

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Patterns of care for people presenting to Australian general practice with musculoskeletal complaints based upon routinely collected data: Protocol for an observational cohort study using the Population Level Analysis and Reporting (POLAR) database
AUTHORS	Haas, Romi; Busija, Ljoudmila; Gorelik, Alexandra; O'Connor, Denise; Pearce, Christopher; Mazza, Danielle; Buchbinder, Rachelle

VERSION 1 – REVIEW

REVIEWER	Smith, Chris University of Leeds Leeds Institute of Health Sciences
REVIEW RETURNED	11-Mar-2021

GENERAL COMMENTS	<p>The manuscript reports the protocol for a study that seeks to use a database containing de-identified electronic health records (POLAR) to investigate patterns of care for people who have presented with musculoskeletal conditions to general practices in Victoria, Australia.</p> <p>The manuscript focuses primarily on issues of coding and variation arising in the use of routinely collected data for research purposes. The authors do not address why it is important to understand patterns of care for musculoskeletal patients. Additionally, the authors do not provide any details regarding what is already known about patterns of care for musculoskeletal patients, how any findings from this study would contribute to the literature relating to patterns of care for musculoskeletal patients, and how any findings might inform future care of musculoskeletal patients.</p> <p>There appear to be significant limitations in the data contained within the POLAR database, such as periods for which data may not be available from a particular practice. Additionally, there does not appear to be any evidence provided as to whether the data contained within the POLAR database is representative of the wider population of Victoria and Australia. High-level total numbers of specific types of records within POLAR, e.g. diagnosis, are reported, from which the authors state that they expect to be able to 'detect variation in patterns of care'. However, no evidence is provided to substantiate this expectation, other than these high-level totals.</p> <p>Further detail is required regarding the rationale for the inclusion/exclusion criteria. It is unclear from the inclusion/exclusion criteria described whether the necessary data will be available to robustly and consistently characterize patients and to construct patterns of care. Many of the decisions regarding</p>
-------------------------	--

	<p>how specific concepts are 'coded' appear to be determined by the 'coding framework' used by Outcome Health. In certain cases, e.g. provider records and diagnoses records, the authors state that 'clinical natural language processing' was used by Outcome Health as part of the coding process. Further detail regarding the provenance of the data, and the specific processes to which the data has been subject prior to receipt by the research team, is required to understand whether concept definitions are sufficiently robust to underpin the study.</p> <p>The manuscript contains no discussion regarding the strengths and the limitations of the specific study design. Discussion focuses on reporting of coding and on the POLAR database in general.</p> <p>A RECORD statement has been completed for the manuscript. However, for a number of STROBE items, it is not clear whether the section named in the 'location in manuscript where items are reported' provides information in the form, and with the detail, required for that item.</p> <p>Dates over which data will be considered are provided (2014-2018), but dates over which the study itself will be undertaken are not included.</p> <p>Further details in relation to each section of the paper are provided below.</p> <ul style="list-style-type: none"> - Abstract -- Introduction --- Highlights the issue of repeatability in research based on routinely collected health data and the inconsistency with which required details have been documented in previous work --- Should "... (POLAR) data space system..." be "... (POLAR) database system"? --- Should "Its focus" be "It's focus..."? --- I wasn't clear on whether the paper was aiming to define an --- approach "used to classify" or to critique an existing approach --- It would be helpful if the authors clarified exactly what was meant by "framework" - is it a framework by which researchers can 'classify and identify eligible records from the POLAR database'? Is the 'approach' the same as the 'framework'? It would assist the reader if further clarity was provided in this respect. -- Methods and analysis --- Data will be obtained from general practices across three primary health networks in South Eastern Victoria – is this a subset of the data in POLAR? If so, on what basis is the subset defined? Does each primary health network contain all general practices or a subset of general practices? If it is a subset of general practices, on what basis are general practices included? Further clarity in this regard would assist the reader in determining the patient population whose records might be considered within the study, and whether, for instance, the population is representative of the wider population in Victoria and Australia.
--	--

	<p>--- Inclusion criteria for the patient population could be more clearly presented. I found the use and positioning of the bracketed text, e.g. '(aged 18 years)', to be unclear. Additionally, I found the phrase 'who received care' to be unclear. It appears that patients are included if they have a relevant diagnosis from a GP (and meet the other inclusion/exclusion criteria). Does the 'receipt of care' represent additional criteria, or does it refer to the fact that the diagnosis has been received as a result of an interaction with a GP?</p> <p>--- Further clarity is required on what it means for records to be 'identified deductively' and for the research team to code data 'inductively'. What is being coded? Is the 'low back', 'shoulder, knee of neck conditions' etc referenced earlier in the section?</p> <p>--- Quality checks are mentioned but the basis of these checks is unclear.</p> <p>--- Use of the term 'data files' is not clear. Are there multiple 'data files' in POLAR, or produced as part of the process of identifying the study population? If so, why? For example, do the separate data files contain data from different general practices?</p> <p>--- What is the relationship between the eligible records and study cohort? It appears from the description that the study cohort will be a subset of the eligible records, but it was not clear to me how this subset will be selected.</p> <p>--- Why will multiple 'customized databases' be created for analysis? Why are different databases needed? In what way will they be customized? Can the planned analysis not be undertaken through, for example, execution of queries or definition of views over a single study database?</p> <p>--- Study will seek to characterize patients, GPs, and the care received by patients (imaging requests, prescriptions for pain relief and referrals).</p> <p>-- Ethics and dissemination</p> <p>--- The study has ethical approval from the Cabrini and Monash University Human Research Ethics Committees, and there are plans to dissemination to the controller of the POLAR database (Outcome Health), and to academic audiences via journals and conferences.</p> <p>- Article summary</p> <p>-- It would be helpful for the authors to be clearer on what is means to 'classify, identify and merge' eligible records. For example, I would expect that you would need to identify records before classifying them. Also, why is it necessary to merge records? Do the records to be merged relate to the same patient, or is this referring to the merging of the records for all patients into a single 'database'?</p> <p>-- Is the 'systematic approach' referenced the same as the 'framework' referenced in the abstract? Does the use the 'framework' facilitate a 'systematic approach'?</p>
--	---

It is stated that 'the study sample is expected to be large and representative of GP patterns of care'. This appears to be conflating separate aspects: i) size and representativeness of the patient population – i.e., is the patient population representative of the wider population in South Eastern Victoria, Victoria and Australia?, and ii) representativeness of the patterns of care – i.e., are the patterns of care that are determined by the study representative of actual patterns of care? Given the statement regarding the sample, and in relation to a previous point regarding the coverage of the POLAR database, I would infer that not all general practices in South Eastern Victoria are included in POLAR. It is important for authors to clarify the inclusion/exclusion criteria for general practices on POLAR.

-- The point relating to inaccurate and incomplete information is not limited to this study and is a consideration for any study that seeks to utilize routinely collected data.

-- The authors highlight that 'GP consultations cannot be linked to a specific health condition because the reason for the clinical encounter is not available in the POLAR database', but it is unclear if or how this limits the study. I would expect that many patients discuss multiple conditions within a single GP consultation, which will be reflected in the data that is routinely collected in relation to that consultation. If the study is limited by the inability to attribute a single health condition to a GP consultation in an unambiguous manner, further details are required regarding these limitations.

- Introduction

-- Authors provide a brief background of primary care in Australia and previous initiatives to capture data relating to activities in general practice.

-- Advantages of longitudinal data contained in general practice clinical information systems to examine changes over time are highlighted

-- Authors highlight the inclusion of referrals by GPs as a difference between the Medicine Insight and POLAR databases in Australia.

-- Challenges of using routinely collected data are discussed, namely the transformation of text values to codes as source of variation between studies. Further clarity would be welcomed the specific variation(s) that are being referenced. Is it variation in the text values used by clinicians? Is it variation in the way these text values are transformed to codes by the different systems? Is it variation in the codes used to reference particular 'concepts' (e.g. back pain) in different studies? It appears that it is the later. If so, then the authors need to acknowledge the other sources of variation. For example, two studies may use exactly the same codes to define a given concept, but if, for instance, the two studies use data from different clinical information systems, which transform text values to codes in different ways, then this would need to also be considered when comparing the studies.

-- The authors highlight the findings of previous studies, which have found 'definitions of codes or classification algorithms' were inadequately reported, and highlighted the development of the RECORD guidelines to assist in addressing this issue.

-- In relation to an earlier comment, the authors provide further clarity in this section regarding the inclusion/exclusion criteria for general practices in the POLAR database. However, it remains unclear whether all general practices in the POLAR database use the same clinical information system. It would also be helpful to understand how representative the 30% of general practices included in the POLAR database are of the wider population of general practices in South Eastern Victoria, Victoria and Australia. The authors also provide further clarity regarding the 'data files' that comprise the POLAR database. These 'data files' appear to represent the data from the different 'tables' in a relational database. I was not clear on why merging of records from these 'data files' is a complex task once relevant records have been identified/selected. Do the primary and foreign keys associated with records in a relational database not enable this?

-- The authors highlight previous studies that have used the POLAR database. They highlight that 'these studies have not reported the coding approach used to classify data into meaningful categories or the processes used to merge data files into a customized database for analysis'. It would be very helpful for the authors to clarify what is meant by a 'coding approach'. For example, is this the set of codes used to represent a particular concept (e.g. back pain)? Additionally, it would be helpful for the authors to clarify what is meant by 'meaningful categories'. What constitutes 'meaningful'? Does this mean 'concepts' that are to be considered by the study? What is a category? Does this mean a particular 'class' of data, such as prescriptions, referrals etc?

-- Should 'This manuscript presents a study protocol investigating...' be 'This manuscript presents the protocol for a study investigating...'?

-- The 'framework' appears to comprise of the 'methods used to classify and identify eligible records from the POLAR database...'. This does not appear to address the issue highlighted in the previous paragraph regarding the 'processes used to merge data files'. It would be helpful if the authors were clear on what constitutes the framework, and the specific issue(s) that the framework aims to address.

-- The introduction provides no specific information regarding musculoskeletal conditions and why it is important to understand patterns of care for musculoskeletal conditions. It appears to be written from a very general perspective, i.e. general practice as a whole, focusing on issues of coding, variation and reporting in the POLAR database. I would expect to see further context in the introduction regarding musculoskeletal conditions to provide the rationale for the study.

-- The objectives subsection of the introduction provides objectives which relate specifically to patterns of care for people with musculoskeletal conditions. The objectives seem generally sensible. However, it is unclear how fulfilment of the objectives will contribute to existing knowledge and inform future care or

	<p>research. Further clarity would be welcomed regarding what constitutes a 'management type'. Additionally, with respect to the objective to 'Examine longitudinal changes in management...', does this refer to changes in management within or between patients, or both? Finally, should 'during 2014...' be 'between 2014...', or does the study plan to look at changes within the two stated years (2014 and 2018)? If it is the two years only, then why have these two years been chosen?</p> <p>- Methods</p> <p>-- Study design</p> <p>--- Description is clear.</p> <p>-- Data source</p> <p>--- See previous point regarding whether 'data space' should read 'database'.</p> <p>--- The distinction between 'practice' and 'provider' is unclear and would benefit from clarification.</p> <p>--- I would argue that the description of the POLAR database can be simplified. A description of the internal structure of the database, e.g. '...based on eight relational files...', seems to be low-level for this section. The referenced table is helpful in providing a simple textual summary of the database. Whilst Figure 1 is somewhat helpful in showing the relationships between the different tables in the schema, it may be clearer to readers if a standard approach, such as UML, was used to represent the entities and relationships that comprise the schema.</p> <p>-- Setting</p> <p>--- The study will include data 'relating to all patients diagnosed with an eligible musculoskeletal condition and a face-to-face GP consultation with a GP'. It is unclear whether the diagnosis occurs within the GP consultation. From the phrasing, it appears that a patient must have 'an eligible musculoskeletal condition' and a 'face-to-face GP consultation', but the relationship between the two is not clear. For example, could the diagnosis occur outside of a GP consultation? Would patients for whom this is the case be included? Also, the inclusion of 'with a GP' does not appear to be necessary.</p> <p>--- There are no details provided regarding the clinical information systems used by the practices in the POLAR database. This has been highlighted in an earlier comment. It is unclear whether the general practices in POLAR use the same clinical information system. This is important to clarify as the clinical information system used by the general practice is a potential source of variation that must be considered.</p> <p>--- It is unclear what is meant by 'patient-level follow-up data' and how this is different to the other data that will be included. Clarification on this would be helpful.</p> <p>-- Participants</p>
--	---

	<p>--- See earlier comment relating to abstract about the use of bracketing.</p> <p>--- Brief justification of the different age criteria is provided.</p> <p>--- Table 2 could be clearer in communicating the eligibility criteria. It may help to separate the inclusion/exclusion criteria for the patient population, and the inclusion/exclusion criteria for data relating to this patient population.</p> <p>--- As mentioned in a previous comment, further clarity is required regarding 'deductive' and 'inductive' approaches.</p> <p>--- Further information is required regarding the approach or 'framework' used by Outcome Health to code the data. Do Outcome Health manually go through text values entered by clinicians and associate codes with these values? If so, how is this undertaken? What are the components of their framework? How do they ensure consistency within and between records etc...? Validity of the study is very much contingent on these details.</p> <p>-- Variables</p> <p>--- Please note, I am not a clinician. Therefore, I cannot review the appropriateness of the inclusion/exclusion criteria for variables from a clinical perspective.</p> <p>--- Provider records appear to contain a 'provider type' field which can be used to determine whether the provider is a GP. Further details would be helpful as to the nature of this field, why 'clinical natural language processing' is required, and the specific form of this processing.</p> <p>--- Diagnoses are coded using SNOMED-CT. It appears that 'clinical natural language processing' defined by Outcome Health was used to obtain SNOMED-CT codes from text values. Evidence is referenced regarding the accuracy of this coding process. Again, further details would be helpful regarding the specific form of the 'clinical natural language processing'.</p> <p>--- Patients with an 'eligible musculoskeletal diagnosis during 2014-2018' were included. This appears to be repeating the inclusion/exclusion criteria for the study population. I would expect this section to provide details regarding the data items included for the study population.</p> <p>--- Details are provided of specific SNOMED terms that were excluded. Some of the rationale for inclusion/exclusion was unclear to me, for instance, the exclusion of general musculoskeletal terms that 'could not be attributed to a specific body region'. It would be helpful if the authors could provide further clarity regarding the specific set of musculoskeletal conditions for which they are interested in determining patterns of care.</p> <p>--- Diagnoses of co-morbidities were included, which again appear based on Outcome Health's coding system. Further detail regarding this coding process would again be helpful. For example, how are 'chronic musculoskeletal conditions' identified?</p>
--	--

	<p>--- As a more general point, the patients included in the study population will be at different stages in the course of their disease in, for example, 2014. Will data relating to musculoskeletal conditions be included for a period prior to 2014? If not, how will patients at different stages in the course of their disease be distinguished. Additionally, patients may be at different stages in the course of other co-morbidities in, for example, 2014. Will data relating to these comorbidities be included for a period prior to 2014? How will the presence/absence of co-morbidities be determined? Do diagnoses need to have occurred within a particular time period (e.g. 5 years) prior to the musculoskeletal condition diagnosis that determines inclusion in the population? It was not clear whether the data would be available to robustly determine the presence/absence of comorbidities for all patients in the population.</p> <p>--- Coding provided within the POLAR database is used to identify for specific types of activities. It would be helpful to know the origin of values of the 'activity type' field. For instance, are they associated with the records automatically by the clinical information system, are they derived from values input by clinicians through some post-processing carried out prior to, or following, incorporation into the POLAR database?</p> <p>--- The comment above is also relevant to the referral records.</p> <p>--- Further detail is required regarding the inclusion/exclusion criteria for medications. It was unclear why relevant medications could not be determined a priori. Why was it necessary for two study authors to determine these from the prescription file?</p> <p>--- It was unclear what was meant by the 'test data file' in the description of the imaging records. It appears that Outcome Health do not, or have not yet, coded imaging tests, and that it was necessary for the authors to define a process by which imaging records could be coded in an automated manner due to the number of records. Records were classified by anatomical region and then sub-classified according to the type of test. The authors provide details of challenges in the coding process, such as test names that did not reference a test, and provided details of how these were resolved. Authors excluded test names that they determined not to 'indicate an imaging test or procedure'. However, might there be cases where such entries provide the only reference to a test?</p> <p>-- Data access and cleaning</p> <p>--- It was unclear of the relevance of the 'database records from 1997' that are referenced in this section. What is the relevance of 1997 to the study?</p> <p>--- Details are provided regarding the exclusion of records based on missing values and values that were deemed erroneous.</p> <p>--- Authors state that 'Data from practices proceeding a gap in reporting of one year or more during the study period will also be deemed ineligible'. Why would there be a gap in the data for a practice? Absence of data for one or more years within the period from 2014-2018 will significantly impact any ability to study longitudinal trends of more than one year. Additionally, it will</p>
--	--

	<p>impact the ability to robustly characterize patients with respect to, for instance, co-morbidities. For example, what is a patient was diagnosed with a particular co-morbidity in year for which data is missing. Additional complexity is added by the fact that the years for which data may be missing appears to vary between practices. Therefore, robust and consistent characterization of, for instance, co-morbidities for patients across all practices will be a very significant challenge.</p> <p>-- Approach to dataset creation</p> <p>--- This section seems to replicate details previously provided. As outlined in a previous comment, it is unclear why four different databases are required, as opposed to the creation of a single database over which specific queries and views can be defined to support different aspects to the analysis.</p> <p>-- Sample size consideration</p> <p>--- The authors report the number of records in the POLAR database at the time of extract. However, these numbers relate to all records, rather than the records that match the inclusion/exclusion criteria that they have previously defined. The expectation of a 'sufficiently large sample' is not substantiated. Additionally, it is unclear what the authors mean by 'sufficiently large'. What sample size would be 'sufficiently large' to 'detect variation in patterns of care'?</p> <p>-- Analyses</p> <p>--- Authors report that the lists of codes used to 'define eligible variables' have been made available on clinicalcodes.org.</p> <p>--- It is stated the 'Patterns and timing of management...will be examined and compared between 2014 and 2018'. As mentioned in a previous comment, is this comparison between two years, i.e. 2014 and 2018, or between all years in the period from 2014 to 2018.</p> <p>--- The authors plan to use 'group-based trajectory modelling to identify groups of individuals with distinct patterns of prescription opioid use over time'. However, given the limitations of the data highlighted in previous comments, e.g. missing years for practices, it is unclear how this can undertaken robustly.</p> <p>--- Analysis will be undertaken relative to the 'first (index) eligible musculoskeletal diagnosis'. However, my understanding from the inclusion/exclusion criteria is that this 'first (index)...diagnosis' is the first occurrence within the period 2014-2018 and the patient may have had relevant diagnoses prior to 2014. Therefore, patients may be at different stages in the course of their disease at the 'first (index)...diagnosis' and may receive different management as a consequence.</p> <p>--- The authors plan a large number of different analyses. The scope of these analyses appears to be very wide. The analyses may benefit from a significant degree of simplification, both in their performance and in their description within this manuscript.</p> <p>-- Patient and public involvement</p>
--	--

	<p>--- No patient and public involvement will be undertaken in the study.</p> <p>- Discussion</p> <p>-- Whilst the authors state that the protocol 'conforms to the RECORD guidelines by describing the coding process', much of the coding appears to be undertaken by Outcome Health using their 'coding framework'. Therefore, any interpretation of the concepts used within the study needs to consider both this coding framework and the inclusion/exclusion criteria detailed in this manuscript.</p> <p>-- The authors have made some of the code lists in available in the ClinicalCodes online repository, which is beneficial for future studies.</p> <p>-- It is stated that the POLAR database 'is expected to contain a large and representative study population'. Expectation of a representative study population is not the same as providing evidence that the study population is indeed representative. If there is no existing evidence regarding whether the population within the POLAR database is representative of the wider population in Victoria and Australia, such evidence needs to be obtained prior to any studies being undertaken using the data from POLAR.</p> <p>-- The authors highlight a strength of the POLAR database as it's 'comprehensive and longitudinal nature' and to 'enable temporal sequences to be examined over time'. However, from the description provided in the 'Data access and cleaning' section, it would appear that there can be significant gaps, e.g. a year, in the data available for practices. This places a significant limitation of the ability to construct robust and consistent temporal patterns for patients from different practices over multiple years.</p> <p>-- Improvements discussed appear to relate more to the clinical information systems used by practices that contribute to POLAR rather than the POLAR database.</p> <p>-- There is no discussion regarding the strengths and the limitations of the specific study design. Discussion focuses on reporting of coding and on the POLAR database in general.</p> <p>- Ethics and Dissemination</p> <p>-- The study has ethical approval from the Cabrini Human REC and Monash University REC. Dissemination to Outcome Health was mentioned in the description in the abstract, but not in this section. 'Participating PHNs' are referenced in this section but not in the description in the abstract. It would improve clarity if the description was consistent between the abstract and this section.</p>
--	---

REVIEWER	Jordan, Kelvin Keele University, School of Primary, Community and Social Care
REVIEW RETURNED	28-Mar-2021

GENERAL COMMENTS	This protocol reports the methods of a thorough study to utilise routine electronic health records (EHR) to describe patterns of
-------------------------	--

	<p>care for musculoskeletal pain in Australian general practice. I just have a few comments.</p> <p>1) The authors rightly say transparency in codes and algorithms in studies using EHR is important. They suggest their protocol will allow replication by other researchers. Whilst I fully agree on the importance of this, it would be useful to consider the scenarios where replication and use of these codes and algorithms may be most feasible. If I understand the protocol, the “raw” information is not coded at the time of entry, but is coded later through natural language processing. This is different to many other nations’ general practice records where symptoms, illnesses and processes of care are coded as the health care professional enters information (SNOMED or ICD, for example). It would be useful to set the context in the background as to how Australian general practice recording and health care system differs to other countries; and how much what is reported here is mainly relevant to researchers using POLAR and how much is more widely generalisable.</p> <p>2) I wondered why osteoarthritis without region specified was excluded? This could exclude a large group of people with musculoskeletal pain, particularly those with generalised OA who may not have a region recorded, given how common OA is in older adults. It seems a shame not to include this.</p> <p>3) How will the researchers know that a prescription is for a relevant musculoskeletal problem? Particularly if someone has several musculoskeletal conditions recorded. Will there be a reason for prescription recorded?</p> <p>4) Will the researchers be able to separate prevalent from incident (or perhaps new episode) consultations? This seems important when characterising management.</p> <p>4) Will the latent class analysis categorising management types take into account order of management approaches and time since consultation?</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Author Response to Reviewer Comments

Thank you to the reviewers for their valuable feedback and suggestions. All comments have been reviewed and our responses to each individual comment is provided below, including the accompanying amendments to the manuscript. Many of Reviewer 1’s comments relate to details of the data processes of Outcome Health. While we have made some changes as outlined below, we believe that Reviewer 1 is requesting a level of detail that is beyond the scope of this paper. Comprehensive descriptions of the methods have been published elsewhere and we refer the reviewer to references 9 (overall nature of POLAR) and 20 (details of the extraction and coding process).

