Author’s response to reviews

Title: Between and Within-site Variation in Qualitative Implementation Research

Authors:

Justin K Benzer (justin.benzer@va.gov)
Irene E Cramer (lmchsc@gmail.com)
Sarah Beehler (sarah.beehler@va.gov)
David C Mohr (david.mohr2@va.gov)
Martin P Charns (martin.charns@va.gov)
James F Burgess (james.burgess@va.gov)

Version: 2 Date: 29 June 2012

Author’s response to reviews: see over
Thank you for the opportunity to revise and resubmit our manuscript. We appreciate the positive responses to our previous manuscript and have endeavored to clarify the issues raised in the review. We thank the reviewer team for providing useful and actionable feedback. We will first address the comments regarding implementation science, and then address the comments regarding qualitative methodology.

Implementation Science Comments

Editor's comment

An additional point to note is the need to locate the existing paper within the literature on qualitative research in complex interventions and implementation (see below for a couple of examples) - of which there is a substantial body of knowledge (both in the UK and US) and to demonstrate how the current paper builds on this literature.

Thank you for this suggestion. We have revised the introduction to position this paper within the multi-site intervention implementation literature. We identify research on organizational context in implementation as a cross-cutting issue in implementation science that is particularly important for multi-site studies.

In addition, we conducted an informal literature review to determine the methodologies typically used to analyze multi-site implementation studies in this journal. We reviewed all of the multisite implementation articles published as a “research article” in this journal over the past two years. We found that most of the recent multi-site studies published in implementation science use only quantitative methods. The three qualitative/mixed methods studies we found used limited sampling. One study only interviewed leaders (1), a second study only interviewed clinicians but found that leaders were an important implementation factor (2). A third study interviewed one leader and worker at a site, but did not report within-site implementation themes (3). Our study identifies both within-site and between-site themes that appear to impact the intervention implementation.

Reviewer 1.
The lack of theory about implementation, and about change, means that the study is not grounded within existing ideas or has a theoretical framework that can inform the study direction, analysis and conclusions.

We agree that grounding this study within an implementation conceptual framework is useful for interpreting the results. We have moved our methodological conceptual framework to the methods section (pages 6-8). We now ground the study within the literature on multi-site interventions. We identify research on organizational context as a cross-cutting concept that is relevant for all multi-site implementation studies. The specific intervention in our study is an example of a locally adapted intervention in response to a mandated organizational change, but organizational context is broadly important to intervention implementation. The purpose of this paper is to demonstrate that analyzing variation in perceptions of the organizational context from different types of informants facilitates the identification of both within-site and between-site implementation themes.
Reviewer 2
Clearly in your study, ‘employee position’ was the key issue that enabled you to explore variation. However, in other sites, other issues might be central – say, closeness to day-to-day implementation of the intervention, or level of disruption to prior work practices … etc. Such factors may be relevant within practitioners of the same ‘employee position’.

We agree that within-position variation is potentially important and creates an additional level of complexity that is typically not addressed in multi-site qualitative implementation research. We have expanded our discussion of this limitation on pages 31-32.

Qualitative Methods Comments
Reviewer 1
1. The complex description of the coding process doesn't actually assist the reader to know how the data were analysed - this could be reduced to a couple of sentences. While coding is important, it is what researchers do with the codes in the analytic process that is of more interest, and this isn't well explained.

Thank you for this feedback, we agree that the analyses are more important than the initial coding procedure. On page 9 we now present one short paragraph describing the initial coding process and one short paragraph orienting the reader to why the data were re-analyzed. On pages 9-11 we have expanded our description of the analyses for the current paper, and note that we guide the reader through our analysis of within-site and between-site variation in the results section.

2. ANY study of merit seeking stakeholder views on implementation of an intervention needs to account for all the viewpoints, so in that sense, this study doesn't offer anything new.

We agree that the novel contribution of this study was not clear enough in our initial draft. We now emphasize that our contribution is explicating the analysis of multi-site implementation studies. Implementation research has to manage budget constraints, and this is particularly a concern for multi-site implementation studies. Some implementation studies limit their data collection to leaders, for example, to limit the cost of qualitative data collection. The need to account for both within and between site variation using a low cost design is somewhat unique to implementation science and these differences are not well discussed in the broader qualitative health literature. To our knowledge there has been very little implementation science research that provides guidance as to which informants to sample and the consequences of incomplete sampling in multi-site implementation studies. Further, given that interventions may be implemented by researchers who do not have formal qualitative training (e.g., physicians) we believe that an example of this type of analysis is an appropriate contribution to the implementation science literature.

3. The references to 'contextual validation' also are unclear.
We have clarified this concept and generally expanded our discussion of triangulation on pages 5-6. As part of our general goal to minimize complexity, we decided to drop the term contextual validation. We were referring to Lincoln and Guba’s extension of the contextual validation concept that was coined by Denzin. Lincoln and Guba (p. 301-307) suggested that credibility can be established for both convergent and divergent data through triangulation. In contrast to the more common use of convergent triangulation to identify common perceptions across informants, patterns of responses can also be identified to establish credibility. We use this concept of “patterns of responses” to establish the credibility of divergent data.

4. Carefully consider the conceptualisation of the data - what is presented are not implementation 'narratives' in the sense that most qualitative researchers would recognise.

To clarify, we had previously labeled a common perception within informant type as a narrative (i.e., primary care narrative). We agree that this terminology was confusing and as part of our effort to simplify the methodology and use clear terminology we have removed the word “narrative” from our paper. We now simply highlight the differences and similarities across informant types in the results section.

5. The table with the presentation of data is confusing, for example I wonder whether 'psychologists are co-located in primary care' (Site Alpha) is a theme - is it more a finding of fact that participants agree on?

We have revised our results section to better support the tables. The tables were not intended to present themes. The purpose of the tables was to highlight what was agreed upon by all informants and the source of the variations. On page 11, we now describe more fully how the results elaborate upon the information provided in the tables.

6. Much of what you have labelled as 'divergent' doesn't read as such, in some cases what you have presented is a slight variation only.

Thank you for pointing out this lack of clarity. A large part of the confusion was due to our distinction between convergent, complementary, and dissonant. Our framework now distinguishes between redundant, convergent, and divergent data. The key difference between convergent and divergent data, as we now state on page 10, is that divergent data “suggests that a phenomenon is interpreted differently across sources.”

In some cases we agree that the distinction between convergence and divergence may be idiosyncratic. However, determining which data are convergent and which data are divergent is not the ultimate purpose of the analyses. The purpose of our analyses is to evaluate the credibility of variation among informant types. If the data indicate convergence between informant types than that minimal variation can be judges as credible. If data indicate some levels of divergence across informant types and that divergence is supported either by “the pattern of responses within sites (i.e., does the divergence fit with the responses of other informants)” or by the “patterns across sites (e.g., does the divergence fit the types of responses from similar informants across sites)” then that variation can be judged to be more credible. We
have eliminated the terms convergence and divergence from the title of our manuscript to further clarify that we are interested in the analysis of variation in multi-site data, rather than convergence and divergence as concepts.

We adopted a relatively strict definition of convergence, and therefore we have presented wider ranges of results that we chose to label divergent. We made this decision for the purpose of demonstrating the utility of variation among the sources. Our analyses highlight what is truly agreed upon in this study (i.e., trivialities) in order to emphasize the value of variation. That is, not only is variation important for a richer description of phenomena, but it is vital to identify meaningful themes rather than generating a trivial taxonomy of concepts. However, we intended to communicate that this was an intentional choice by the research team not a standard qualitative methodology, and this was not communicated well. We have made several revisions to highlight this analysis choice.

First, we have moved most of the discussion of convergence and divergence from the background section to the methods. In this section (page 7-8) we specify that we have adopted a strict criterion for determining what is shared across informants for demonstration purposes. We also note that this is not standard qualitative methodology. Second, in the data analysis section (page 11), we define our criterion for convergence as “redundant data where one source is completely interchangeable with another” and specify that complementary data elaborate on the shared concept. Third, in the results section (page 11), we specify that the convergence is “the descriptive data that was supported by data from all three informants.”

7. Consider how the data can be presented thematically - the separation into sites does not work and is repetitive and descriptive. You may get different insights into the different views expressed if you read inductively from the data instead of trying to fit the data into the categories you have identified.

Thank you very much for this important suggestion that drove our choices in revision. In our prior presentation, we had organized the data by type of key informant within each site. The intention was to demonstrate how conclusions regarding the site would differ if only one informant was selected. We see now that this was not an effective presentation. We have re-organized our results to parallel the tables. The tables present the convergent and divergent data organized by the redundant concepts reported by all three informants. Our results now begin by introducing the redundant concepts and explaining how the convergent and divergent data elaborates the concepts. We do still stratify our analyses by site. This presentation allows us to separate within-site themes from between-site themes. We have greatly expanded our analysis such that analyses now focus on the differences between within-site and between-site themes. We now conclude each section/site with an analysis for a dominant within-site theme and conclude the results section with an analysis of the between-site themes.

Reviewer 2
1. Page 8, 20 & 27: You note that ‘theoretical sampling’ was used. However, I’m unsure how what you describe – interviewing clinic leaders who then forwarded you to staff represent any version of theoretical sampling (either a priori or emergent – as in, the style
of grounded theory). This appears to be, following Patton’s typology, an example of criterion sampling (criteria being clinic lead) followed by either critical case or intensity sampling. Or, you could, as you later refer to it, describe it as ‘stratified’ purposeful sampling.

We agree that our use of theoretical sampling is not very precise. Thank you for this suggestion to use Patton’s typology. We agree that stratified purposeful sampling accurately reflects our methodology.

2. Page 10: I’m not sure what you mean by ‘redundant data’ when you say ‘Convergent narratives were supported by redundant data from all three informant types.’ We intended to communicate that we judged data to be convergent if identical conclusions could be reached from all three sources.

We have clarified this in the data analysis section on page 11.

“Data were judged as redundant if data from each of MH, PCMH, and PC each supported the common concept (i.e., one source is completely interchangeable with another). Concepts were judged as convergent if data from one source elaborated upon the redundant data. Concepts were judged as divergent if data from one source conflicted with the convergent concept.”

This is somewhat repetitive as we describe our rationale more completely in our presentation of convergent and divergent data on page 9, but we felt that this point was important enough to repeat.

“We define redundant data as information obtained across informants that was completely parallel, indicating perceptions of an intervention that were completely shared across informants. We note that complete agreement is a strict criterion adopted for the purpose of demonstrating what was truly shared across informants. We define convergent data as information that elaborated redundant data by demonstrating variation in interpretation of shared phenomena. Variation among convergent data provides a deeper and richer understanding of a concept. Finally, we define divergence as unique data that reflects a departure from the convergent concepts. Divergence suggests that a phenomenon is interpreted differently across sources.”

References


3. Driessen MT, Groenewoud K, Proper KI, Anema JR, Bongers PM, van der Beek AJ. What are possible barriers and facilitators to implementation of a Participatory Ergonomics programme? Implement Sci. 2010;5:64.