Author's response to reviews

Title: More than a checklist: A realist evaluation of supervision of mid-level health workers in rural Guatemala

Authors:

Alison R Hernandez (alison.hernandez@epiph.umu.se)
Anna-Karin Hurtig (anna-karin.hurtig@epiph.umu.se)
Kjerstin Dahlblom (kjerstin.dahlblom@epiph.umu.se)
Miguel San Sebastian (miguel.sansebastian@epiph.umu.se)

Version: 6 Date: 16 January 2014

Author's response to reviews: see over
Dear BMC Health Services Research editors and reviewers,

We greatly appreciate the time you have taken to review our manuscript and provide constructive feedback to help us clarify the situation of actors in the study setting, the formulation of the theoretical proposition, the data collection and analysis process, the presentation of results, and their interpretation. Below we detail the reviewers’ comments in each of these areas and the changes made in light of their comments.

The situation of actors in this setting
2. The roles of the Auxiliary nurses (AN) are never described. Knowing more about their roles, tasks, and responsibilities would help in understanding what the manager is supervising. I feel “in the dark” about what type of care occurs in this settings and who is doing what for patients. (Reviewer 2)

10. Again, I feel as though I am trying to interpret what I’m reading in a vacuum because the roles and responsibilities of the ANs are not described. For example, what types of reports are they writing? What do the terms “institutional requirements” and “coverage levels” refer to (page 10)? Do these have any relationship to patient care? (Reviewer 2)

Further description of the ANs’ roles and “who is doing what” among the other staff at the health post has been added in the 3rd and 4th paragraphs of the sub-section Study setting. Their role is very similar to other mid-level health workers working at the primary care level in other LMIC health systems. They are in the health posts to attend the patients that come to them, but they also have to ensure that mothers and children are receiving preventive care services – for example, immunization, nutritional supplements, prenatal care, post-partum care, family planning – which require them to actively encourage service utilization and go out into the community to identify and follow up on people in need of these services. Institutional requirements dictate that the ANs meet the coverage goals for particular services in the HP’s catchment area, for example, 95% of children under one have a complete vaccination scheme. This is calculated to a certain percentage of children under 1 that must be vaccinated each month for them to be able to meet the coverage goal of 95% for the year. So in effect, these coverage goals quantify the services delivered and do not provide insight into patient care. This orientation is reflective of the historical trend in the health systems of LMICs (that typically depend to some extent on donor-financing) towards focusing on selective primary health care instead of comprehensive primary health care in light of the difficulty of financing and evaluating implementation of the latter. While perhaps it would be possible to include even more extensive detail in the text of the paper regarding ANs’ tasks in primary health care service delivery, we feel that this information provides an adequate base for understanding what the manager is supervising.
Health posts cover a catchment area of around 2000 inhabitants and are typically staffed by two auxiliary nurses (ANs). Auxiliary nurses are mid-level health workers with one year of training provided through ministry-accredited institutions. A lower entry-level education requirement makes the training more accessible than medicine or professional nursing for indigenous rural residents, thus many ANs in primary care services in Alta Verapaz share the culture and language of the population served. In the health posts, ANs provide basic maternal and child health services, attend consultations, detect and provide treatment for diseases and injuries of low complexity, provide health education in the community, and maintain census and epidemiological data for their catchment area. Their work with maternal and child health services, and specific diseases, such as tuberculosis, is structured by ministry programs which specify guidelines for service delivery and documentation, and indicators for coverage goals of priority services.

Other actors involved in supporting the ANs’ work at the health post level include local leaders, community health volunteers, and additional staff. In order to gain support for their activities and for implementation of health promotion strategies, the ANs’ work depends on communication with local leaders, as well as training of and coordination with health volunteers who have roles specified by the health system. In some health posts, there is also additional staff, including an educator, who helps with health promotion and/or a rotating medical student who attends patients visiting the health post, usually with the AN serving as interpreter.

3. How are nurse supervisors trained in management? It would be helpful to know if they get trained in specific competencies consistently. Also, who do they report to and how is their performance evaluated? Are there incentives or disincentives that influence their managerial style? (Reviewer 1)

We agree that this information would be a very interesting addition to the paper. Unfortunately, our knowledge of this is only anecdotal and we have not engaged in a review of professional nurses’ training curriculum, nor have we had a systematic focus on supervisors’ relationship to higher levels of management. Based on personal conversation with nursing school faculty, we know that management is covered in the curriculum and that in the National Nursing School in Alta Verapaz where these supervisors were trained, they include focus on supportive supervision and introduce the concept of accompaniment. But it was not possible to review the curriculum to be able to state with certainty what managerial competencies they receive training in. Regarding the relationship to higher management, our impression was that there is not much attention to evaluation of the supervisors’ performance as a supervisor – instead they work as part of the management team along with the district nurse and the district director, and together they are accountable to the regional level for the performance of the district health services. We’re not sure if their managerial style would be known or recognized by other members of the management team, but there are no incentives or disincentives attached to it. We have added the following text to clarify the organization of management at the district level in the 4th paragraph of Study Setting:
The health post is connected to the district level through the institutional intervention of supervision. Supervision of ANs is conducted by professional nurses with three-year technical degrees or five-year bachelor degrees, and their training includes some focus on their role in managing care by ANs. At the time of this study (2010), district management included a dedicated primary care supervisor who was in charge of the health posts, in addition to the district nurse manager, and district director.

In the last paragraph of Methodological Considerations, we have added text to highlight this as an area in need of further study:

These findings can serve as a starting point for further exploration of supervisors’ interpretation of their role and how their managerial style is generated in order to identify aspects that can be fortified to promote implementation of more supportive supervision in this context, or similar contexts.

**Formulation of the theoretical proposition**

3. The middle range theory comes out of nowhere in the sense that the prior study is not described or cited. This is a major weakness. (Reviewer 2)

We have aimed to clarify the formulation of the theoretical proposition that provided the base for the study by using the term “program theory” and defining it more clearly. We felt the observation that our revised MRT might be better stated as “hypothesis” (comment #16, reviewer 2) reflected that we should re-evaluate our use of the term “middle-range theory”. After returning to the literature on realist evaluation, we found that the term “program theory” more accurately describes the level of abstraction of our initial theoretical proposition. The observations employed in the initial development of this program theory were made during field work for the first study of this dissertation project, which included interviews with supervisors and ANs. The paper where this study was reported focuses on the ANs’ practice and does not include findings from the supervisor interviews, but we have now included it as a reference to indicate which “previous study” we are referring to. The second paragraph of the sub-section Study design has been modified as follows, and the term “program theory” has replaced “MRT” throughout the paper:

The first step of the process was to formulate a program theory articulating the mechanisms of how the activities of supervision contribute to supporting AN performance based on the models that local actors use implicitly to make sense of their actions [20]. Realist evaluation typically starts from a preliminary middle range theory which draws on existing knowledge and theory to articulate the links between outcomes, mechanisms and context elements. However, program theory can serve as an alternative starting point when little is published or known about the intervention under study, and in this case, published knowledge of how supervision of rural peripheral health units in LMIC health systems operates at the inter-personal level is limited [21]. A program theory articulates the set of assumptions, often held implicitly by the designers and implementers of an intervention, that explain the choice and design of the activities of an
intervention and how they are intended to contribute to the desired outcomes [22]. The content of the program theory was developed through the first author’s observations of the organization of supervision of ANs in this setting from previous field work, and was refined through discussion with Guatemalan nurses working in the setting [23].

Data collection process
4. The selection of sites is clear but what criteria were used to determine the n of 5 cases? Was this sufficient to understand the phenomenon. For example, how many cases would need to be collected to see another case such as HP1? (Reviewer 2)

Determination of number of cases needed was based on the assumption that supervision would be relatively similar across health posts, due to the similarity of the overall institutional context. The initial theoretical proposition was fairly straightforward and did not include important rival explanations, though we did collect data to assess the influence of the context of the HP setting. Based on this scenario, we anticipated that five cases would provide sufficient information for assessing the original program theory. In the process of conducting the site visits, HP 1 stood out as a unique case, but the role of supervision as it related to the original theory was not understood until we began more in-depth analysis of the data. Based on the results of the analysis, we agree that 5 cases was not sufficient to fully understand the variation in the phenomenon that this case presented, particularly in terms of its potential relevance to supervision in other health posts. However, while the nature of the HP 1 supervisor’s approach was unique among the cases, it reflected themes that are prevalent among nurses in this setting even if they reach practice in different ways. In light of the reviewer’s comment, we have added the following text in the 5th paragraph of the sub-section Study design:

Replication of case studies provides the opportunity to test and refine the original theoretical proposition [19]. We anticipated that five health post cases selected to represent different regions of Alta Verapaz and different levels of performance would be sufficient to gain insight into the adequacy of the program theory for understanding supervision’s contribution to AN’s motivation and ability to perform in this setting. We selected three health posts with high scores (HP 1 – 3) and two with low scores (HP 4 – 5) from districts in the northeast, southwest and center of the health region as cases. The performance scores used to select these five health posts as well as their characteristics are presented in Table 1.