Reviewer 1, Comment 1:

The manuscript reports the protocol for a study that seeks to use a database containing de-identified electronic health records (POLAR) to investigate patterns of care for people who have

presented with musculoskeletal conditions to general practices in Victoria, Australia. The manuscript focuses primarily on issues of coding and variation arising in the use of routinely collected data for research purposes. The authors do not address why it is important to understand patterns of care for musculoskeletal patients. Additionally, the authors do not provide any details regarding what is already known about patterns of care for musculoskeletal patients, how any findings from this study would contribute to the literature relating to patterns of care for musculoskeletal patients, and how any findings might inform future care of musculoskeletal patients.

Response:

We have now added the information below to address these comments.

Amendment:

1st paragraph of the introduction (pages 1-2):

“General practice plays an essential role in providing primary health care to the population. In Australia 86% of the population visits a general practitioner (GP) multiple times a year¹, and nearly 20% of these consultations are for a musculoskeletal condition². These conditions account for 23% of the years lived with disability in Australia³ and are also a major cause of disability worldwide⁴. Until 2016, the BEACH (Bettering the Evaluation and Care of Health) program provided the most comprehensive data on clinical activities of Australian general practice⁵. The program identified a number of activities that represent low-value care for people with musculoskeletal conditions including an over-reliance on imaging, prescription of opioids, and unnecessary referrals to specialist care^{6 7}. However, in-depth exploration of these activities within the BEACH program is limited by its cross-sectional design, and these data are no longer being collected.”

Final paragraph of the introduction (page 4):

“Study findings will advance existing knowledge about GP care for people with these musculoskeletal conditions and whether it conforms to best evidence-based practice. Differences in care across different musculoskeletal complaints may also inform tailored interventions to improve care and ultimately reduce the burden of disease associated with these musculoskeletal complaints.”

Reviewer 1, Comment 2:

There appear to be significant limitations in the data contained within the POLAR database, such as periods for which data may not be available from a particular practice.

Response:

There is no reason to suggest the POLAR database contains periods for which data may not be available from a particular practice. However, we agree that if this were the case, it could be a significant limitation. This is why we have included a process at the data cleaning stage to assess for this and manage this situation only if it eventuates. The process we have stipulated is similar to that

recommended and used by the Clinical Practice Research Datalink (CPRD) database in the United Kingdom (see reference 27 within manuscript).

Amendment:

We have clarified this in the Data access and cleaning section (page 8).

“We will exclude practices without any activity data during 2014-2018. We will also examine the consistency of activity, test, prescription, and referral data for each practice in each eligible calendar year. If a gap in reporting from any practice is identified for one year or more, only data from the earliest date after which there was no gap will be included.”

Reviewer 1, Comment 3:

Additionally, there does not appear to be any evidence provided as to whether the data contained within the POLAR database is representative of the wider population of Victoria and Australia.

Response:

The data from each practice within the POLAR database will be representative of each practice because it contains all the data from each practice throughout 2014-2018. We do not yet know if the general practices within the POLAR database or the people diagnosed with musculoskeletal conditions within these practices are representative of the wider population of Victoria and Australia. Although approximately 30% of general practices across South-Eastern Victoria are included within the POLAR database, non-accredited, corporate owned practices, and those not using electronic medical records are likely to be under-represented. Previous research has demonstrated comparable prevalence and age-gender distribution of people diagnosed with type 2 diabetes in the POLAR database to those within Australia. This study will add to these findings by comparing the demographic characteristics of people diagnosed with musculoskeletal conditions within the POLAR database to those within the wider populations of Victoria and Australia using data from the Australian national health survey (. Further information about the representativeness of the POLAR database will therefore be provided in the subsequent results manuscript pertaining to the study cohort.

Amendment:

Analyses (page 9):

“Descriptive statistics will be used to summarise the study cohort including the number and type of eligible musculoskeletal conditions; patient demographics; and comorbidities. These will be compared to data from the Australian national health survey to assess the representativeness of the POLAR database to the wider Australian population.”

Discussion (page 10):

“The main strength of this study is that it will facilitate an overview of the care provided by GPs to the same patient(s) over time and thereby enable temporal sequences to be examined. The POLAR database contains all patient-related activity within each practice making it representative

of the included practices. Previous research has demonstrated comparable prevalence and age-gender distribution of people diagnosed with type 2 diabetes within the POLAR database to those within Australia²⁷. This study will add to these findings by assessing the representativeness of people with musculoskeletal conditions within the POLAR database to the wider populations of Victoria and Australia.”

Reviewer 1, Comment 4:

High-level total numbers of specific types of records within POLAR, e.g. diagnosis, are reported, from which the authors state that they expect to be able to ‘detect variation in patterns of care’. However, no evidence is provided to substantiate this expectation, other than these high-level totals.

Response:

Since this manuscript is a protocol, we do not yet know the number of records that match the inclusion/exclusion criteria. We have therefore estimated an expected sample size based on the size of a previous cohort with type 2 diabetes from the POLAR database and compared this with the sample size required for our planned sequence analyses that will be used to examine overall patterns of care. We have also moved the sample size consideration section to below the analyses because this is now based on the planned analyses.

Amendment:

Sample size consideration (page 10):

“Sequence analysis will require the largest sample size of our planned analyses and will therefore form the basis of our sample size consideration. We plan to examine the following six management types: non-surgical referrals, surgical referrals, allied health referrals, opioid prescription, X-ray and/or ultrasound requests, and MRI and/or CT scan requests. This provides a total of 720 potential sequence combinations. Based on a recommended 20 to 30 subjects per subgroup³¹, we estimate a sample size of between 14,400 and 21,600 will be required to differentiate between each sequence combination or pattern of care. Recent use of the POLAR database using data from approximately 200 general practices identified 20,514 active adult patients with type 2 diabetes before July 2016³². Our extract is based on 301 general practices from 2014 to 2018 and since the prevalence of diabetes is less than that of musculoskeletal conditions³³, we expect a sample size of more than 20,000.”

Reviewer 1, Comment 5:

Further detail is required regarding the rationale for the inclusion/exclusion criteria. It is unclear from the inclusion/exclusion criteria described whether the necessary data will be available to robustly and consistently characterize patients and to construct patterns of care. Many of the decisions regarding how specific concepts are ‘coded’ appear to be determined by the ‘coding framework’ used by Outcome Health. In certain cases, e.g. provider records and diagnoses records, the authors state that ‘clinical natural language processing’ was used by Outcome Health as part of the coding process. Further detail regarding the provenance of the data, and the specific processes to which the data has been subject prior to receipt by the research team, is required to understand whether concept definitions are sufficiently robust to underpin the study.

Response:

Further detail regarding the rationale for the inclusion/exclusion criteria has been added to the manuscript. We have described the methods used to classify and select eligible records from the POLAR database and how relational data files will be merged into a patient-centred database to ensure the patients will be robustly and consistently characterized and to construct patterns of care in a systematic manner. Prior research (see references 9 and 20) and extensive discussions with Outcome Health prior to receiving the data extract revealed the necessary data files were available for this analysis. We have described Outcome Health's approach to coding at the beginning of the variables section and within each of the specific variables as required. For further detail we refer to reference 20.

Amendment:

Variables (page 5):

"Preparatory work to classify and select eligible records has been completed as part of the protocol process. In circumstances where Outcome Health has previously coded data (e.g., diagnosis records), we used this coding to select eligible records that fitted our inclusion criteria. In circumstances where there was no coding (e.g., imaging tests), we coded the data into categories and then selected eligible records. Outcome Health's approach to coding used clinical natural language processing to automatically code structured narrative text within the electronic medical record following by a manual process for quality checking and correction²⁰. For example, this allowed the free text items 'back pain', 'low back pain', and 'lumbar pain' to all sit under the same diagnostic code. Where possible, coding was conducted using a standardised classification system. For example, diagnoses are coded using SNOMED CT-AU terminology²¹ and prescriptions are coded according to the Anatomical Therapeutic Chemical (ATC) classification system²². In cases where there is no standardised classification system available (e.g., providers and referrals), Outcome Health used a similar process to code these variables into relevant categories (e.g., type of health care provider). Clinical natural language processing conducted by Outcome Health has previously demonstrated accurate coding of over 95% of the narrative text to SNOMED CT-AU terms in a sample of approximately 57,000 diagnosis records²⁰. Our approaches to coding and/or selecting eligible records for each variable are described in detail below."

Reviewer 1, Comment 6:

The manuscript contains no discussion regarding the strengths and the limitations of the specific study design. Discussion focuses on reporting of coding and on the POLAR database in general.

Response:

The discussion has been amended to focus on the strengths and the limitations of the study design rather than the POLAR database in general.

Amendment:

See amended discussion (pages 10-11).

Reviewer 1, Comment 7:

A RECORD statement has been completed for the manuscript. However, for a number of STROBE items, it is not clear whether the section named in the 'location in manuscript where items are reported' provides information in the form, and with the detail, required for that item.

Response:

We have reviewed each item within the RECORD statement to check the information contained within the manuscript is provided with sufficient detail required for each item. For clarity, we have completed the 'location in manuscript where items are reported' in the corresponding STROBE and RECORD items. We have also added a brief description of the information that is referred to within the manuscript.

Amendment:

See revised RECORD statement.

Reviewer 1, Comment 8:

Dates over which data will be considered are provided (2014-2018), but dates over which the study itself will be undertaken are not included.

Response:

Data analysis will be undertaken throughout 2021.

Amendment:

Setting (page 5):

"Data analyses will be completed by the end of 2021."

Abstract

Reviewer 1, Comment 9:

Should "... (POLAR) data space system..." be "... (POLAR) database system"?

Response:

Thank-you for this suggestion.

Amendment:

Throughout the manuscript, reference to the POLAR data space system has been changed to POLAR database.

Reviewer 1, Comment 10:

Should “Its focus” be “It’s focus...”?

Response:

We have used the possessive form of ‘its’ to indicate ownership rather than a contraction for it is. We are happy for the editor to change this if necessary.

Reviewer 1, Comment 11:

I wasn’t clear on whether the paper was aiming to define an --- approach “used to classify” or to critique an existing approach.

