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)	Protocol: Virtual online consultations - advantages and limitations (VOCAL) study
AUTHORS	Greenhalgh, Trisha; Vijayaraghavan, Shanti; Wherton, Joe; Shaw, Sara; Byrne, Emma; Campbell-Richards, Desiree; Bhattacharya, Satya; Hanson, Philippa; Ramoutar, Seendy; Gutteridge, Charles; Hodkinson, Isabel; Collard, Anna; Morris, Joanne

VERSION 1 - REVIEW

REVIEWER	Amie Goodin, PhD University of Kentucky, USA
REVIEW RETURNED	12-Aug-2015

GENERAL COMMENTS	- Please update the protocol to describe in greater detail how screening for inclusion in the study will be conducted (e.g., Will **all** patients visiting these two facilities be screened as potential subjects if they sign the consent form prior to contact by the research team? What questions are specifically being asked of subjects in order to determine eligibility by the research team and how is that information collected prior to the intervention?). *This is a key element of the study protocol and **must** be addressed.* - Related to the comment above- The sampling protocol describes a method of patient selection to purposefully "…seek maximum variety in social…and personal circumstances, and in health and IT literacy," yet it would be very difficult to identify appropriate candidates having low socioeconomic status and/or low IT literacy considering that the technology is not accessible to many patients satisfying those descriptions. With this in mind, can the research team establish a firmer set of definitions or selection criteria when determining which patients to include? - The "meso-level" seems more aptly described as another facet of the micro-level, because the data collection protocol as written appears to target patient-facing staff and administrators. Please clarify which staff and personnel with whom you will speak in the work environment (e.g., do you seek those who have direct involvement or only tangential involvement with the implementation or administration of the remote videoconferencing service?). It is unclear how the corporate partners mentioned in the meso-level sampling section fit into this picture. If these actors represent "institutional support" this would need to be gauged via

some other means unless the research team is prepared to ask the interviewed individuals to wholly represent their respective institutions.

- o The data analysis section for the meso-level portion of the study further demonstrates that the "institutional level" is ill defined because analysis is focused on patient-facing staff and administrators. Data collected from these actors, which is indeed very valuable for determining the impact of this service on workflow and ground-level administration, does not provide evidence of the institutional factors purported to be explored in the objectives and described in the sampling strategy. The proposed series of studies are already quite ambitious and it would not detract from the potential impact by simplifying the objectives of this one component.

- What videoconferencing software/application is being used for patient consultations? Skype is mentioned in the literature review but no software or application is explicitly stated in the methodology.
  - o Related to the above- How does the team assess technological compatibility once a patient is enrolled in the study (or is this part of the screening prior to enrollment)? If a qualifying patient has access to the appropriate hardware, but does not have a videoconferencing program installed, will the research team assist with installation and setup?

- The literature review contained numerous references to previous works that struggled with poor subject retention. Knowing that "no-shows" and dropouts are a risk inherent to the micro-level study design, will new patients be enrolled to fill the slots of those who do not follow through with participation (i.e., does the research team require that a minimum of 45 videoconferencing sessions be recorded before findings are reported)?

- For micro-level data recording- if a patient requests or agrees to participate in follow up visits via remote videoconferencing, will these visits also be recorded and used for study or is only the first encounter recorded? How many follow up visits and over what time period would be included in the body of collected data if that is the case?

The ambitious sets of analyses proposed in the VOCAL study have the potential to contribute a great deal to this field. I hope that you will choose to address the comments above, particularly the first comment, and wish you the best with study implementation.

REVIEWER	Changmi Jung
	Johns Hopkins University, U.S.A.
REVIEW RETURNED	22-Aug-2015

| GENERAL COMMENTS | The study deals with varying level of analysis from micro (agent study) to macro (national policy) with regard to video conference |

medical consultation. If completed, the study has a great potential to fill the gap in the existing literature where we find a lack of understanding of interaction between agents, organizational routine, and national-level policy perspective. The study is still in the premature status as there is no preliminary data analysis yet, and thus I expect that the study bears a fruit in the near future.

As the study is qualitative, it seems to me that the authors have taken proper approach to address the research question. However, there is a lack of descriptive statistics of the data (patient demographics, physician information, etc.) which will enrich the understanding of the readers about the study subjects.

There is some clarification issues with the data as well. For instance, authors mentioned that the patients' engagement increased (13 percent 'did not attend' rate for Skype patients compared to 28% 'did not attend' rate of the rest - page 9). What is the percentage of 'did not attend' pre/post for each patient population? The authors need to keep in mind that the patient populations might be systematically different. Skype patients are likely to be more engaged in keeping their own health by using online network, and might be more closely watching their condition to begin with. Thus, it is important to compare the difference between pre/post introduction of such service.

In addition, the paper can be richer by including survey targeting patient users and physician users. Survey study can provide more solid evidence of the users' perception. Also, we can utilize it to confirm the researchers' interpretation of the users' behavior on the video-taping.

It would be great if the authors add one more question in macro level data collection: which healthcare policy needs to be revised/updated in order to scale up the video conference consultation?

Meso-level analysis might be able to apply organizational learning theory. What the authors mentioned by "what gets done, by whom and how" is regarded as Transactive Memory System. The study can be extended to Org. learning study by observing the routinization of the new process.

Also how human resources are utilized in the new service, and how it affects the existing operation would be another subject that the authors might be interested in exploring.

The study is certainly interesting and the authors well addressed the questions and good rationale behind them. Currently there is no result to discuss, and thus I hope to see the development of the study in the future. Also, the study can be a great case article if written in more personal level.

VERSION 1 – AUTHOR RESPONSE

REVIEWER 1
Please update the protocol to describe in greater detail how screening for inclusion in the study will be conducted (e.g., Will all patients visiting these two facilities be screened as potential subjects if they sign the consent form prior to contact by the research team? What questions are specifically being asked of subjects in order to determine eligibility by the research team and how is that information

collected prior to the intervention?). This is a key element of the study protocol and must be addressed.

We accept the criticism that the inclusion/exclusion criteria were inadequately defined and have rectified this on page 15. However, it's important to point out that we are seeking to DEVELOP inclusion/exclusion criteria for Skype consultations through a qualitative and emergent approach – hence we do not want to impose a rigid set of criteria at the outset (indeed, as we now say in the paper, that would be clinically dangerous given that there is no good data on this topic in the literature). We need to allow clinicians to play (and justify) their hunches.

Related to the comment above- The sampling protocol describes a method of patient selection to purposefully "...seek maximum variety in social...and personal circumstances, and in health and IT literacy," yet it would be very difficult to identify appropriate candidates having low socioeconomic status and/or low IT literacy considering that the technology is not accessible to many patients satisfying those descriptions. With this in mind, can the research team establish a firmer set of definitions or selection criteria when determining which patients to include?

Actually, the setting where this study is based (Newham, London) challenges this statement, and we have published pilot data to substantiate this claim (see for example http://bjdvd.co.uk/index.php/bjdvd/article/view/84/181). Despite being one of the most deprived and multi-ethnic boroughs in the country, Newham has a higher level of broadband connection than the UK average. Furthermore, Black and Asian residents are MORE likely to have a broadband connection than those who self-classify as White (probably partly because many use Skype to keep in touch with family abroad).

As Dr Vijayaraghavan reflected in a recent BMJ blog (http://blogs.bmj.com/bmj/2015/02/03/shanti-vijayaraghavan-using-mobile-technology-to-empower-patients/): "....we then wanted to see if we could engage people who were traditionally labelled as 'harder to reach'—those who were most disengaged as a result of social isolation and who never turned up to appointments at all. This was part of a wider ambition at Newham to completely change the way we deliver care to ensure that we make it as easy as possible for people to engage with us. We made it clear to these patients that we would hold one scheduled consultation with them each year when we would be able to examine them, and that they could contact us directly in the interim if experiencing any problems managing their diabetes. Many of the patients traditionally labelled 'hard to reach' began getting in touch with us proactively. They welcomed the newfound access that the project and Skype gave them."