Clarification of the selection and recruitment of interview participants
5. The process and rationale for selection of individual participants and how they are recruited to participate are not described in sufficient detail. In each setting was there only one supervisor or many. If many how was the one selected?
6. A particular concern is that the ANs might have felt that participation was not voluntary because it appears they were recruited by the supervisor. That could cause them to bias their responses to interview questions. (Reviewer 2)
1c. Why was only 1 AN interviewed per site (most had 2)? (Reviewer 1)

The concerns related to recruitment and the reason for only interviewing one AN were addressed by adding the following text to the second paragraph of the sub-section on Data Collection:

_The visits were arranged by first contacting the nurse supervisors to inform them about the study, arrange a date to interview them and obtain the telephone numbers of the ANs. The first author then contacted the ANs to explain the nature of the study and set up a date for the visit to the health post...Only one of the two ANs on staff was available for interview per site due to the fact that they had begun scheduled vacation for the holidays._

Clarification about the number of supervisors was added in the fourth paragraph of the sub-section on Study Setting:

_At the time of this study (2010), district management included a dedicated primary care supervisor who was in charge of the health posts, in addition to the district nurse manager and district director._

7. What was asked in the interview? What type of interview approach? Open ended? What were the questions? What was the nature of the “relevant activities” in the documents collected? (Reviewer 2)

Information about the content of the interviews was moved from the paragraph on the development of the protocol to the third paragraph of the sub-section Data Collection and further elaboration was added in response to these questions:

_The interviews lasted from 20 minutes to one hour, and included open-ended questions to the supervisors and ANs about what supervision consisted in, how activities were perceived to function, how the activities influenced the ANs, as well as questions to ANs about their motivation and their perceptions of their work. The interviews with the community members explored the ANs’ relationship with the community and the nature of community involvement in supporting the health post._

Clarification of the nature of relevant activities was provided in the fourth paragraph on Data Collection:

_Notes from the review of documents in the health post reflecting contact with the community through health promotion activities, home visits, training for volunteers, meetings with community leaders as well as issues raised during supervisor visits and notifications sent by the ministry during 2009-2010 and other observations were recorded in field notes._
Data analysis process
8. It would be useful to have the definitions of the concepts used for coding.
9. What coding approach was used? Who did the coding? Was there more than one coder? What criteria for rigor were used? At what point did data analysis begin? Were all data collected or were data collected and analyzed simultaneously? (Reviewer 2)

**********************************************************************

Perhaps the use of the term “code” was misleading. Large segments of the transcript texts were labelled by themes from the program theory, like “monitoring”, and by categories of contextual information, like “AN characteristics”, “physical conditions”. There was little reduction of the data at this stage of the analysis. Instead, the case reports included many relevant quotes interspersed with descriptive and interpretive text. In this way, the research team was able to get involved in the analysis at this stage and still have contact with the voices of the participants in making their assessments of the initial analysis. We have added greater detail in the sub-section Data Analysis on the analytical approach and the research team’s interaction around the interpretation of the data in order to give better insight into the processes contributing to the credibility and confirmability of the findings.

The interviews were transcribed verbatim and the first stage of analysis focused on writing case reports for each of the health posts, following the strategies proposed by Yin of “relying on theoretical propositions” and “developing a case description”(ref). An index of themes was developed to organize the data from different sources related to the operation of supervision activities described in the initial program theory, as well as characteristics of relevant actors and the context of the health post. The themes were applied to the interview transcripts using NVivo 8.0 software, applied to field notes by hand, and analyzed on a case-by-case basis. Case reports were prepared by the first author which provided a thorough orientation to the health post context, including the physical conditions, the background and characteristics of the ANs and the supervisor, community involvement with the health post, as well as impressions of performance based in activities reflected in the register and observations. Reports also provided detailed description of the activities and outcomes of supervision, including extensive quotes from supervisors and ANs, and initial reflection on mechanisms and the appropriateness of the program theory. The research team reviewed each case report as it was produced. Inclusion of extensive quotes and description provided a basis for discussion around the interpretation of the nature of supervision as well as the influence of contextual factors on performance in each case.

