Response to reviewers

This manuscript was previously submitted to the British Journal of General Practice. We received insightful and constructive comments from three reviewers in the round of revision. Below are our point-by point responses to the comments by reviewers. In the revised manuscript, revisions are highlighted by red text.

Reviewer: 1

1. The data available did not indicate whether the usual GP seen was the named GP, so it was possible to have high continuity score (i.e. UPC) with a GP other than the named GP.

The observation is correct (and we also highlight this feature in the discussion section). Nothing to address.

2. The study design is not clearly described. It was a follow-up (or cohort?) study of people who developed one of a set of chronic conditions, follow-up being for up to 4 years, with a minimum of 1 year follow-up (I think). I'm unsure when the study end date was, but recruitment began in 2009 and ceased in 2015, meaning the data are several years old. The analysis compared people with a named GP and people who did not have a named GP. Because of the choice patients and practices had over the named GP policy, it was possible that patients and practice staff had similar views on the importance of continuity for people with chronic conditions, but one group of practices used the named GP approach to promote continuity and the other practices did not.

We have revised section 2.3 to clarify that we followed a cohort of patients for up to four years after they were first diagnosed with any of the listed chronic conditions. (See also the response to the next comment.)

As usual in register studies, we need to rely on historical data. Using "old" data can be defended given the purpose of the study, which was not to describe the current level of continuity but to measure associations. Hopefully, the historical associations are informative even though the data is from an earlier period. In fact, using any later data would also force us to include the covid-19 period, which would lead to serious concerns for both interpretation and generalization. By recruiting cases between 2009 up to 2015, we are able to follow all patients for four years (see question below) and we don't need to use data from 2020 i.e., the year the pandemic started.

The reviewer correctly notes that using the named GP approach was voluntary in our study setting. This is also highlighted in the discussion section.

3. The achieved duration of follow-up should be stated. Were all patients followed-up for 4 years, or was the mean length of follow-up shorter than 4 years?

We have clarified in section 2.3 that patients were followed for up to four years; the follow-up was shorter if the patient died within four years or ceased to be registered at any primary care center in the region. All patients who remained alive and registered throughout the four-year follow up were followed for four years.

4. The UPC is a reasonable measure of relationship continuity to employ; it is probably the most frequently used measure in research studies in which information from records is available. The study is concerned with relationship continuity rather than informational or management continuity, and statement to this effect could be usefully added to the introduction.

These observations are correct. We have added information in the introduction that we studied relational continuity.

5. The patient sample is a relatively small proportion of all those who were newly diagnosed with a chronic condition (Figure S1). For inclusion, patients had to have one of a list of ten chronic conditions, but otherwise be free of chronic conditions.

These observations are correct. In particular, a large fraction of the potential study population had another chronic condition outside our set of relevant conditions, which led to a large reduction in the number of observations. We have highlighted this in the revised Supplementary material; see also the revised Figure S1.

6. Levels of continuity were quite high – 46-48% for all diagnoses and 64-68% for chronic conditions. People who did not have a named GP achieved continuity of 46% for non-chronic conditions and 64% for chronic conditions, so it seems they had a more or less 'usual GP' whether or not they had a named GP. The UPC was 0.52 in a local study in England (Hull et al. Br J Gen Pract 2022; DOI: https://doi.org/10.3399/BJGP.2022.0043), and 0.64 in a study from Amsterdam (Winkel MT et al. Br J Gen Pract 2022; DOI: https://doi.org/10.3399/BJGP.2022.0038), although levels may be higher in some countries. UPC seems to be expressed as a proportion rather than a percentage in most studies.

These observations are correct. Nothing to address.

7. The sample size is rather small for studying the secondary outcomes of ED visits, out-of-hours visits, hospitalizations and mortality.

We do not fully agree with the reviewer here. There are 66,063 observations in these models. Table 3 shows that 18%, 35%, and 23% of the sample made an out-of-hours visit, visited the ED, or was hospitalized during the four-year follow-up. Thus, these are not rare outcomes in our sample. When it comes to mortality, the proportion of people dying is low (1-4% over the follow-up period), so the power issue is more concerning; however, in practice, we find statistically significant differences in the final two years of the follow-up, suggesting that the sample size is adequate.

8. The units in Table 1 and 2 appear to be sample moments. These are likely to be difficult for many readers to interpret.