The above comment is a clinician's impression, not research data, but it conveys an important point – that our stereotypes about who will 'engage' with remote consulting may well be incorrect. Hence, we will offer the service to anyone for whom it is considered clinically appropriate and collect data on who actually takes it up. The 'maximum variety sample' described in the paper is a standard approach for qualitative studies: we take all suitable participants initially, then if we find that all recruits to date are Asian men (say), we stop recruiting Asian men for a while. The sample will not be statistically representative, nor does it have to be to answer our research questions.

Please note also that this study is predominantly qualitative and asks primarily "what is quality….?". Defining demographic sub-groups of 'suitable' patients with any level of statistical robustness would need a different design and is beyond the scope of this study (though also an important question!).

The "meso-level" seems more aptly described as another facet of the micro- level, because the data collection protocol as written appears to target patient- facing staff and administrators. Please clarify which staff and personnel with whom you will speak in the work environment (e.g., do you seek those who have direct involvement or only tangential involvement with the implementation or administration of the remote videoconferencing service?). It is unclear how the corporate partners mentioned in the meso-level sampling section fit into this picture. If these actors represent "institutional support" this would need to be gauged via some other means unless the research team is prepared to ask the interviewed individuals to wholly represent their respective institutions.

We respectfully wonder whether this reviewer is familiar with organisational ethnography, which is a standard qualitative technique used by organisation and management scholars to capture meso-level data. The dataset was described in our paper as follows: "Data collection will be predominantly

ethnographic, consisting of physically visiting the different departments (clinical, administrative, executive) and undertaking naturalistic interviews – that is, asking people on the job what they are doing and why they are doing it. [...] as well as collecting key documentation such as existing standard operating procedures and any informal guides and notes made by staff to help them do their job. The dataset for the meso analysis will thus consist of field notes (to be typed up and annotated as soon as practicable after the field visit), plus documents, charts and other artefacts supplied by staff." Elsewhere, we make clear that we will be including people from "estates, finance and clinical informatics" departments. Hence we are very explicit that we will be interviewing staff who are NOT patient-facing (as well as those who are), and using the powerful technique of ethnography along with document analysis to build up a picture of the organisation (and the different departments within the organisation). We have removed the phrase "corporate partners" which we agree was unclear. We ask both the reviewer and the editors to note our firm view that in sociological studies of organisational innovation and change, it is poor research practice to begin with rigid definitions and categories of who will be interviewed or observed. If needed, we can supply previous studies by ourselves and others that took an emergent approach to sampling.

The data analysis section for the meso-level portion of the study further demonstrates that the "institutional level" is ill defined because analysis is focused on patient-facing staff and administrators. Data collected from these actors, which is indeed very valuable for determining the impact of this service on workflow and ground-level administration, does not provide evidence of the institutional factors purported to be explored in the objectives and described in the sampling strategy. The proposed series of studies are already quite ambitious and it would not detract from the potential impact by simplifying the objectives of this one component.

The meso level objective was defined as follows: "At meso level, to illuminate and explore the socio-technical microsystem that supports the remote consultation, thereby identifying how organisations can best support the introduction and sustainability of this service model in areas where it proves acceptable and effective". The analysis section for the meso-level data states "...we will use both diagrams and narrative as synthesising devices to draw together a visual representation and a linked verbal account of the human and technical interactions and interdependencies on which the successful execution of the remote consultation depends. We will also draw on Feldman's notion of the organisational routine – defined as "a repetitive, recognizable pattern of interdependent actions, involving multiple actors" – whose potential and use in the healthcare setting we have previously described theoretically and applied empirically."

We have used precisely this approach in a number of previous studies to produce a meso-level anlaysis – see for example this BMJ paper on safety routines in repeat prescribing: http://www.bmj.com/content/343/bmj.d6788.full. Again, the reviewer has incorrectly assumed that we were planning to interview only patient-facing staff. Back offices have routines too, and we'll be studying those. See for example this paper, in which we focused exclusively on 'back office' routines: http://www.biomedcentral.com/1472-6963/15/177

What videoconferencing software/application is being used for patient consultations? Skype is mentioned in the literature review but no software or application is explicitly stated in the methodology.

