

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)	Understanding negative feedback from South Asian patients: experimental vignette study
AUTHORS	Burt, Jenni; Abel, Gary; Elmore, Natasha; LLoyd, Cathy; Benson, John; Sarson, Lara; Carluccio, Anna; Campbell, John; Elliot, Marc; Roland, Martin

VERSION 1 - REVIEW

REVIEWER	Claudia Cooper UCL, UK
REVIEW RETURNED	31-Jan-2016

GENERAL COMMENTS	Interesting and well reported. I would suggest making it clear in the methods whether or not the scales were completed on the computer or in an interview - the discussion talks about social desirability so I assume the latter - if so would help to say in methods. No other suggestions
-------------------------	--

REVIEWER	Khaled Mohammed Mayo Clinic, USA
REVIEW RETURNED	22-Mar-2016

GENERAL COMMENTS	I have 2 points need more explanation: 1. In the methods section, please provide a clear description of the study sampling and sample size calculation. Outcomes of interest need to be defined clearly. 2. in the discussion section, please expand more about the limitations of using vignette design.
-------------------------	---

REVIEWER	Glenys Parry School of Health and Related Research University of Sheffield UK
REVIEW RETURNED	10-May-2016

GENERAL COMMENTS	This is a well-designed, carefully conducted and clearly reported study. It has some limitations as an analogue study but these are adequately acknowledged and discussed. The results are unexpected and of significant interest.
-------------------------	--

	<p>This paper reports an experimental analogue study to investigate whether the poorer GP care experience consistently reported by South Asian patients is accounted for by ethnic group differences in reporting, rather than differences in care quality. Of its type, this is a well-designed, carefully conducted and clearly reported study. Equal numbers of people from White British and Pakistani backgrounds, stratified for age (55 years or above and under 55), rated video vignettes of GP consultations scripted to demonstrate higher or poorer quality communication. The findings were surprising, in that they were in the opposite direction to that found in patient surveys of GP care; Pakistani respondents rated the 'poor quality' consultations more favourably than did the White British. This effect was particularly marked for the older respondents and was not accounted for by socio-demographic covariates.</p> <p>The analogue study defends against threats to internal validity by standardising and experimentally manipulating the communication quality of the consultation. The extent to which it is externally valid, so that its results can be used to explain the survey results, depends on how well the analogue represents the real-world situation. This study was able to investigate communication between GP and patient; however this was only one aspect of the quality of GP care assessed in the GP patient survey and this perhaps should be acknowledged in the discussion.</p> <p>The study design depends on the premise that it is possible to disaggregate 'objective' quality of communication from the patient's reported experience of communication. Whilst of course it is possible to achieve reliable consensus among external raters on 'good' and 'poor' communication in videos of GP consultations, it is arguable that what is actually experienced by the patient is what constitutes quality. This opens the possibility that what the patients in the survey are reporting as poorer communication may genuinely differ from the 'poor communication' demonstrated in the vignettes.</p> <p>Whilst the paper is admirably clear, there were two points which could be considered. First, it is not clear why only same-ethnicity pairs were used in the videos rather than mixed ethnicity. Second, the title of the paper would be more accurate as 'Understanding negative feedback from South Asian ethnicity patients' rather than 'minority ethnic patients', as other ethnic minorities were not investigated.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Claudia Cooper, UCL

Interesting and well reported. I would suggest making it clear in the methods whether or not the scales were completed on the computer or in an interview - the discussion talks about social desirability so I assume the latter - if so would help to say in methods. No other suggestions.

Author response:

We have inserted the following sentence into the methods section when describing the data collection procedures: "All interview questions and ratings were completed verbally, with responses recorded by interviewers directly onto the CAPI software. "

Reviewer 2: Khaled Mohammed, Mayo Clinic

1. In the methods section, please provide a clear description of the study sampling and sample size calculation. Outcomes of interest need to be defined clearly.

Author response:

We have added clarification of our sampling approach to the existing description of our sample size calculation. This section now reads: "We worked with a UK market research company, Ipsos MORI, to collect the data. We aimed to recruit 1,120 adult respondents who self-identified as either Pakistani or White British, across a broad age range. Each respondent was asked to rate three vignettes. Our sample size calculation was based on data from the national GP Patient Survey, as we used the same communication questions for our respondents as are used in this national survey. Inclusion of 560 Pakistani respondents and 560 White British respondents gave over 80% power to detect a 3.1 point difference (on a 0-100 scale) seen between these two groups after controlling for age, gender, deprivation, self-rated health and practice. As ethnic disparities are largest in older ages, we aimed to recruit equal numbers above and below the age of 55 within each ethnic group."

We have additionally added a sentence at the beginning of the methods section to clarify the primary outcome of interest. This now reads: "We undertook an experimental vignette study in which videos of simulated GP-patient consultations were shown to two groups of people, who were asked to rate the quality of the communication within each consultation. The primary outcome of interest was communication score (on a scale of 0-100)."