It is correct that these are sample moments. We believe that the mean is easy to interpret for most readers, but we agree that the variance and skewness may be challenging. Note that the standard deviation, which readers may be more familiar with, can be approximated by the square root of the sample variance. Notably, the main purpose of presenting the three moments is to show that the entropy balancing strategy manages to generate adequate weights. In the revised manuscript, we have expanded on the text in the manuscript to clarify what the moments are (see Table 1; note that

this tables now also incorporates the information from Table 2 in the previous submission).

9. Review by a statistician is needed. The study uses the method of moments which is an alternative to the method of maximum likelihood. I am not an expert in this approach, and the views of a statistician would be wise. Entropy balancing weights are a means of matching the two groups of patients (with and without a named GP). What is the value of reporting an unadjusted model (model 1), or indeed model 2? The descriptive statistics appear to present proportions, but the table captions do not make this clear. Presumably the figures in the top line of Table 4 and Table 5 are beta coefficients for the comparison of having versus not having a named GP. Beta-coefficients for other variables are not included.

Actually, we did not use the method of moments estimator – we used least squares for estimation. We suspect that the reviewer was misled by the term "moment", which here simply refers to the properties of the distributions studied. In statistics, the mean/average is called "the first moment", the variance is "the second moment", the skewness is the third moment etc. To avoid confusing readers, we have replaced the term "linear regression" by "least squares regressions" in section 2.6.

The value of reporting an unadjusted model can of course be disputed. We included it partly for transparency, and partly to highlight how the outcomes of a simple mean comparison can lead to haphazard conclusions. We think that it is helpful to show how patient and practice characteristics combined explain the unadjusted difference.

The descriptive statistics in Table 3 are means and standard deviations. For the binary variables, it is true that the mean is the same as the proportion. We have noted this in the table note.

Regarding Table 4 and 5, the referee correctly inferred that the figures in the top lines are the beta coefficients. To clarify this, we have added this information in the table notes. Please note that there are no other beta coefficients to report: in model 2, patient characteristics are taken into account of via the entropy balancing weights, and in model 3, the practice characteristics are captured by the practice fixed effects.Using weighted regression (model 2) means that we accomplish the same thing as if we had entered the variables are controls, i.e., adjustment for confounders, but we do not obtain any estimate of the partial correlations between the patient characteristics (such as age, for instance) and continuity of care. The practice fixed effects (model 3) are essentially a set of dummy variables, one for each practice. By including such variables, we compare patients within the same practice. Thus, this model adjusts for observable as well as unobservable practice characteristics. Note that this implies that we obtain no estimates of the partial correlations between specific practice characteristics (such as rural location, for instance) and continuity.

10. What primary care centre characteristics were used in the adjustments (model 3)? No information on primary care characteristics is provided in the descriptives tables. Table S3 presents data on patient characteristics but not practice characteristics. Continuity is affected by size of practice, the numbers of GPs and whether they work full- or part-time, the numbers of support staff and their roles in consulting with patients. Which if any of these were accounted for? Patient demand is also a key determinant of achieved continuity, with continuity falling as the demand for care from the population increases.

Our statistical approach in model 3, which includes practice fixed effects, implies that we account for all practice characteristics that are constant over the study period. Typically, the examples provided by the referee fall into this category (as the practice size, staff etc change little over time for most practices). We elaborate on the interpretation of the fixed effects estimates in the section Statistical analysis in the revised manuscript.

Reviewer: 2
Major points

Introduction

1. Study motivation: The introduction might benefit from further highlighting the controversy on the necessity of the named GP policy/arrangement for the important patient subgroups in primary care (i.e., patients with chronic conditions in this study).

In the introduction of the revised manuscript, we elaborate more on the need to rethink primary care to realise continuity in modern-day general practice, with reference to the report suggested by the referee (see comment 8.).

Method

2. What is the operational definition of the onset of a specific disease? Was a "wash-out" period applied to exclude those with pre-existing diagnoses before the index date?

We have clarified in section 2.3 of the revised manuscript that the definition of onset was the first date when a diagnosis code in our set of relevant conditions was observed for a patient, going back as far as we could before the index date (i.e., back to 2005). Notably, this means that the length of the wash-out period differs between patients with different index dates.

3. What is the timeframe used for the definition of exposure? The same date/month/year of the first diagnosis of a specific disease? It would be better to elaborate more on the Figure S1.

