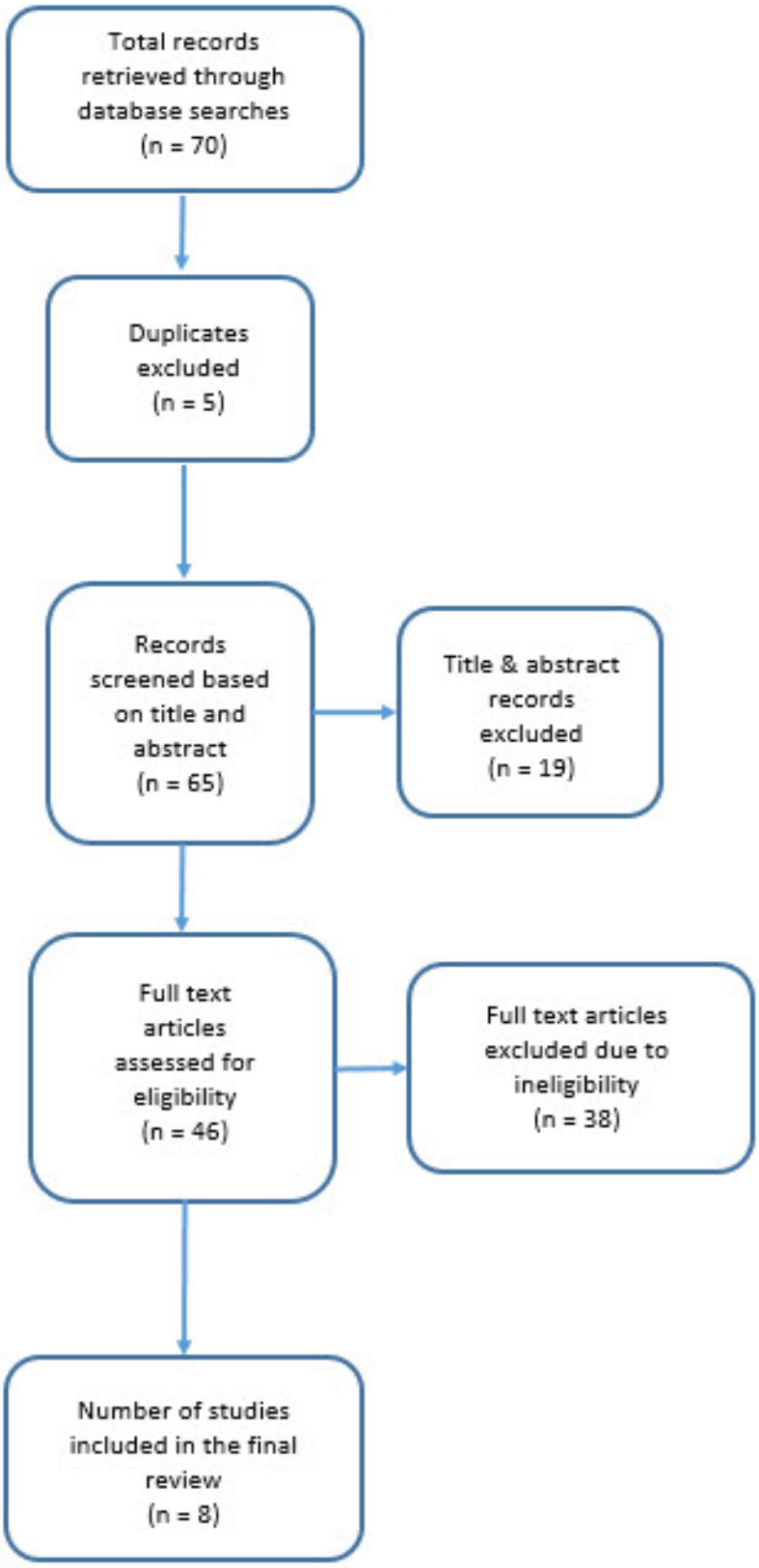
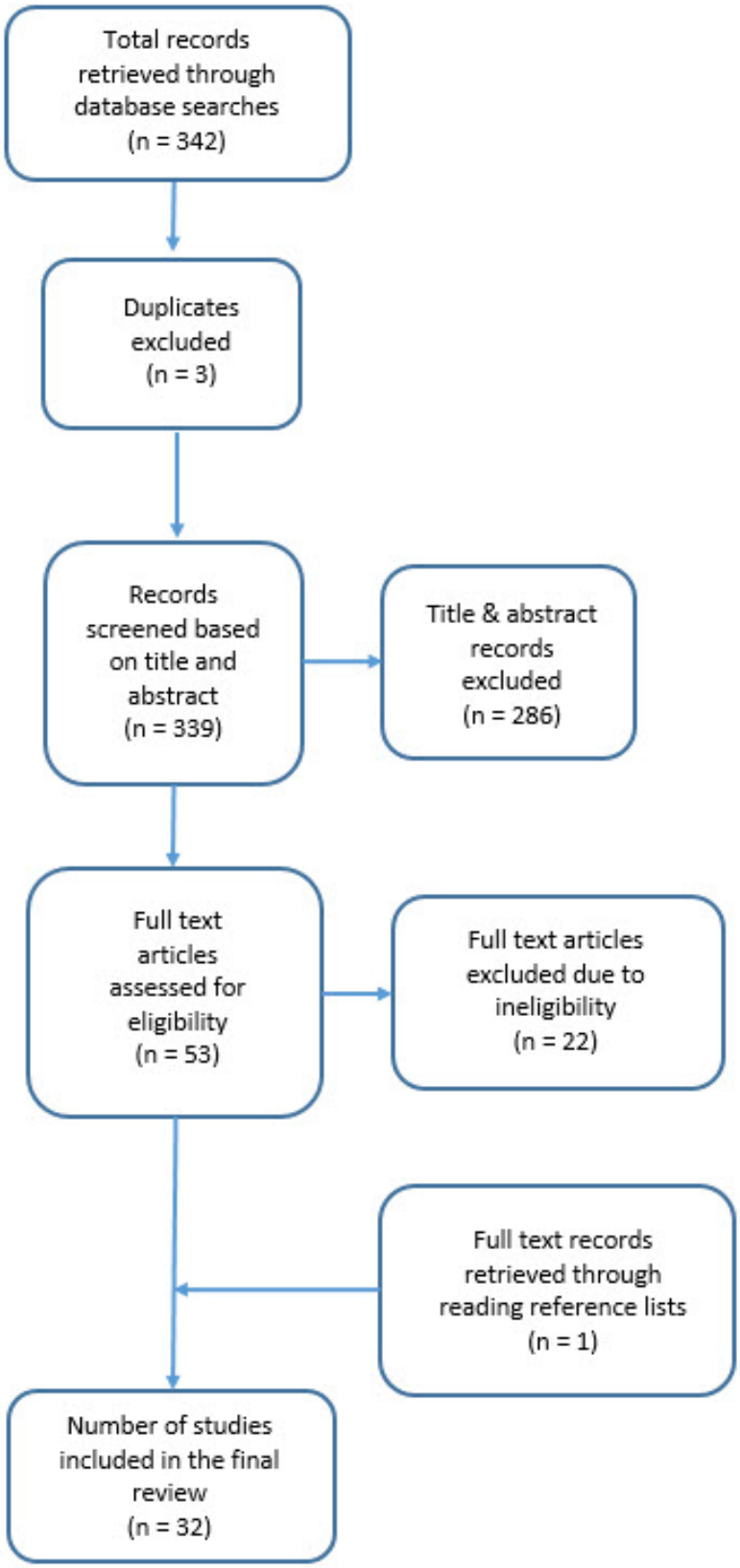


Figure 1b. PRISMA diagram: Question 2



Appendix 7: Data appraisal items

Figure 2: Data appraisal items included in Quality Assessment Tool of Studies of Diverse Design (QATSDD)

Quality Criteria

- Explicit theoretical framework
- Statement of aims/objective in body of report
- Clear description of research setting
- Evidence of sample size considered in terms of analysis
- Representative sample of reasonable size
- Description of procedure for data collection
- Rationale for choice of data collection tool
- Detailed recruitment data (no. approached, declined etc.)
- Statistical assessment of reliability & validity of measurement tools (quantitative)
- Fit between study objectives & method of data collection
- Fit between study objectives & content of data collection tool
- Fit between study objectives and method of analysis
- Good justification for method of analysis
- Assessment of reliability of analytic process (qualitative)
- Evidence of user involvement in design (e.g. pilot work)
- Strengths & limitations critically discussed

Appendix 8: Assessment of UCV Algorithm

Figure 3: Assessment of unwarranted clinical variation algorithm