2. In the discussion section, please expand more about the limitations of using vignette design.

Author response:

We have expanded the rationale for (and limitations of) using the vignette design in the discussion section. The amended paragraph now reads:

"Previous examinations of inequalities in patient experience between ethnic groups have commonly relied on real-world data such as that generated through surveys, in which it is difficult to distinguish whether differences are attributable to variations in care or variations in the reporting of that care. 3-9 Large-scale video recording of actual GP-patient consultations, an external assessment of their communication quality, and the comparison of this to reported patient experiences of care would enable us to develop a more robust "real-world" understanding of the drivers of variations in reported experience, but the utility of such an undertaking must be balanced against its many challenges. Our experimental design enables us to control the content of the consultations being rated by respondents in order to efficiently explore how differences in reporting may explain the disparities in minority ethnic experience in real-life surveys. We chose to focus on communication as this is a key component of quality of care, yet one where certain minority ethnic groups report consistently poor experience of their interactions with clinicians. 7-9 The study builds on previous vignette research by using multiple video vignettes manipulating several key attributes. 12,13 Video vignettes have so far been little employed in this field, in spite of evidence of viewers perceiving them as realistic and enabling immersion in the situation at hand, although well-crafted vignettes are essential to ensure good construct validity. 14 In the US, Weinick et. al. reported no evidence of differences among White, African American and Latino evaluations of doctor-patient communication in vignettes when using an "Always-to-Never" response scale; they concluded that variations within national surveys on such

items for these groups were likely to reflect differences in real-life experiences. 13 In our study, however, we found substantially more positive ratings by Pakistani in comparison to White British respondents.”

Reviewer 3: Glenys Parry, University of Sheffield

This is a well-designed, carefully conducted and clearly reported study. It has some limitations as an analogue study but these are adequately acknowledged and discussed. The results are unexpected and of significant interest.

The analogue study defends against threats to internal validity by standardising and experimentally manipulating the communication quality of the consultation. The extent to which it is externally valid, so that its results can be used to explain the survey results, depends on how well the analogue represents the real-world situation. This study was able to investigate communication between GP and patient; however this was only one aspect of the quality of GP care assessed in the GP patient survey and this perhaps should be acknowledged in the discussion.

Author response:

We deliberately set out to focus on communication as this is a key driver of satisfaction with care, and an aspect of the quality of care where we see large and consistent differences between certain minority ethnic groups and the White British majority. We therefore were not attempting to explain all survey results, but those relating to communication alone: we have added this clarification to the discussion section as follows: “We chose to focus on communication as this is a key component of quality of care, yet one where certain minority ethnic groups report consistently poor experience of their interactions with clinicians. 7-9”

The study design depends on the premise that it is possible to disaggregate ‘objective’ quality of communication from the patient’s reported experience of communication. Whilst of course it is possible to achieve reliable consensus among external raters on ‘good’ and ‘poor’ communication in videos of GP consultations, it is arguable that what is actually experienced by the patient is what constitutes quality. This opens the possibility that what the patients in the survey are reporting as poorer communication may genuinely differ from the ‘poor communication’ demonstrated in the vignettes.

Author response:

We acknowledge this point, and agree that patients and “objective” external assessors may draw on different concepts of what constitutes “good” communication – we currently have another paper under review which examines precisely this point. However, our manipulation of the vignettes we used to include both “poor” and “good” communication was undertaken simply to generate a range of scenarios to ensure generalisability of our findings. It would be possible to conduct this study with only one vignette, but that runs the risk of accidentally including something within the vignette that is unacceptable to one or other group of respondents – to avoid this, we created sixteen. As we state in the paper: “To ensure generalisability and to avoid the chance inclusion of a characteristic or event which, unknown to us, might systematically be rated differently by our two groups of participants, we produced a series of 16 vignettes for this study.”

Whilst the paper is admirably clear, there were two points which could be considered. First, it is not clear why only same-ethnicity pairs were used in the videos rather than mixed ethnicity. Second, the title of the paper would be more accurate as ‘Understanding negative feedback from South Asian ethnicity patients’ rather than ‘minority ethnic patients’, as other ethnic minorities were not investigated.

Author response:

We have added a clarification of our use of same-ethnicity pairings only to the ‘simulated consultations’ section in the methods: as outlined above, the characteristics of the participants are not a crucial part of the design, but were simply varied to ensure generalisability: “The restriction of vignettes to same-ethnicity pairings, rather than including mixed pairings, is a function of wishing to introduce some variation to ensure generalisability whilst keeping the number of vignettes to a manageable number. “

We have changed the title in accordance with the recommendation, with which we agree: it now stands as “Understanding negative feedback from South Asian patients: experimental vignette study”

Many thanks indeed for your further consideration of our paper.