This vagueness is deliberate. We're studying a service (remote consulting), not a technology. Our study design is technology-agnostic. We are currently using Skype but we do not want to tie ourselves to that software if a strategic-level decision is made in the trust to change to a different software or application (or, for example, if some or all participating departments decide to respond to patient demand for using Facetime). (We also don't want Skype using the VOCAL protocol paper as advertising material!)

Related to the above- How does the team assess technological compatibility once a patient is enrolled in the study (or is this part of the screening prior to enrollment)? If a qualifying patient has access to the appropriate hardware, but does not have a videoconferencing program installed, will the research team assist with installation and setup?

'Technological compatibility' is assessed in a similar way to assessing whether the person's telephone works – if they get through to the clinic, they're technically compatible. As with patients' telephones, the trust does not offer technical support or installation. What it does do is assure people that if they encounter technical problems they can be seen promptly and as needed in the face to face clinic. The beauty of Skype is that most people responding to the invitation to try remote consultations are already using it, often very skillfully, to link with family and friends. In a previous pilot, the diabetes team used Adobe conferencing software and patients had to download and install it, and because of that it generated a lot more problems than it solved.

The literature review contained numerous references to previous works that struggled with poor subject retention. Knowing that "no-shows" and dropouts are a risk inherent to the micro-level study design, will new patients be enrolled to fill the slots of those who do not follow through with participation (i.e., does the research team require that a minimum of 45 videoconferencing sessions be recorded before findings are reported)?

Yes. 45 is our target. We will probably have to recruit considerably more to capture that many in the final dataset.

For micro-level data recording- if a patient requests or agrees to participate in follow up visits via remote videoconferencing, will these visits also be recorded and used for study or is only the first encounter recorded? How many follow up visits and over what time period would be included in the body of collected data if that is the case?

This is a qualitative study of emergent change, hence we will take an emergent approach to this question. We would certainly want to capture some follow-up consultations, but our key research question is "what is quality….?" and in order to address that, we plan to see how the Skype experience unfolds for different groups of patients.

The ambitious sets of analyses proposed in the VOCAL study have the potential to contribute a great deal to this field. I hope that you will choose to address the comments above, particularly the first comment, and wish you the best with study implementation.

Many thanks for a most helpful review which has helped us improve the paper and also given us much to think about as we do the empirical work.

REVIEWER 2
The study deals with varying level of analysis from micro (agent study) to macro (national policy) with regard to video conference medical consultation. If completed, the study has a great potential to fill the gap in the existing literature where we find a lack of understanding of interaction between agents, organizational routine, and national-level policy perspective. The study is still in the premature status as there is no preliminary data analysis yet, and thus I expect that the study bears a fruit in the near future.

Thanks. This is exactly the challenge, to get this field to the starting blocks. We see it as (potentially) illuminative and hypothesis-generating rather than hypothesis-testing.

As the study is qualitative, it seems to me that the authors have taken proper approach to address the research question. However, there is a lack of descriptive statistics of the data (patient demographics, physician information, etc.) which will enrich the understanding of the readers about the study subjects.

We agree this was an omission, and have added some inclusion criteria for patients and a broad description of staff (page15). Because this is a naturalistic study (we're studying the real world, not doing an experiment), we will include whatever staff are allocated to the remote service. We don't need to predefine a specific group (as we would in an experiment).

There is some clarification issues with the data as well. For instance, authors mentioned that the patients' engagement increased (13 percent 'did not attend' rate for Skype patients compared to 28% 'did not attend' rate of the rest - page 9). What is the percentage of 'did not attend' pre/post for each patient population? The authors need to keep in mind that the patient populations might be systematically different. Skype patients are likely to be more engaged in keeping their own health by

using online network, and might be more closely watching their condition to begin with. Thus, it is important to compare the difference between pre/post introduction of such service.
Good point. We have commented on page 9 that the 'DNA' rates are against different denominators and cautioned against over-interpreting these data. But see above for a comment on demographics. Actually a lot of patients in the young adult diabetes clinic who successfully connect to the service using Skype are ones who have been previously hard-to-reach (e.g. have not attended a diabetes check-up in several years). To our surprise, Skype seems to fit a different demographic from the 'tech-savvy middle class patient'.
In addition, the paper can be richer by including survey targeting patient users and physician users. Survey study can provide more solid evidence of the users' perception. Also, we can utilize it to confirm the researchers' interpretation of the users' behavior on the video-taping.
We haven't been funded to do a survey in this study. In a previous set-up study funded by Health Foundation, we did surveys of patients and staff.
It would be great if the authors add one more question in macro level data collection: which healthcare policy needs to be revised/updated in order to scale up the video conference consultation?
Our macro-level research question is "What is the national-level context for the introduction of virtual consultations in NHS organisations, and what measures might incentivise and make these easier?" This gives us plenty of scope for addressing policy. The reviewer is very perceptive in that we too wanted to find out which policy was the one we should unpack and explore. Our preliminary data were surprising – whilst politicians and policymakers are all taking about Skype / remote consulting, it is only mentioned at the very vaguest level in the documents we've studied. But all that is 'work in progress' and for a later paper.
Meso-level analysis might be able to apply organizational learning theory. What the authors mentioned by "what gets done, by whom and how" is regarded as Transactive Memory System. The study can be extended to Org. learning study by observing the routinization of the new process.
Actually we're using a more sociological lens, with a slightly different interpretation of the word 'routine'. See the links in our response to reviewer 1 for some examples of where we've made this work well in the healthcare context.
Also how human resources are utilized in the new service, and how it affects the existing operation would be another subject that the authors might be interested in exploring.
Yes, this is definitely part of our meso level dataset.
The study is certainly interesting and the authors well addressed the questions and good rationale behind them. Currently there is no result to discuss, and thus I hope to see the development of the study in the future. Also, the study can be a great case article if written in more personal level.
Thank you.

VERSION 2 – REVIEW

REVIEWER	Amie Goodin, PhD University of Kentucky, USA
REVIEW RETURNED	27-Oct-2015

GENERAL COMMENTS	It was my pleasure to review the revised protocol for the VOCAL study. The revisions submitted for the sampling procedure in particular (pages 15 and 16) were both thoughtful and thorough, and have satisfied my primary concern from initial review. I do have three minor comments for other sections of the revised protocol. -The response document contains a reference to a recent study conducted by several members of the research team that challenges the notion that patients with low SES may be under-represented in the VOCAL study due to limited access to the required technology

(DAWN: Diabetes Appointments via Webcam in Newham). The reference was a fascinating read and I strongly suggest that it be added to the background section with the brief explanation as to why Newham and London may be "atypical" in terms of previous reports of technology access. This departure from previous reports of limited access in low SES populations has major (and exciting!) implications for future studies and trials using these technologies for health care delivery.

-The response document contains clarification regarding meso-level participant selection and some minor editing was conducted in the protocol proper (page 17). The elimination of the phrase "corporate partners" did help me better understand both target and intent. The response document contains clarification to my comments as well as comments from Reviewer #2 that the team is choosing to analyze the meso-level data from a sociological lens, and there is indeed a citation in the protocol that refers the reader to a previous discussion of the theory behind the meso-level design. I am assuming that the write up after data collection will include a more rigorous recounting of this meso-level theory and so my comment from first review is satisfactorily addressed.

-The response document mentions a deliberate vagueness about naming the technology platform patients choose to connect with the clinicians. You provide an excellent rationale for the choice to avoid brand names in your data collection section, while also clarifying in the response document that the study design is "technologically agnostic." I agree with your stated rationale about brand name avoidance and suggest that it would be beneficial for your team to explicitly state this research design choice in your micro-level data collection section (if excluded, the Skype-heavy literature review gives the impression that it is the only platform by which a patient could connect). Not only will this minor addition clarify the procedure, it will also reinforce how unique this study effort is in terms of exploring health care solutions driven by practical and available resources that truly meet patient needs.