Response:

This study has described the systematic approach used to classify and select eligible records.....rather than to critique an existing approach. We refer to the introduction (page 4):

“This manuscript presents a protocol for a study investigating patterns of GP care for people with a low back, neck, shoulder and/or knee condition in Victoria, Australia. It describes the methods used to classify and select eligible records from the POLAR database and how relational data files will be merged into a patient-centred database. This systematic approach will guide future research by enabling researchers interested in using routinely collected health data, and the POLAR database in particular, to answer other clinically relevant questions about general practice care.”

Reviewer 1, Comment 12:

It would be helpful if the authors clarified exactly what was meant by “framework” - is it a framework by which researchers can ‘classify and identify eligible records from the POLAR database’? Is the ‘approach’ the same as the ‘framework’? It would assist the reader if further clarity was provided in this respect.

Response:

Thank-you for this suggestion.

Amendment:

For consistency, we have now used the term ‘systematic approach’ instead of ‘framework’ to describe the methods used to classify and identify eligible records and the process used to merge data files into a patient-centred database throughout the manuscript.

Reviewer 1, Comment 13:

Data will be obtained from general practices across three primary health networks in South Eastern Victoria – is this a subset of the data in POLAR? If so, on what basis is the subset defined? Does

each primary health network contain all general practices or a subset of general practices? If it is a subset of general practices, on what basis are general practices included? Further clarity in this regard would assist the reader in determining the patient population whose records might be considered within the study, and whether, for instance, the population is representative of the wider population in Victoria and Australia.

Response:

The POLAR database includes a subset of general practices from each Primary Health Network (PHN) based on consent from these practices. This has been clarified throughout the manuscript and the potential implications on representativeness of the study sample have been added to the discussion.

Amendment:

Abstract (page 2):

“Patient-related data will be obtained through electronic health records from a subset of general practices across three primary health networks (PHN) in South Eastern Victoria.”

Introduction (page 3):

“The POLAR database draws data from every consultation occurring for millions of patients in approximately 30% of general practices across South-Eastern Victoria⁸, an area that comprises more than half of Victoria’s population⁹. Inclusion is based on practice consent so this volume is increasing exponentially as more practices consent to add their data and as more consultations occur over time.”

Analyses (page 9):

“Descriptive statistics will be used to summarise the study cohort including the number and type of eligible musculoskeletal conditions; patient demographics; and comorbidities. These will be compared to national health survey data to assess the representativeness of the POLAR database to the wider population.”

Discussion (page 10):

“The POLAR database contains all patient-related activity within each practice making it representative of the included practices. Previous research has demonstrated comparable prevalence and age-gender distribution of people diagnosed with type 2 diabetes within the POLAR database to those within Australia³². This study will add to these findings by assessing the representativeness of people with musculoskeletal conditions within the POLAR database to the wider Australian population.”

Reviewer 1, Comment 14:

Inclusion criteria for the patient population could be more clearly presented. I found the use and positioning of the bracketed text, e.g. '(aged 18 years)', to be unclear. Additionally, I found the phrase 'who received care' to be unclear. It appears that patients are included if they have a relevant diagnosis from a GP (and meet the other inclusion/exclusion criteria). Does the 'receipt of care' represent additional criteria, or does it refer to the fact that the diagnosis has been received as a result of an interaction with a GP?

Response:

We have removed the bracketed text referring to age restrictions from the abstract in order to keep within the word limit of the abstract. These age restrictions are now clarified within the participants section. Patients are eligible for inclusion if they have a relevant diagnosis from a GP and received at least one face-to-face consultation with a GP during the study period. 'Receipt of care' refers to the fact that a patient had to have consulted a GP in a face-to-face consultation. This has now been clarified.

Amendment:

Abstract (page 2):

"Data for patients with a low back, neck, shoulder and/or knee condition and who received at least one GP face-to-face consultation between 1/01/2014 and 31/12/2018 will be included."

Participants (page 5):

"The study cohort will include people diagnosed during 2014 to 2018 inclusive with a low back, neck, shoulder and/or knee condition, limited to age 45 years and over except for low back which will be limited to age 18 years and over."

Reviewer 1, Comment 15:

Further clarity is required on what it means for records to be 'identified deductively' and for the research team to code data 'inductively'. What is being coded? Is the 'low back', 'shoulder, knee of neck conditions' etc referenced earlier in the section?

Response:

A deductive approach refers to using Outcome Health's coding approach to select eligible records. An inductive approach refers to conducting our own coding in order to select eligible records. We appreciate these terms are unclear so have instead described the process we used to select eligible records. The term 'relational' data files has been added to the abstract and further explanation is provided within the variables section of the methods on page 5.

Amendment:

Sentence deleted from abstract and explained further within the variables section of the methods on page 5.

The term 'relational' data files has now been added in the abstract for clarity.

"Relational data files with eligible and valid records will be merged to select the study cohort....."

Reviewer 1, Comment 16:

Quality checks are mentioned but the basis of these checks is unclear.

Response:

The quality of general practice electronic medical record data is variable because data are entered by GPs during routine consultations rather than for the purposes of research. It is therefore important to conduct comprehensive data quality checks to identify and exclude patients with non-continuous follow-up or patients with poor data recording.

Amendment:

Abstract methods and analysis (page 1):

"Data quality checks will be conducted ~~and to identify and exclude patients with poor data recording and/or non-continuous follow-up.~~ Relational data files with eligible and valid records will be merged to select the study cohort and ~~create customised databases for analyses~~ GP management (consultations, imaging requests, prescriptions and referrals) conducted between diagnosis and 31/12/2018."

Reviewer 1, Comment 17:

Use of the term 'data files' is not clear. Are there multiple 'data files' in POLAR, or produced as part of the process of identifying the study population? If so, why? For example, do the separate data files contain data from different general practices?

Response:

Data files are the relational data files provided by POLAR. e.g. provider, practice, patient, activity, diagnosis, referral, prescription, imaging. This was explained within the introduction as follows: "Data are provided to research users in a relational database that organises data into files that can be merged based on common data fields."

Amendment:

The term 'relational' data files has now been added in the abstract for clarity (page 1).

“Relational data files with eligible and valid records will be merged to select the study cohort and create customised databases for analyses GP management (consultations, imaging requests, prescriptions and referrals) conducted between diagnosis and 31/12/2018.”

Reviewer 1, Comment 18:

What is the relationship between the eligible records and study cohort? It appears from the description that the study cohort will be a subset of the eligible records, but it was not clear to me how this subset will be selected.

Response:

The study cohort will be patients with a low back, neck, shoulder and/or knee condition and who received at least one GP face-to-face consultation between 1/01/2014 and 31/12/2018. To select the study cohort, relational data files with eligible and valid records will be merged.

Amendment:

Abstract (page 1):

“Relational data files with eligible and valid records will be merged to select the study cohort and create customised databases for analyses the care received by GPs (consultations, imaging requests, prescriptions and referrals) between diagnosis and 31/12/2018.”

Reviewer 1, Comment 19:

Why will multiple ‘customized databases’ be created for analysis? Why are different databases needed? In what way will they be customized? Can the planned analysis not be undertaken through, for example, execution of queries or definition of views over a single study database?

Response:

The planned analysis will be undertaken through execution of queries and we therefore do not require different databases for analysis. Reference to ‘customised databases’ has now been replaced with ‘patient-centred database’ throughout the manuscript.

Amendment:

Abstract (page 1):

“Relational data files with eligible and valid records will be merged to select the study cohort and create customised databases for analyses the care received by GPs (consultations, imaging requests, prescriptions and referrals) between diagnosis and 31/12/2018.”

Strengths and limitation of this study (page 2):

"This is the first study to our knowledge to report the codes and algorithms used to classify, ~~identify~~ select and merge eligible records from the POLAR database into a patient-centred ~~customised~~ database to facilitate analysis of general practice patterns of care."

Reviewer 1, Comment 20:

It would be helpful for the authors to be clearer on what it means to 'classify, identify and merge' eligible records. For example, I would expect that you would need to identify records before classifying them. Also, why is it necessary to merge records? Do the records to be merged relate to the same patient, or is this referring to the merging of the records for all patients into a single 'database'?

Response:

It was necessary to classify or code records into categories prior to selecting eligible records for inclusion. This coding or classification was conducted by Outcome Health for provider, activity, diagnosis, referral and prescription records. However, in the case of imaging requests, this was conducted by the research team and is outlined in this protocol. For clarity, we have changed the term 'identify' to 'select' throughout the manuscript when referring to this process of selecting eligible records for inclusion within each relational data file.

Since POLAR data are provided to research users in a relational database, data from eligible files (e.g. patient, practice, provider, diagnosis, consultations, imaging, prescriptions and referrals) need to be merged to create a patient-centred database for analysis. This refers to the merging of records relating to the same patient and has now been clarified throughout the manuscript. Further clarity has also been provided within the strengths and limitations section.

Amendment:

Strengths and limitations of this study (page 2):

"This is the first study to our knowledge to report the codes and algorithms used to classify, ~~identify~~ select and merge eligible records from the POLAR database into a ~~customised~~ patient-centred database to facilitate analysis of general practice patterns of care."

Reviewer 1, Comment 21:

Is the 'systematic approach' referenced the same as the 'framework' referenced in the abstract? Does the use of the 'framework' facilitate a 'systematic approach'?

Response:

The 'systematic approach' is the same as the 'framework' referenced in the abstract.

Amendment:

As per our response to Comment 12 above regarding the abstract, for consistency, we have now used the term 'systematic approach' instead of 'framework' to describe the methods used to classify and identify eligible records throughout the manuscript.

Reviewer 1, Comment 22:

It is stated that 'the study sample is expected to be large and representative of GP patterns of care'. This appears to be conflating separate aspects: i) size and representativeness of the patient population – i.e., is the patient population representative of the wider population in South Eastern Victoria, Victoria and Australia?, and ii) representativeness of the patterns of care – i.e., are the patterns of care that are determined by the study representative of actual patterns of care? Given the statement regarding the sample, and in relation to a previous point regarding the coverage of the POLAR database, I would infer that not all general practices in South Eastern Victoria are included in POLAR. It is important for authors to clarify the inclusion/exclusion criteria for general practices on POLAR.

Response:

As per our response to Comment 13 above, the POLAR database includes a subset of general practices from each PHN based on consent from these practices. This has been clarified throughout the manuscript and the potential implications on representativeness of the study sample have been added to the discussion. The strength of this study relating to representativeness has been reframed to extend upon previous research that has assessed the representativeness of an alternate patient population in the POLAR database (see reference 32 within the manuscript).

Amendment:

Strengths and limitations of this study (page 2):

~~"The study sample is expected to be large and representative of GP patterns of care across South Eastern Victoria provided to patients with musculoskeletal conditions. This study will extend previous research that has assessed the representativeness of POLAR data to GP care across the wider Australian population"~~

Abstract methods and analysis (page 1):

"Patient-related data will be obtained through electronic health records from a subset of general practices across three primary health networks (PHN) in South Eastern Victoria."

