Reviewer's report

Title: Efficiency and effectiveness evaluation of an automated multi-country patient count cohort system

Version: 2  Date: 19 January 2015

Reviewer: Björn Schreiweis

Reviewer's report:

Major Compulsory Revisions
1. Line 186ff.: How was the estimation of the number of patients visiting the clinic per year and being eligible to participate estimated?
2. For the results on the effectiveness you should at least give a short summary on the responses to the questions on the common and current PF process.
3. Line 275ff.: What are the boundaries of the CI? Mention them also in the text.
4. Line 283f.: Is the lower boundary equal to zero?
5. Line 298ff.: I think the sentence is already part of the discussion. It should be only part of the discussion.
6. Line 319ff. you mention that accurate results are only possible with well documented and structured data in the EHR. I think there is sufficient literature to support this thesis as it is common understanding. Mention at least one article. Otherwise link to the differences between UKM and AP-HP.
7. In line 325ff. you mention an important benefit from the EHR4CR PF, but I think your conclusion could be even more like “the system reduces the workload on personnel from the clinics.”
8. Can you explain why you had to create your own terminologiy instead of using several of the many standardized already available?
10. Line 383f.: This sentence is not understandable without comparing your tables. Describe it a bit more clearly, so it is easily understandable in the text or add a table comparing the results of UKM and AP-HP in one table.
11. Line 394-400: Is already in the results section. Remove it from the results as it fits in better in the discussion.

Minor Essential Revisions
1. Line 167ff.: Try to rephrase the sentence as right now readers might get
confused by the mixed use of anonymity and pseudonyms

2. Line 173ff.: Instead of “… evaluation were to be the same 10 … the data used would be from the year 2012.” Better use: “… evaluation were the same 10 … the data used was from 2012.”

3. Line 208f.: “The temporal constraints of the queries were adapted according to the date shift of the EHR4CR data elements previously explained.”

4. Numbers up to twelve should be written as words.

5. Line 305: “…patients and none false-…”

6. Line 323ff.: instead of “… using the EHR4CR PF system has proven to be not only faster …” better “…using the EHR4CR PF system has proven not only to be faster …”

7. Line 343: It should be “tests” instead of “test”

8. Line 347: As you are talking about the EHR4Cr consortium it should be “is has” instead of “they have”

9. Line 363ff.: “… the ETL processes need to be improved …”

Discretionary Revisions
1. Line 405: “no” instead of “none”

2. Line 407: better “despite” than “in spite of”

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests.