Author’s response to reviews

Title: Regional differences in reasons for consultation and general practitioners’ spectrum of services in Northern Germany – Results of a cross-sectional observational study

Authors:

Ingmar Schäfer (in.schaefer@uke.de)
Heike Hansen (h.hansen@uke.de)
Thomas Ruppel (thomas.ruppel@gesundheitsrecht.de)
Dagmar Lühmann (d.luehmann@uke.de)
Hans-Otto Wagner (h.wagner@uke.de)
Agata Kazek (a.kazek@uke.de)
Martin Scherer (m.scherer@uke.de)

Version: 2 Date: 17 Dec 2019

Author’s response to reviews:

On the following pages you will find answers to each comment. To facilitate reading, the original comments were copied into the text. All changes are highlighted in the manuscript.

Reviewer 1: Catherine Darker

1.1. Every country grapples with the concept of healthcare utilisation and the readiness of a country's system to respond effectively to that community needs. This is the essence of health system planning. This is particularly important in primary care - the first point of contact of people with their health service. - With this in mind I think that you may have to explain a little context for your international readers in relation to Germany and how the primary care services are formed and funded. This can differ from country to country. For example, in Ireland GP's are primarily independent small businesses that have contracts of service to the national health service provider; whereas in the UK GPs are employees of the State. This is important in terms of the flexibility and therefore the responsiveness of the GP to alter the type of service they are providing to meet the needs of their population. Can you please provide enough information so as that readers can contextualise this situation in Germany? For example, I am thinking that a lot of it is provided through health insurance - are some health insurers more/less likely to cover/refund particular types of procedures over others? This may be an alternative explanation for some of the findings. I think this is important to be able to comment on the interaction effects between presentation of disease and provision of services.

Thank you for this helpful suggestion! We now included a paragraph giving context for international readers: “In Germany, for example, GPs are mostly self-employed in independent small businesses. They can work individually in private practices (individual practice) or together with other physicians – either in private practices where each physician submits his own claims (group practice) or in private
practices where the claims of all physicians are combined and submitted as one bill (joint practice). It is also possible that GPs work as employees or self-employed in medical care centres (“Medizinische Versorgungszentren”). In the statutory health insurance system, which includes approximately 90% of the German population, about 95% of the services are determined by the Federal Joint Committee. The remainder of the services can be defined individually by the statutory health insurance companies based on strict legal requirements [1].”

1.2. You state that the study was based on GP and patient interviews but only present GP data - why exclude patients data?

We used the same standardized instrument for assessing consultation reasons and services in patients and GPs, but assessed the patient data on the individual level (=each patient rated his own consultation reasons and services within the last 3 months) and GP data on the collective level (=GPs described their complete practice population). However, consultation reasons and services were manifold and heterogeneous on the patient level. As we had a comparably small sample size (811 patients, which corresponds to a mean of approximately 4 patients per practice), we did not consider this sample as representative for the respective practice population with regard to consultation reasons and services. In our view, the GP interview therefore was the only appropriate data source for the analysis of these data.

1.3. Why an initial plan to recruit 80 GPs per region?

We already addressed this question in our answers to the comments in the first revision, in which we answered: “As pointed out in the study protocol, due to the investigation of multiple outcomes and the observational character of the study, we could not make a sample size calculation. We therefore planned the sample size based on our experience with similar studies.”

The limitations of the predefined sample size is already discussed in the strengths and weaknesses section: “As this was an observational study with multiple outcomes, we were unable to carry out a sample size calculation and therefore might have missed some differences between the regions due to limited statistical power.”

1.4. I understand that the study was interview based rather than audit - can you comment on why you did not verify even a subset through a clinical audit of records? This would have determined the validity and reliability of the responses you were getting from the interviews.

We decided against such an audit for three reasons. First, it is difficult to extract the data needed for our study from the EDP systems. There are different software programs, of which many are not capable of displaying summary statistics. Also, the regular documentation in the practices does not correspond to the standardized instruments in our study. For example, in Germany, about 90% of the services in the GP practice are paid by flat rate. Most categories in our assessment of services would therefore have missing values if we analysed the remaining summary statistics. Additionally, consultation reasons are different from ICD diagnoses displayed in summary statistics and can only in part be translated into ICPC, which also would cause many blanks in our analysis. The only valid way to perform such an audit would be to select a large, representative sample of each practice population and to perform an analysis of the free text entries of the GPs for each patient, which would need a lot of time in each practice – and therefore would have needed more personal and financial resources than available in the study.
Second, data collection was already very time consuming for the study GPs. In addition to the interviews we also conducted patient recruitment in each practice for which GPs (together with our staff) had to create a list of all patients from the GP’s EDP system (which often required technical support by the software companies), check the patients’ inclusion and exclusion criteria, make a random selection of patients and address and send the recruitment letters. Effectively, we spent between 4 and 8 hours in one to three visits in each practice during which the GPs had to support our study – despite the fact that the practices were open in this time frame and they also had to care for their patients. Conducting an additional audit of clinical records would have resulted in a higher time demand for GPs and therefore less GPs willing to participate in our study and more GPs discontinuing study participation.