Introduction (page 3):

"The POLAR database draws data from every consultation occurring for millions of patients in approximately 30% of general practices across South-Eastern Victoria, an area that comprises more than half of Victoria's population⁹. Inclusion is based on practice consent."

Discussion (page 10):

“The POLAR database contains all patient-related activity within each practice making it representative of the included practices. Previous research has demonstrated comparable prevalence and age-gender distribution of people diagnosed with type 2 diabetes within the POLAR database to those within Australia³². This study will add to these findings by assessing the representativeness of people with musculoskeletal conditions within the POLAR database to the wider Australian population.”

The point relating to inaccurate and incomplete information is not limited to this study and is a consideration for any study that seeks to utilize routinely collected data.

Response:

We agree that this point is a consideration for any study that seeks to utilise routinely collected data. We have therefore re-framed this limitation to relate to incomplete information stemming from medications for pain relief and referrals to allied health providers that may be generated outside general practice.

Amendment:

Strengths and limitations of this study (page 2):

“Inaccurate and incomplete information entered in the electronic health record by GPs in the clinical setting can limit the validity and volume of data analysed. These data are likely to underestimate actual allied health visits as some of these do not require a GP referral in Australia; some prescriptions for pain relief are available without a prescription so these data will also be underestimated.”

Reviewer 1, Comment 23:

The authors highlight that ‘GP consultations cannot be linked to a specific health condition because the reason for the clinical encounter is not available in the POLAR database’, but it is unclear if or how this limits the study. I would expect that many patients discuss multiple conditions within a single GP consultation, which will be reflected in the data that is routinely collected in relation to that consultation. If the study is limited by the inability to attribute a single health condition to a GP consultation in an unambiguous manner, further details are required regarding these limitations.

Response:

Further details regarding these limitations have now been provided.

Amendment:

Strengths and limitations of this study (page 2):

“GP consultations cannot be linked to a specific health condition because the reason for the clinical encounter is not available in the POLAR database. It is possible not all patterns of care provided for the study cohort will be directly related to a musculoskeletal condition as reasons for GP consultations, referrals, and prescriptions are not mandated by the source Electronic Medical Records (EMRs).”

Introduction

Reviewer 1, Comment 24:

Challenges of using routinely collected data are discussed, namely the transformation of text values to codes as source of variation between studies. Further clarity would be welcomed the specific variation(s) that are being referenced. Is it variation in the text values used by clinicians? Is it variation in the way these text values are transformed to codes by the different systems? Is it variation in the codes used to reference particular 'concepts' (e.g. back pain) in different studies? It appears that it is the later. If so, then the authors need to acknowledge the other sources of variation. For example, two studies may use exactly the same codes to define a given concept, but if, for instance, the two studies use data from different clinical information systems, which transform text values to codes in different ways, then this would need to also be considered when comparing the studies.

Response:

We agree that there can be multiple sources of variation when using routinely collected data for research purposes. Further clarity around the sources of variation has now been provided.

Amendment:

3rd paragraph introduction (page 3):

"Differences in patient information management tools and data extraction tools result in variability in both the information captured and ways in which this information is coded. In particular, the way in which text values (diagnoses, examination findings, test results and medications) are transformed by systems to codes which can be a source of variation within and between studies.

Reviewer 1, Comment 25:

In relation to an earlier comment, the authors provide further clarity in this section regarding the inclusion/exclusion criteria for general practices in the POLAR database. However, it remains unclear whether all general practices in the POLAR database use the same clinical information system. It would also be helpful to understand how representative the 30% of general practices included in the POLAR database are of the wider population of general practices in South Eastern Victoria, Victoria and Australia.

The authors also provide further clarity regarding the 'data files' that comprise the POLAR database. These 'data files' appear to represent the data from the different 'tables' in a relational database. I was not clear on why merging of records from these 'data files' is a complex task once relevant records have been identified/selected. Do the primary and foreign keys associated with records in a relational database not enable this?

Response:

Data is extracted from two different clinical information systems, covering ninety percent of included general practices. This has now been clarified in the data source section of the methods.

Although approximately 30% of general practices across South-Eastern Victoria are included within the POLAR database, non-accredited, corporate-owned practices, and those not using electronic medical records are likely to be under-represented. We plan to discuss this further within the manuscript pertaining to the study cohort.

Merging of the records from the 'data files' to create a patient-centred database (i.e. to identify the study cohort) is a complex task even after the relevant records have been selected. The complexity is related to the clinical question and the sequencing of the merges rather than in the mechanics of the merge. This process requires the merging of 5 different 'data files' (patient, practice, provider, activity and diagnosis) and the sequencing of this required careful consideration to ensure we retained all relevant records (see Figure 2). For example, we could not limit diagnosis records to 2014-2018 initially as we would therefore be unable to identify relevant comorbidities (i.e. chronic disease diagnoses prior to the study period). Another example is that the provider file needed to be merged with both the activity and diagnosis files to identify activities/consultations provided only by GPs and diagnoses made only by GPs. The complexity of this task has been clarified in the approach to dataset creation section of the methods.

Amendment:

Data source (page 4):

"Data is extracted from two different clinical information systems, covering ninety percent of included general practices. All data is extracted using the Hummingbird data extraction tool."

Approach to dataset creation (page 8):

"This process will require the merging of five relational data files (patient, practice, provider, activity and diagnosis) in a specific sequence to ensure all relevant records are retained. For example, we will not limit diagnosis records to 2014-2018 until after we have selected relevant comorbidities."

Reviewer 1, Comment 26:

The authors highlight previous studies that have used the POLAR database. They highlight that 'these studies have not reported the coding approach used to classify data into meaningful categories or the processes used to merge data files into a customized database for analysis'. It would be very helpful for the authors to clarify what is meant by a 'coding approach'. For example, is this the set of codes used to represent a particular concept (e.g. back pain)? Additionally, it would be helpful for the authors to clarify what is meant by 'meaningful categories'. What constitutes 'meaningful'? Does this mean 'concepts' that are to be considered by the study? What is a category? Does this mean a particular 'class' of data, such as prescriptions, referrals etc?

Response:

We are referring to the process used to classify and select eligible records. That is, the set of codes used to represent a particular diagnosis and the system used to group or categorise codes (e.g. SNOMED). For consistency, we have rephrased as follows.

Amendment:

Introduction (pages 4-5):

“However, these studies have not reported the ~~coding approach used to classify data into meaningful categories~~ methods used to classify and select eligible records or the processes used to merge data files into a patient-centred database for analysis.”

Reviewer 1, Comment 27:

Should 'This manuscript presents a study protocol investigating...' be 'This manuscript presents the protocol for a study investigating...'?

Response:

Thanks for this suggestion.

Amendment:

Introduction (page 4):

“This manuscript presents a protocol for a study ~~protocol~~ investigating patterns of care.....”

Reviewer 1, Comment 28:

The 'framework' appears to comprise of the 'methods used to classify and identify eligible records from the POLAR database...'. This does not appear to address the issue highlighted in the previous paragraph regarding the 'processed used to merge data files'. It would be helpful if the authors were clear on what constitutes the framework, and the specific issue(s) that the framework aims to address.

Response:

As per our responses to Comments 12 and 21 above, we have now used the term 'systematic approach' instead of 'framework' to describe the methods used to classify and identify eligible records and the process used to merge data files into a patient-centred database throughout the manuscript.

Reviewer 1, Comment 29:

The introduction provides no specific information regarding musculoskeletal conditions and why it is important to understand patterns of care for musculoskeletal conditions. It appears to be written from a very general perspective, i.e. general practice as a whole, focusing on issues of coding, variation and reporting in the POLAR database. I would expect to see further context in the introduction regarding musculoskeletal conditions to provide the rationale for the study.

Response:

Thank-you for this suggestion. We have justified this protocol based on the need for a systematic approach regarding how to use the POLAR database for research purposes. More specific information regarding why it is important to understand patterns of care for musculoskeletal conditions had been added to the introduction and will be provided in the manuscripts pertaining to the results. We have also added some brief background information in the introduction.

Amendment:

1st paragraph introduction (pages 2-3):

“General practice plays an essential role in providing primary health care to the population. In Australia 86% of the population visits a general practitioner (GP) multiple times a year¹, and nearly 20% of these consultations are for a musculoskeletal condition². These conditions account for 23% of the years lived with disability in Australia³ and are also a major cause of disability worldwide⁴. Until 2016, the BEACH (Bettering the Evaluation and Care of Health) program provided the most comprehensive data on clinical activities of Australian general practice⁵. The program identified a number of activities that represent low-value care for people with musculoskeletal conditions including an over-reliance on imaging, prescription of opioids, and unnecessary referrals to specialist care^{6,7}. However, in-depth exploration of these activities within the BEACH program is limited by its cross-sectional design, and these data are no longer being collected.”

Reviewer 1, Comment 30:

The objectives subsection of the introduction provides objectives which relate specifically to patterns of care for people with musculoskeletal conditions. The objectives seem generally sensible. However, it is unclear how fulfilment of the objectives will contribute to existing knowledge and inform future care or research. Further clarity would be welcomed regarding what constitutes a ‘management type’. Additionally, with respect to the objective to ‘Examine longitudinal changes in management...’, does this refer to changes in management within or between patients, or both? Finally, should ‘during 2014...’ be ‘between 2014...’, or does the study plan to look at changes within the two stated years (2014 and 2018)? If it is the two years only, then why have these two years been chosen?

Response:

We have added information to the introduction about how the fulfilment of the objectives will contribute to existing knowledge and inform future care.

In this context, management type refers to GP consultations, referral to other relevant health care providers, prescription for pain relief and imaging requests.

Longitudinal change in management will incorporate changes between different patients and changes within the same patients followed up over more than 1 year. The focus of this manuscript and analysis is at the level of the GP rather than the patient.

During 2014 and 2018 refers to between 2014 and 2018 inclusive.

Amendment:

Final paragraph introduction (page 4):

“Study findings will advance existing knowledge about GP care for people with these musculoskeletal conditions and whether it conforms to best evidence-based practice. Differences in care across different musculoskeletal complaints may also inform tailored interventions to improve care and ultimately reduce the burden of disease associated with these musculoskeletal complaints.”

Objective 4 (page 4):

“Examine the longitudinal changes in GP management for these conditions ~~during~~ between 2014 and 2018 inclusive”

Methods

Data source

Reviewer 1, Comment 31:

See previous point regarding whether ‘data space’ should read ‘database’.

Response:

As per our response to Comment 9 above, throughout the manuscript, reference to the POLAR data space system has been changed to POLAR database.

Reviewer 1, Comment 32:

The distinction between ‘practice’ and ‘provider’ is unclear and would benefit from clarification.

Response:

Practice relates to an organisation where a patient may present for management whereas provider relates to the person providing the management e.g. GP, physiotherapist, psychologist. Provider is clarified under Variables as follows (page 5):

“This is coded by Outcome Health according to the professional background of the healthcare provider delivering the service (e.g., GP, nurse).”

Reviewer 1, Comment 33:

I would argue that the description of the POLAR database can be simplified. A description of the internal structure of the database, e.g. ‘...based on eight relational files...’, seems to be low-level for this section. The referenced table is helpful in providing a simple textual summary of the database. Whilst Figure 1 is somewhat helpful in showing the relationships between the different

tables in the schema, it may be clearer to readers if a standard approach, such as UML, was used to represent the entities and relationships that comprise the schema.

