PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The Relationship Between Quality of Care and Choice of Clinical Computing System: Retrospective Analysis of Family Practice Performance Under the UK's Quality and Outcomes Framework
AUTHORS	Kontopantelis, Evangelos; Buchan, Iain; Reeves, David; Checkland, Kath; Doran, Tim

VERSION 1 - REVIEW

REVIEWER	Ross Koppel, Ph.D, FACMI
REVIEW RETURNED	23-May-2013

GENERAL COMMENTS	This is a work of serious scholarship, well written, and well presented. Nevertheless, I'm torn about the piece and its findings. 1. As they admit, it's unclear if the "quality" differences they find are due to the way the different systems report quality measures or if they systems have differential effects on quality of care. 2. As they also admit, the differences – which seem massive as presented in the paper – are only a few percentage points….often just 1.4%, but appear statistically significant because the sample size (really population size) is so massive. So, the title of this paper could be easily changed from "significant differences found across EHRs" to "no meaningful differences found across EHRs" – a reality they acknowledge. They could have begun to address this by examining the way the systems work (which they note they did not) but who the heck am I to tell them to write a different paper? It's more than enough to offer this good work, and suggest to others (or for themselves) to perform the needed additional work to help answer the question.

REVIEWER	Rogers, Anne University of Southampton, Faculty of Health Sciences NO competing interests. I have published in the past with two of the authors and am currently a co-applicant on a grant with David Reeves
REVIEW RETURNED	03-Jun-2013

GENERAL COMMENTS	This is a very interesting article which explores the impact of computing systems on reporting of QOF outcomes. This is clearly discussed and presented. It would add even more to the article if the implications of the findings could be drawn out more. For example does it matter that the discpreancy in systems results in £602 more for some practices rather than others? This seems a small amount so is the difference important or suggestive of a fairly equitable

	system. The socio-demographic differences in reporting of outcomes is worthy of further note - e.g. what might be the implications of the gender bias? The discussion section is rather short at present so there is room to expand the discussion and implications.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Ross Koppel, Ph.D, FACMI
Univ of Pennsylvania, Sociology Dept & Sch of Medicine
This is a work of serious scholarship, well written, and well presented. Nevertheless, I'm torn about the piece and its findings.
1. As they admit, it's unclear if the "quality" differences they find are due to the way the different systems report quality measures or if they systems have differential effects on quality of care.
Response:
Unfortunately, that's more usually than not the case in observational studies in which the chance for bias is higher and the mechanisms of differences are difficult to pinpoint. Nevertheless, some (if not more) clinicians argue that improved measurement and recording is, at the very least, a necessary prerequisite for 'real' improvement in quality of care – if not actual improvement in quality of care.

2. As they also admit, the differences – which seem massive as presented in the paper – are only a few percentage points….often just 1.4%, but appear statistically significant because the sample size (really population size) is so massive. So, the title of this paper could be easily changed from "significant differences found across EHRs" to "no meaningful differences found across EHRs" – a reality they acknowledge.
Response:
The reviewer probably means the key message in the article summary box. We agree about significance and we have tried to avoid whenever possible. However, although the differences might look rather negligible they must be considered in the UK population context and possibly in relation to the strength of other predictors: a) even a small difference, as we discuss in page 15, affects many thousands of patients nationally; b) we found it extraordinary that of all the potential predictors of performance we included in the models, choice of system was the strongest (stronger than area deprivation, area of GP qualification, etc.). So we feel the differences, albeit appearing small, should not be disregarded. We have made changes to the messages in the article summary box (page 17) and to the discussion (page 15) to better reflect this.

3. They could have begun to address this by examining the way the systems work (which they note they did not) but who the heck am I to tell them to write a different paper? It's more than enough to offer this good work, and suggest to others (or for themselves) to perform the needed additional work to help answer the question.
Response:
We did consider delving deeper into this, in the context of this paper, but we felt there were too many additional questions and corresponding analyses that would lead to a report-size publication. Therefore we decided to highlight the fact that there exist some small but not negligible system differences in the first instance, and hopefully further explore mechanisms in future work.

Reviewer: Anne Rogers
University of Southampton, Faculty of Health Sciences NO competing interests. I have published in the past with two of the authors and am currently a co-applicant on a grant with David Reeves
1. This is a very interesting article which explores the impact of computing systems on reporting of QOF outcomes. This is clearly discussed and presented. It would add even more to the article if the implications of the findings could be drawn out more. For example does it matter that the discrepancy in systems results in £602 more for some practices rather than others? This seems a small amount so

is the difference important or suggestive of a fairly equitable system. The socio-demographic differences in reporting of outcomes is worthy of further note - e.g. what might be the implications of the gender bias? The discussion section is rather short at present so there is room to expand the discussion and implications.

Response:

We have now expanded the discussion section to further address these two points. The differences in remuneration are indeed negligible, in the context of the average practice bill. In addition we have further discussed the main socio-demographic differences and referenced previous work that examined recorded diabetes care at the patient level and quantified differences for various population groups (including males and females). That work found that differences were relatively small for most groups and tended to further diminish over time.