We have clarified in section 2.3 that the patients' exposure status was not updated during the follow-up period, but treated as a time-invariant feature given by the status on the index date. Note that updating the exposure status would be problematic empirically, as the decision to change status may be driven by the experience of previous encounters with one's named GP or other GPs at the practice. That is, patients' decisions not to remain registered with at named GP essentially "a part of the treatment".

4. The outcome measure: The measurement of UPCI might be influenced by the length of follow-up. The author conducted a sensitivity analysis where the UPCI was estimated with a shorter follow-up of 2 years, which did help to test the robustness of the results. What is the distribution of the length of the follow-up period of the participants in the primary analysis? When is the end of the follow-up? The author mentioned "each patient could be follow-up for up to +/- 4 years around the dates of first chronic disease diagnosis", which needs to be further clarified. Perhaps it would be helpful to add a figure for the timeline to illustrate the study design.

We have added information about the distribution of the length of follow up for the participants with at least three visits (section XX). Note that, because it is more likely to reach the threshold of at least three visits if one remains uncensored, the level of attrition was lower for this is subset of the study

population. Overall, among the patients with at least 3 visits, 95% were uncensored throughout the whole four-year period; 97% were uncensored at least 3 years, 98.5% at least two years, and 99.5% at least one year. There was no statistically differential attrition (either in terms of death, no longer being registered at a PCC, or the combination of the two reasons for attrition) after accounting for patient and PCC characteristics in this subsample (added in Results section). We have also included a table (new Table 1) in section 2.3 to clarify the timeframe used to compute outcomes for censored and uncensored observations.

Results

5. Table 1: What is the sample size in different groups (with/without named GP)? The timeframe used for some covariates seems confusing and misleading. Better to add a footnote for clarification.

We have added information on the sample size. We have also revised a table note to explain the timeframe of the covariates.

6. Table 3: The numbers of observations seem confusing here. Better to elaborate more on the structure of the dataset and the observations referred to for each outcome in the table. It might be helpful to indicate the actual sample size in each group.

The table does display the actual sample size for each variable, it is just that some outcomes were not defined for patients with fewer than three primary care visits. We have clarified this in the table note to explain why the number of observations vary.

7. Table 4 and 5: The table might need to be reshaped to avoid misunderstandings. It would be better to put the model specification above the model estimates.

We have expanded the table note to clarify that model 1 (columns 1 and 4) is unadjusted, model 2 (columns 2 and 5) applies the entropy weights, and model 3 (columns 3 and 6) includes the fixed effects.

Discussion

8. Implication: From another perspective, the findings of this study also implied that a more flexible model/pattern for continuity of care could serve as an alternative in situations where the named GP policy is not available. https://bjgp.org/content/66/649/396

This is an interesting angle on our results. In the revised manuscript, we mention at the end of 4.4 that the named GP policy is neither a necessary nor sufficient prerequisite to realise continuity.

Minor points

9. I would suggest merging Table 1 and Table 2 in the same table for comparison (i.e., before and after balancing).

We have merged Tables 1 and 2 into one.

Reviewer: 3

1. The authors conclude that having a 'named' GP does not lead to a higher UPC index. However, there's no information on the policy of PCP's for deciding which GP sees a certain patient with a chronic disease with a need for a consultation. Most PCPs might have a policy for this situation and that might lead to their finding that UPC is not enhanced when there is a 'named' GP.

In response to this comment, we have noted among the study limitations that one potential interpretation of the finding is that PCPs have another policy that achieves the same purpose. Unfortunately, we have no information on the routines used at the 150 practices in the region (which may also have changed over the period).

2. They do no reflect on whether the degree of UPC they found in their study is satisfactory or too low in relation to the benefits associated with continuity of care. i think such a reflection would add to the usefulness of their discussion paragraph. However, their conclusion that a mandatory 'named' GP will not lead to a higher UPC is correct based on their findings.

It falls outside the scope of the study to judge whether the observed degree of UPC is satisfactory is outside the scope of the study. However, in response to the comment, we now acknowledge in the discussion section that the UPC is far from the maximum, suggesting that the results are not due to a "ceiling effect" and that there is room for improvement.

3. It's not entirely clear how they calculated the UPC. I also miss a reference for this outcome measure.

We have clarified the description of the UPC. We also provide a reference, (Pollack et al. 2016).