Response:

Thank-you for this suggestion. Figure 1 has now been replaced with a UML representing the relationship between each data file and its variables. Accordingly, Table 1 has been removed.

Amendment:

Data source (page 4):

“The database structure is based on eight relational files, each containing de-identified practice, provider, and/or patient codes (Figure 1). These common fields (Figure 1) allow merging of the data files so that databases can be configured for specific research purposes. ~~A list of data files and examples of variables in the POLAR database are presented (Table 1).”~~

See Figure 1. Database structure

Setting

Reviewer 1, Comment 34:

The study will include data ‘relating to all patients diagnosed with an eligible musculoskeletal condition and a face-to-face GP consultation with a GP’. It is unclear whether the diagnosis occurs within the GP consultation. From the phrasing, it appears that patient must have ‘an eligible musculoskeletal condition’ and a ‘face-to-face GP consultation’, but the relationship between the two is not clear. For example, could the diagnosis occur outside of a GP consultation? Would patients for whom this is the case be included? Also, the inclusion of ‘with a GP’ does not appear to be necessary.

Response:

To be included in the study cohort, a patient must have an eligible musculoskeletal condition diagnosed by a GP and at least one face-to-face GP consultation during the study period. The inclusion around at least one face-to-face GP consultation was to ensure the patient had received care from the GP at some stage. This has been clarified within the Setting section. However, the diagnosis did not necessarily have to occur during a GP consultation. This is because the diagnosis could be added to the patient’s electronic medical record after an inpatient hospital stay or correspondence from a specialist. This has now been clarified in the text relating to Participants.

Amendment:

Setting (pages 4-5):

“Our study will include data collected over five calendar years from January 2014 until December 2018 relating to all patients diagnosed by a GP with an eligible musculoskeletal condition and who received at least one face-to-face GP consultation ~~with a GP~~.”

Participants (page 5)

“Patients with an eligible diagnosis and age will also have received at least one GP face-to-face consultation during the study dates. The musculoskeletal diagnosis will not have to occur during a GP consultation since it is an eligible diagnosis that could result from consultation with other health care providers.”

Reviewer 1, Comment 35:

There are no details provided regarding the clinical information systems used by the practices in the POLAR database. This has been highlighted in an earlier comment. It is unclear whether the general practices in POLAR use the same clinical information system. This is important to clarify as the clinical information system used by the general practice are a potential source of variation that must be considered.

Response:

As per our response to Comment 25, the general practices in POLAR use 3 different clinical information systems, 2 of which cover 90% of the general practices. This has now been clarified in the data source section of the methods and acknowledged in the discussion as a potential source of variation.

Amendment:

Data Source (page 4):

“Data is extracted from two different clinical information systems, covering ninety percent of included general practices. All data is extracted using the Hummingbird data extraction tool⁹”.

Reviewer 1, Comment 36:

It is unclear what is meant by ‘patient-level follow-up data’ and how this is different to the other data that will be included. Clarification on this would be helpful.

Response:

Patient-level follow-up indicates that each included patient in the study cohort will be followed up until 31st December 2018. This has now been clarified in text.

Amendment:

Setting (page 5):

“Patient-level Follow-up data will be from the time of the initial recorded diagnosis to 31st December 2018 will be included.”

Participants

Reviewer 1, Comment 37:

See earlier comment relating to abstract about the use of bracketing.

Response:

As per our response to Comment 14, we have removed the bracketed text referring to age restrictions.

Amendment:

Abstract (page 2):

“Data for patients with a low back, neck, shoulder and/or knee condition and who received at least one GP face-to-face consultation between 1/01/2014 and 31/12/2018 will be included.”

Participants (page 5):

“The study cohort will include people diagnosed during 2014 to 2018 inclusive with a low back, neck, shoulder and/or knee condition, limited to age 45 years and over except for low back which will be limited to age 18 years and over.”

Reviewer 1, Comment 38:

Table 2 could be clearer in communicating the eligibility criteria. It may help to separate the inclusion/exclusion criteria for the patient population, and the inclusion/exclusion criteria for data relating to this patient population.

Response:

Thank-you for this suggestion.

Amendment:

Table 2 renumbered as Table 1 (page 13):

An additional row has been added to separate the inclusion/exclusion criteria relating to the patient population to the inclusion/exclusion criteria relating to patient management (outcomes).

Table 2 has also been renumbered as Table 1.

Reviewer 1, Comment 39:

As mentioned in a previous comment, further clarity is required regarding ‘deductive’ and ‘inductive’ approaches.

Response:

See response to Comment 15 above. A deductive approach refers to using Outcome Health’s coding approach to select eligible records. An inductive approach refers to conducting our own coding in order to select eligible records. We appreciate that this was not clear so have instead described the process we used to select eligible records.

Amendment:

“In circumstances where Outcome Health has previously coded data (e.g., diagnosis records), we used this coding to select eligible records that fitted our inclusion criteria. In circumstances where there was no coding (e.g., imaging tests), we coded the data into categories and then selected eligible records.”

Reviewer 1, Comment 40:

Further information is required regarding the approach or framework’ used by Outcome Health to code the data. Do Outcome Health manually go through text values entered by clinicians and associate codes with these values? If so, how is this undertaken? What are the components of their framework? How do they ensure consistency within and between records etc...? Validity of the study is very much contingent on these details.

Response:

The coding approach used by Outcome Health uses clinical natural language processing to automatically code structured narrative text followed by a manual process for quality checking and correction. This process has now been described at the beginning of the variables section rather than within each specific variable. We also draw the reviewer’s attention to reference 20 which outlines the processes in detail.

Amendment:

Variables (Page 5):

“Outcome Health’s approach to coding used clinical natural language processing to automatically code structured narrative text within the electronic medical record following by a manual process for quality checking and correction²⁰. For example, this allowed the free text items ‘back pain’, ‘low back pain’, and ‘lumbar pain’ to all sit under the same diagnostic code. Where possible, coding was conducted using a standardised classification system. For example, diagnoses are coded using SNOMED CT-AU terminology²¹ and prescriptions are coded according to the Anatomical Therapeutic Chemical (ATC) classification system²². In cases where there is no standardised classification system available (e.g., providers and referrals), Outcome Health used a similar process to code these variables into relevant categories (e.g., type of health care provider). Clinical natural language processing conducted by Outcome Health has previously demonstrated accurate coding of over 95% of the narrative text to SNOMED CT-AU terms in a sample of approximately 57,000 diagnosis records²⁰.

Variables

Reviewer 1, Comment 41:

Provider records appear to contain a 'provider type' field which can be used to determine whether the provider is a GP. Further details would be helpful as to the nature of this field, why 'clinical natural language processing' is required, and the specific form of this processing.

Response:

The 'provider type' field refers to the professional background of the healthcare provider delivering the service (e.g., GP, nurse). It was necessary to code this information because healthcare providers other than a GP may be nested within a general practice. Details regarding the clinical natural language processing have now been provided earlier within the manuscript.

Amendment:

Provider records (page 5):

"Healthcare providers other than a GP may be nested within a general practice. To limit all diagnoses, consultations, referrals, and prescriptions to those made by only GPs we used coding within the provider type field conducted by Outcome Health. This is coded by Outcome Health according to the professional background of the healthcare provider delivering the service (e.g., GP, nurse)."

Reviewer 1, Comment 42:

Diagnoses are coded using SNOMED-CT. It appears that 'clinical natural language processing' defined by Outcome Health was used to obtain SNOMED-CT codes from text values. Evidence is referenced regarding the accuracy of this coding process. Again, further details would be helpful regarding the specific form of the 'clinical natural language processing'.

Response:

Further details regarding the specifics of clinical natural language processing are available in reference 20 provided.

Amendment:

Details regarding the clinical natural language processing have now been provided prior to this section.

Variables (page 5):

"Outcome Health's approach to coding used clinical natural language processing to automatically code structured narrative text within the electronic medical record following by a manual process for quality checking and correction²⁰."

Reviewer 1, Comment 43:

Patients with an 'eligible musculoskeletal diagnosis during 2014-2018' were included. This appears to be repeating the inclusion/exclusion criteria for the study population. I would expect this section to provide details regarding the data items included for the study population.

Response:

This section explains the process used and the rationale for selecting eligible diagnoses. The diagnoses included within the study population are provided in Table 2. We appreciate there is some repetition but felt it was necessary to stipulate that all patients with an eligible musculoskeletal diagnosis were included regardless of whether they had a prior musculoskeletal condition.

Reviewer 1, Comment 44:

Details are provided of specific SNOMED terms that were excluded. Some of the rationale for inclusion/exclusion was unclear to me, for instance, the exclusion of general musculoskeletal terms that 'could not be attributed to a specific body region'. It would be helpful if the authors could provide further clarity regarding the specific set of musculoskeletal conditions for which they are interested in determining patterns of care.

Response:

Thank-you for this suggestion. Eligible musculoskeletal conditions are described in Table 1.

Amendment:

See Table 1

Reviewer 1, Comment 45:

Diagnoses of co-morbidities were included, which again appear based on Outcome Health's coding system. Further detail regarding this coding process would again be helpful. For example, how are 'chronic musculoskeletal conditions' identified?

Response:

Once the diagnoses are coded, Outcome Health (using experienced clinicians) has created further overarching groups – all diabetes codes into a single diabetes category. Key chronic disease groups are utilised as a qualifier as well.

Amendment:

The following additional information has now been provided along with a reference within the diagnoses records section (page 6):

"Using experienced clinicians, Outcome Health has further categorised SNOMED diagnoses into overarching groups and utilised key chronic disease groups as a qualifier⁹. For example, free text

such as 'low back pain' or 'angina' could be qualified as a chronic disease if present for six months or more."

Reviewer 1, Comment 46:

As a more general point, the patients included in the study population will be at different stages in the course of their disease in, for example, 2014. Will data relating to musculoskeletal conditions be included for a period prior to 2014? If not, how will patients at different stages in the course of their disease be distinguished? Additionally, patients may be at different stages in the course of other co-morbidities in, for example, 2014. Will data relating to these comorbidities be included for a period prior to 2014? How will the presence/absence of co-morbidities be determined? Do diagnoses need to have occurred within a particular time period (e.g. 5 years) prior to the musculoskeletal condition diagnosis that determines inclusion in the population? It was not clear whether the data would be available to robustly determine the presence/absence of comorbidities for all patients in the population.

Response:

We have identified our cohort by using the first recorded eligible musculoskeletal diagnosis by a GP from 1st January 2014. We acknowledge that some patients may have received a similar diagnosis prior to this date but that if it were a recurring musculoskeletal condition for more than six months, it would have been identified as a comorbidity. The presence/absence of comorbidities will be based on Outcome Health's coding of chronic disease groups. They do not need to have occurred within a particular time period as it is expected that a chronic disease has the capacity to remain problematic at any time and will not be limited to a particular time period.

Amendment:

We plan to acknowledge this as a limitation within the discussion of the manuscripts pertaining to the results.

Reviewer 1, Comment 47:

Coding provided within the POLAR database is used to identify for specific types of activities. It would be helpful to know the origin of values of the 'activity type' field. For instance, are they associated with the records automatically by the clinical information system, are they derived from values input by clinicians through some post-processing carried out prior to, or following, incorporation into the POLAR database?

Response:

See below

Amendment:

Activity records (page 6):

“Activity records are coded in POLAR according to the type of consultation provided (e.g., telehealth, visit, telephone). Each time a note is recorded in the narrative section it is coded by the EMR and this is extracted by POLAR.”

Reviewer 1, Comment 48:

The comment above is also relevant to the referral records.

Response:

Referral records were coded using clinical natural language processing of structured text according to the type of health care provider referred to.

Amendment:

Variables (page 5):

“In cases where there is no standardised classification system available (e.g., providers and referrals), Outcome Health used a similar process to code these variables into relevant categories (e.g. type of health care provider).”

Reviewer 1, Comment 49:

Further detail is required regarding the inclusion/exclusion criteria for medications. It was unclear why relevant medications could not be determined a priori. Why was it necessary for two study authors to determine these from the prescription file?

Response:

Medications according to ATC category were determined a priori as per eligibility criteria in Table 1. To ensure we didn't miss any potentially eligible medications, two study authors searched according to both ATC category and medication name. This allowed us to find any relevant medications that were classified into an ATC category we did not specifically search under (e.g. although fentanyl is an opioid, it is classified as an opioid anesthetic rather than an opioid analgesic) or may not have been coded to an ATC category.

Amendment:

Prescription records (page 7):

“To ensure we included all potentially eligible medication names, we searched by both ATC classification and by medication name from the prescription file during 2014-2018.”

Reviewer 1, Comment 50:

It was unclear what was meant by the 'test data file' in the description of the imaging records. It appears that Outcome Health do not, or have not yet, coded imaging tests, and that it was necessary for the authors to define a process by which imaging records could be coded in an automated manner due to the number of records. Records were classified by anatomical region

and then sub-classified according to the type of test. The authors provide details of challenges in the coding process, such as test names that did not reference a test, and provided details of how these were resolved. Authors excluded test names that they determined not to 'indicate an imaging test or procedure'. However, might there be cases where such entries provide the only reference to a test?

Response:

The test data file contains pathology and radiology tests requested by the GP. This has now been clarified. We excluded test names that we deemed not to indicate an imaging test or procedure (e.g., 'report', 'findings', 'results'). We agree it is plausible that there may be some cases where such entries might provide the only reference to a test. However, in most circumstances test names such as 'report' or 'results' were an additional reference to the actual test requested. We chose to exclude these terms because we would rather underestimate than overestimate imaging tests or procedures requested. Given that our coding accounted for more than 95% of the identified imaging test names, we accept that if in some cases these terms were the only reference to a test, this would only account for a very small proportion of imaging tests.

Amendment:

Imaging records (page 7):

"The test data file within POLAR contains pathology and radiology tests requested by the GP."

Data access and cleaning

Reviewer 1, Comment 51:

It was unclear of the relevance of the 'database records from 1997' that are referenced in this section. What is the relevance of 1997 to the study?

Response:

1997 was the inception of the POLAR database. This has now been clarified in text.

Amendment:

Data access and cleaning (page 8):

"Outcome Health provided the research team with access to all POLAR database records from since inception (1997)."

Reviewer 1, Comment 52:

Authors state that 'Data from practices proceeding a gap in reporting of one year or more during the study period will also be deemed ineligible'. Why would there be a gap in the data for a practice? Absence of data for one or more years within the period from 2014-2018 will significantly impact any ability to study longitudinal trends of more than one year. Additionally, it will impact the ability to robustly characterize patients with respect to, for instance, co-morbidities. For example,

what is a patient was diagnosed with a particular co-morbidity in year for which data is missing. Additional complexity is added by the fact that the years for which data may be missing appears to vary between practices. Therefore, robust and consistent characterization of, for instance, co-morbidities for patients across all practices will be a very significant challenge.

Response:

As per our response to Comment 2 above, there is no reason to suggest the POLAR database contains periods for which data may not be available from a particular practice. However, we agree that if this were the case, it could be a significant limitation (for all the reasons listed by the reviewer). This is why we have included a process at the data cleaning stage to assess for this and manage this situation only if it eventuates. The process we have stipulated is similar to that recommended and used by the Clinical Practice Research Datalink (CPRD) database in the United Kingdom (see reference 27 within the manuscript). This process analyses data for each practice separately in order to allow for variation between practices. We have clarified this in the Data access and cleaning section. As a side note, preliminary analysis has not identified any gaps although there were some practices that did not start general practice activity collection until some point between 2014 and 2018.

Amendment:

Data access and cleaning (page 8):

“We will exclude practices without any activity data during 2014-2018. We will also examine the consistency of activity, test, prescription, and referral data for each practice in each eligible calendar year. If a gap in reporting from any practice is identified for one year or more, only data from the earliest date after which there was no gap will be included.”

Approach to dataset creation

Reviewer 1, Comment 53:

This section seems to replicate details previously provided. As outlined in a previous comment, it is unclear why four different databases are required, as opposed to the creation of a single database over which specific queries and views can be defined to support different aspects to the analysis.

Response:

This section expands on that within the abstract. The planned analysis will be undertaken through execution of queries and we therefore do not require different databases for analysis. Reference to ‘customised databases’ has now been replaced with ‘patient-centred database’ throughout the manuscript.

Amendment:

Approach to dataset creation (page 8):

“We will use a systematic process to systematically exclude ineligible records in order to merge data and select the study cohort (Figure 2). This process will require the merging of five relational data files (patient, practice, provider, activity and diagnosis) in a specific sequence to ensure all relevant records are retained. For example, we will not limit diagnosis records to 2014-2018 until after we have selected relevant comorbidities. A customised patient-centred database will be prepared to examine.....”

Sample size consideration

Reviewer 1, Comment 54:

The authors report the number of records in the POLAR database at the time of extract. However, these numbers relate to all records, rather than the records that match the inclusion/exclusion criteria that they have previously defined. The expectation of a ‘sufficiently large sample’ is not substantiated. Additionally, it is unclear what the authors mean by ‘sufficiently large’. What sample size would be ‘sufficiently large’ to ‘detect variation in patterns of care’?

Response:

Since this manuscript is a protocol, we do not yet know the number of records that match the inclusion/exclusion criteria. We have therefore estimated an expected sample size based on the size of a previous cohort with type 2 diabetes from the POLAR database and compared this with the sample size required for our planned sequence analyses that will be used to examine overall patterns of care. We have also moved the sample size consideration section to below the analyses because this is based on the planned analyses.

Amendment:

Sample size consideration (page 10):

“Sequence analysis will require the largest sample size of our planned analyses and will therefore form the basis of our sample size consideration. We plan to examine the following six management types: non-surgical referrals, surgical referrals, allied health referrals, opioid prescription, X-ray and/or ultrasound requests, and MRI and/or CT scan requests. This provides a total of 720 potential sequence combinations. Based on a recommended 20 to 30 subjects per subgroup³¹, we estimate a sample size of between 14,400 and 21,600 will be required to differentiate between each sequence combination or pattern of care. Recent use of the POLAR database using data from approximately 200 general practices identified 20,514 active adult patients with type 2 diabetes before July 2016³². Our extract is based on 301 general practices from 2014 to 2018 and since the prevalence of diabetes is less than that of musculoskeletal conditions³³, we expect a sample size of more than 20,000.”

Analyses

Reviewer 1, Comment 55:

It is stated the ‘Patterns and timing of management...will be examined and compared between 2014 and 2018’. As mentioned in a previous comment, is this comparison between two years, i.e. 2014 and 2018, or between all years in the period from 2014 to 2018.

Response:

As per our response to Comment 30 above, this is a comparison between all years in the period from 2014 to 2018 and has been amended for clarity.

Amendment:

“Patterns and timing of management (imaging requests, prescriptions and referrals) for people with eligible low back, neck, shoulder and knee conditions will be examined and compared between each year within the five-year study period 2014 and 2018 and relative to time of diagnosis using trend analyses.”

Reviewer 1, Comment 56:

The authors plan to use ‘group-based trajectory modelling to identify groups of individuals with distinct patterns of prescription opioid use over time’. However, given the limitations of the data highlighted in previous comments, e.g. missing years for practices, it is unclear how this can be undertaken robustly.

Response:

In relation to Comment 58 below, we have decided to remove this analysis in order to narrow the scope of our analysis.

Amendment:

~~“We will use group-based trajectory modelling to identify groups of individuals with distinct patterns of prescription opioid use over time.”~~

Reviewer 1, Comment 57:

Analysis will be undertaken relative to the ‘first (index) eligible musculoskeletal diagnosis’. However, my understanding from the inclusion/exclusion criteria is that this ‘first (index)...diagnosis’ is the first occurrence within the period 2014-2018 and the patient may have had relevant diagnoses prior to 2014. Therefore, patients may be at different stages in the course of their disease at the ‘first (index)...diagnosis’ and may receive different management as a consequence.

Response:

We have identified our cohort by using the first recorded eligible musculoskeletal diagnosis by a GP from 1st January 2014. We acknowledge that some patients may have received a similar diagnosis prior to this date but that if it were a recurring musculoskeletal condition, it would have been identified as a comorbidity.

Amendment:

We plan to acknowledge this as a limitation within the discussion of the manuscript pertaining to the results rather than the protocol.

“Although we expect to be able to identify sequences of GP care provided to people with musculoskeletal conditions relative to the time at which they have first sought care from a GP, we do not expect all patients to be at a similar stage in the course of their condition. We identified our cohort by using the first recorded eligible musculoskeletal diagnosis by a GP from 1st January 2014. It is therefore possible that some patients may have sought other care prior to receiving a GP diagnosis and others may have received a similar diagnosis prior to 1st January 2014. In the latter case, if the musculoskeletal condition was recurring, we expect it would have been identified as a comorbidity.”

Reviewer 1, Comment 58:

The authors plan a large number of different analyses. The scope of these analyses appears to be very wide. The analyses may benefit from a significant degree of simplification, both in their performance and in their description within this manuscript.

Response:

We agree that the scope of these analyses is very wide but also necessary in order to gain a full picture of patterns of care provided by GPs for people with musculoskeletal conditions. This is part of the underlying reason for publishing a protocol. In order to address this, we plan on a number of different publications to report the results. This will entail 5 separate publications:

1. Description of study cohort
2. Descriptive analysis of referrals to other healthcare providers
3. Descriptive analysis of imaging requests
4. Descriptive analysis of prescriptions for pain relief
5. Overall patterns and sequences of care provided by GPs for people with musculoskeletal conditions

We have removed our planned analysis of group-based trajectory modelling to identify groups of individuals with distinct patterns of prescription opioid use over time and where possible have simplified the analyses section.

Amendment:

See amended analyses section (pages 8-10).

Discussion

Reviewer 1, Comment 59:

Whilst the authors state that the protocol ‘conforms to the RECORD guidelines by describing the coding process’, much of the coding appears to be undertaken by Outcome Health using their ‘coding framework’. Therefore, any interpretation of the concepts used within the study needs to consider both this coding framework and the inclusion/exclusion criteria detailed in this manuscript.