Third, we know from experience in another study in which we performed such an audit that our ethics committee would only approve the audit if we obtained informed consent from all patients of which we analyse data. This would have delayed our study and would have made it very difficult to stick to the time plan of the project.

However, we agree that verifying the interview data by an audit of medical records would have been a good way to determine the data quality. We therefore now added the missing audit to our discussion of the strengths and weaknesses of the study: “As in all studies based on interviews, the physicians’ answers might also have been influenced by memory gaps, errors or social desirability. This relates, in particular, to the assessment of the consultation reasons and the service spectrum which were mainly retrieved from memory and which could not be verified through a clinical audit of records.”

1.5. What role does deprivation have to play in the study findings? I know that you did quota sampling looked at balancing out by density the three areas of urban, rural and environs but how does this address the important feature of presentation/consultation frequency and also the type of presenting complaint?

Thank you for this important comment. The quota sampling was not made to balance out the three areas of urban, rural and environs, but to represent underserved and oversupplied administrative districts in the data set, which should also address the differences between deprived and undeprived districts. We now revised the respective passage in the methods section in order to clarify the way in which the quota sampling was conducted: “The goal was to represent as many individual administrative districts of the survey area as possible. We stratified the sample a) by regional category based on a pre-defined sample size of at least 80 GPs in each category and b) within the regional categories by counties and independent cities proportionally to the respective population size in each district. In each regional category, a 25% maximum deviation from the recruitment plan was accepted [11].”

Additionally, we added a passage about deprivation into the strengths and weaknesses section: “By stratifying the sample by the individual counties and independent cities we were able to represent both, medically undersupplied and overserved as well as socioeconomically deprived and undeprived districts in our data set.”

Reviewer 2: Sofia Eilat

2.1. Methods: One of the comments in the previous review was related to the questionnaire that is still not supplemented but described. There is still some concern related to the part dealing with reasons for visits. This is a crucial part in the questionnaire and I feel there is still a gap in the provided information: How was the question phrased? Is recall the only way to have the information? Is there information from the medical records?
We now included a translation of the phrasing of the questions into the methods section: “Additionally, we documented 99 different reasons for consultation from 17 areas / organ systems (“How often (per day/week/month/year, rarely, never) do you see patients with the following reasons for consultations (including home visits)?”) and healthcare services involving 38 different procedures (“How often (per day/week/month/year, rarely, never) do you provide the following services (including home visits)?”). This assessment was based on a standardised instrument developed on the basis of the International Classification of Primary Care (ICPC-2) [11;13-14]. The resulting data were described as the proportion of the respective consultation reasons or the respective services of all consultation reasons or services documented at the respective practice.”

The question if recall was the only way to have the information or if there was information from the medical records was already answered in the manuscript. We changed the wording in order to clarify this: “Data were supplied by memory recall. However, GPs were allowed to check their medical records if they considered it necessary.”

2.2. Results: The description of the various ways of employment (p. 6, l. 1-11) fits more to the method section.

We now moved the description of the various ways of employment from the results section into the introduction section, because it was made part of the description of the context of health care, which was suggested by reviewer 1 (cf. our answer to comment 1.1.).

2.3. Conclusions: This section too long and repeats the results and bring unnecessary evidence

As suggested, we shortened the Conclusions section. Sentences repeating the results and discussing the evidence were moved to the Discussion/Comparison with the literature section.

2.4. Tables: In table 4 and 4A needs to add a description of the used models

We did not find any missing information in the mentioned tables. In our opinion, the statistical model used for table 4 and D (formerly A4) is already described as each independent variable constitutes one line in the table. Transformations are described in the respective lines. The independent variable and the statistical methods are mentioned in the headline of the table.

2.5. Tables: The order of the tables is confusing the 4A comes before 3.

We now changed the numeration of the tables in the additional files from “A1 – A4” to “A – D” in order to have a more consistent order of the tables.