Response:

Thank-you for this suggestion. We have this as follows in the discussion.

Amendment:

Discussion paragraph 3 (page 11):

“Constraints within the POLAR database may potentially limit the reliability of this study’s findings although these are problems inherent in the use of any extracted data. Variability in workflows and recording behaviour introduces potential biases and the different clinical information systems used by the practices within POLAR may result in variability in the information entered. The objective of POLAR is to remove as much variability as possible by using and being transparent about the coding process. High accuracy of diagnostic coding by Outcome Health has been previously demonstrated²⁰.”

Reviewer 1, Comment 60:

It is stated that the POLAR database ‘is expected to contain a large and representative study population’. Expectation of a representative study population is not the same as providing evidence that the study population is indeed representative. If there is no existing evidence regarding whether the population within the POLAR database is representative of the wider population in Victoria and Australia, such evidence needs to be obtained prior to any studies being undertaken using the data from POLAR.

Response:

Thank-you for this suggestion. In accordance with our response to Comment 22 above, the strength of this study relating to representativeness has been reframed to extend upon previous research that has assessed the representativeness of an alternate patient population in the POLAR database. We have also clarified in the analysis section that we plan to assess the representativeness of the people with musculoskeletal conditions within the POLAR database to the wider populations of Australia by comparing the characteristics of the study cohort with the Australian population using national health survey data.

Amendment:

Analyses (page 9):

“Descriptive statistics will be used to summarise the study cohort including the number and type of eligible musculoskeletal conditions; patient demographics; and comorbidities. These will be compared to national health survey data to assess the representativeness of the POLAR database to the wide population.”

Discussion (page 10):

“The POLAR database contains all patient-related activity within each practice making it representative of the included practices. Previous research has demonstrated comparable prevalence and age-gender distribution of people diagnosed with type 2 diabetes within the POLAR database to those within Australia³². This study will add to these findings by assessing the representativeness of people with musculoskeletal conditions within the POLAR database to the wider Australian population.”

Reviewer 1, Comment 61:

The authors highlight a strength of the POLAR database as its 'comprehensive and longitudinal nature' and to 'enable temporal sequences to be examined over time'. However, from the description provided in the 'Data access and cleaning' section, it would appear that there can be significant gaps, e.g. a year, in the data available for practices. This places a significant limitation of the ability to construct robust and consistent temporal patterns for patients from different practices over multiple years.

Response:

As per our response to Comment 2 above, there is no reason to suggest the POLAR database contains periods for which data may not be available from a particular practice. However, we agree that if this were the case, it could be a significant limitation. Our approach to data cleaning follows recommendations by an established UK GP database and addresses this potential limitation. We agree that we should not highlight this as a strength since the consistency of the data has not yet been assessed. We plan to evaluate this as a strength or limitation in our reporting of the results.

Amendment:

Discussion (page 10):

~~"The main strength of the POLAR database this study is that it is expected to contain a large and representative study population of general practice within south-eastern Victoria. Its comprehensive and longitudinal nature will facilitate an overview of the care provided by GPs to the same patient(s) over time and thereby enable temporal sequences to be examined."~~

Reviewer 1, Comment 62:

Improvements discussed appear to relate more to the clinical information systems used by practices that contribute to POLAR rather than the POLAR database.

Response:

We have removed this section and re-framed the discussion to focus on the strengths and limitations of this study rather than the POLAR database in general.

Amendment:

See amended discussion (pages 10-11).

Reviewer 1, Comment 63:

There is no discussion regarding the strengths and the limitations of the specific study design. Discussion focuses on reporting of coding and on the POLAR database in general.

Response:

The discussion has been amended to focus on the strengths and the limitations of the specific study design rather than the POLAR database in general.

Amendment:

See amended discussion (pages 10-11).

Reviewer 1, Comment 64:

The study has ethical approval from the Cabrini Human REC and Monash University REC. Dissemination to Outcome Health was mentioned in the description in the abstract, but not in this section. 'Participating PHNs' are referenced in this section but not in the description in the abstract. It would improve clarity if the description was consistent between the abstract and this section.

Response:

We have amended to ensure consistency between the information provided in the abstract and in the ethics and dissemination section.

Amendment:

Ethics and Dissemination (page 11):

"The study findings will be reported to Outcome Health, ~~the~~ participating PHNs, disseminated in peer-reviewed academic journals and presented in national and international conferences."

Reviewer 2**Reviewer 2, Comment 1:**

The authors rightly say transparency in codes and algorithms in studies using EHR is important. They suggest their protocol will allow replication by other researchers. Whilst I fully agree on the importance of this, it would be useful to consider the scenarios where replication and use of these codes and algorithms may be most feasible. If I understand the protocol, the "raw" information is not coded at the time of entry, but is coded later through natural language processing. This is different to many other nations' general practice records where symptoms, illnesses and processes of care are coded as the health care professional enters information (SNOMED or ICD, for example). It would be useful to set the context in the background as to how Australian general practice recording and health care system differs to other countries; and how much what is reported here is mainly relevant to researchers using POLAR and how much is more widely generalisable.

Response:

Thank-you for this suggestion. We have added a sentence to the introduction to clarify that coding is not embedded in the clinical process and therefore needs to be conducted specifically for

research purposes. We have also added a section to the discussion to acknowledge that although the process used to identify and select the eligible study population and variables used may only be applicable to health care systems that do not embed coding within the clinical process, the approach used to examine patterns of care over time can be applicable to other conditions in general practice.

Amendment:

Introduction (page 3):

“Unlike in other countries, coding is not embedded within the clinical process and therefore needs to be conducted specifically for research purposes.”

Discussion (page 10):

“Although our coding process may only be applicable to systems that do not embed coding in the clinical process, this approach can also be adapted to examine patterns of care over time for other conditions in general practice.”

Reviewer 2, Comment 2:

I wondered why osteoarthritis without region specified was excluded? This could exclude a large group of people with musculoskeletal pain, particularly those with generalised OA who may not have a region recorded, given how common OA is in older adults. It seems a shame not to include this.

Response:

We chose not to include these participants in our cohort as one of the objectives was to compare the patterns of management for people with low back, shoulder, knee and neck conditions. Diagnosis of generalised OA would not allow us to do this but could certainly be the focus of further research. We have now clarified this within the manuscript.

Amendment:

Introduction (page 4):

“Differences in care across different musculoskeletal complaints may also inform tailored interventions to improve care and ultimately reduce the burden of disease associated with these musculoskeletal complaints.”

Objective 1 (page 4):

“Describe and compare the management (number, type and timing of imaging tests and procedure requests, prescriptions for pain relief, and referrals to other health providers) provided by GPs to people with low back, shoulder, knee and neck conditions”

Analyses (page 9):

“Results will be stratified by affected body region.”

Reviewer 2, Comment 3:

How will the researchers know that a prescription is for a relevant musculoskeletal problem? Particularly if someone has several musculoskeletal conditions recorded. Will there be a reason for prescription recorded?

Response:

Unfortunately there will not be a reason for prescription recorded. This has been acknowledged as a limitation of the POLAR database and is the reason underpinning our decision to analyse those with multiple body regions affected by a musculoskeletal condition separately to those with an eligible diagnosis affecting a single body region. We have also acknowledged the implication that not all patterns of care will be directly attributable to a musculoskeletal condition as a limitation of this study.

Amendment:

Analyses (page 9):

“One of the limitations of the POLAR database is that it does not capture reasons for the clinical encounter or management types (imaging request, prescription or referral). To account for the subsequent uncertainty in attributing management types to a particular diagnosis for those with multiple musculoskeletal conditions, participants with eligible musculoskeletal diagnoses from multiple body regions will be analysed separately to those with eligible diagnoses in one body region.”

Discussion (page 11):

“In addition, it is possible not all patterns of care for the study cohort will be directly attributable to a musculoskeletal condition because reasons for GP consultations, referrals, and prescriptions are not mandated in the source EMRs.”

Reviewer 2, Comment 4:

Will the researchers be able to separate prevalent from incident (or perhaps new episode) consultations? This seems important when characterising management.

Response:

We have identified our cohort by using the first recorded eligible musculoskeletal diagnosis by a GP from 1st January 2014. We acknowledge that some patients may have received a similar diagnosis prior to this date but that if it were a recurring musculoskeletal condition, it would have been identified as a comorbidity. We therefore expect that the first recorded eligible musculoskeletal

diagnosis will represent the first time the participant is seeking care from the GP in the majority of cases. Unfortunately the POLAR database does not allow us to directly link these to a consultation. We have now clarified that the GP consultations examined will be due to all-causes (i.e. not necessarily a musculoskeletal condition) and that a limitation of the POLAR database is that it does not capture reasons for consultations. We also plan to acknowledge in the manuscript pertaining to the results that although we have identified our cohort using the first recorded eligible musculoskeletal diagnosis, that this may not necessarily represent a new diagnosis.

Amendment:

Analyses (page 9):

“Descriptive statistics will also be used to summarise the number and type of GP all-cause consultations, imaging tests and procedures requested, prescriptions for pain relief, and referrals to other health providers for the study cohort.”

“One of the limitations of the POLAR database is that it does not capture reasons for the clinical encounter or management types (imaging request, prescription or referral). To account for the subsequent uncertainty in attributing management types to a particular diagnosis for those with multiple musculoskeletal conditions, participants with eligible musculoskeletal diagnoses from multiple body regions will be analysed separately to those with eligible diagnoses in one body region.”

Discussion (page 11):

“In addition, it is possible not all patterns of care for the study cohort will be directly attributable to a musculoskeletal condition because reasons for GP consultations, referrals, and prescriptions are not mandated in the source EMRs.”

Reviewer 2, Comment 5:

Will the latent class analysis categorising management types take into account order of management approaches and time since consultation?

Response:

Thank-you for this suggestion. Latent class analyses categorising management types will take into account both the order of management approaches (sequence) and time since diagnosis (timing). This has been clarified within the text. Upon further reflection, since we are particularly interested in both the timing and sequence of each management type, we have decided to use sequence analysis rather than latent class analyses.

Amendment:

Analyses (page 10):

~~“Sequence Latent class~~ analysis will be used to categorise management types of people with eligible musculoskeletal conditions into similar groups based on observed characteristics²⁵. This will take into account both the time relative diagnosis and sequence of each management type. We will use this to identify the most frequently used combinations and sequences of management and the patient- and practice-related variables that correlate with each management combination.”

Additional Changes

To account for differences in the follow-up period for each included participant and increasing uncertainty over time since diagnosis in linking a management type to a diagnosis, our primary analysis will examine management types during the first year after index diagnosis for each participant. We will also conduct a sensitivity analysis including the entire follow-up period. For prescriptions, the primary analysis will include the entire follow-up period because repeated prescriptions over a period of more than a year are anticipated.

Amendment:

Analyses (page 9):

“Primary analysis will include analysis of each management type provided for each participant during the first year after their index diagnosis. A sensitivity analysis will be conducted including the entire follow-up period until 31st December 2018. For prescriptions, the primary analysis will include the entire follow-up period because repeated prescriptions over more than one year are anticipated.”