The second stage of analysis focused on examining patterns across cases. Based on iterative review of the reports, the research team identified common and unique characteristics of supervision, and points of interest for further analysis. To facilitate comparison, text from the case reports was sorted into the three categories of supervision activities (monitoring, individualized support and accompaniment). Analysis focused on capturing patterns of similarity and variation in the implementation of supervision activities and their outcomes, uncovering indications of the mechanisms underlying the activities that explain how they contributed to the different outcomes observed, and understanding the influence of context on both the process of supervision and the outcomes. The account of the operation of supervision
articulated in the program theory was then reexamined and the theory was revised in light of the empirical findings.

In-depth analysis of the data collected did not commence until data collection was complete. However, initial impressions of the cases were recorded in field notes after each site visit. We have now stated this at the end of the sub-section Data Collection.

Presentation of results

Clarification of other factors influencing AN performance
11. Some description of the participants such as ages and years of experience would be informative. Differences in education etc. Did any of personal characteristics play a role in the behaviors noted?
19. The discussion contained what looked to be findings—the material on context should be presented in the findings section and integrated where needed to interpret the results. Seeing it in the discussion makes it seem to be an afterthought.
(Reviewer 2)
*********************************************************************************************
We have added more information introducing the context of the health posts in the 1st paragraph of the Results section.

The characteristics of the health posts are presented in Table 1. All health posts were located within one hour travel from the local health center, though availability of transportation varied. Staffing included one to two ANs and among those interviewed, three were female and two were male and years of experience ranged from one to six. The dispersion of the population covered varied among health posts with both high and low performance scores. Though most of the communities covered were within 30 minutes walking, HP 2 and 5 included villages that were more than one hour’s walking distance. Regarding physical conditions, the services available (water, electricity, cellular signal) were also variable. Only HP 2 and 3 reported current structural problems (roof and wall damage), while ANs at the other health posts recounted successful efforts to mobilize community support to improve physical conditions.

Presentation of monitoring and individualized support
12. You have not described the “mechanisms” of monitoring and individualized support. Instead, the results describe “what” the managers do—not “how.” For example, how did the nurse supervisors at sites 1-3 accomplish knowing the needs of their staff? Knowing that would help inform the supervisors at sites 4 and 5.
13. Overall, the findings are over synthesized, providing little description or actual data. Thus, the voices of participants do not come through as strongly as they could. Revising to include more data and direct quotes would help to bring these participants into view and would make the paper more compelling. On page 12, the authors begin to give more data to support their interpretations of what is occurring in these sites. This is helpful and the paper is more interesting here. Also, throughout the section for “accompaniment” the mechanisms or behaviors used to accomplish supervision are clearer and the data are more alive. More data is needed in the prior two sections along with more description.
14. At the bottom of page 12, there is an abrupt change in how data are presented with the shift from findings synthesized across sites to just site 1. Need a transition sentence or two.
15. I believe the paper would read better and would have more impact if the section on mechanisms was integrated into the sections on three supervisory approaches.
(Reviewer 2)

Reflection on these comments led us to significantly revise the presentation of the results. We have followed the suggestion in comment #15 to integrate the presentation of the mechanisms into the description of the activities and their outcomes. In this way, we have tried to provide a more dynamic description of how supervision was implemented that captures the experiences of the participants. In the previous version, we had stated that accompaniment in humanized support was infused in the other activities. With the new organization, the depiction of the contrast of the humanized approach is also more “infused”, and thus avoids the abrupt transition. Please see pp. 11-18 of the revised manuscript to review these changes.

16. It is unclear to me how the revised middle range theory is formed given the main finding is from only one HP and this is a primary feature of the theory (e.g., being attentive to the AN’s assessment of own needs....). I would suggest using these data to make some hypotheses rather than try to state as a middle-range theory. The mid range theory does not seem fully supported by the data but hypotheses would be appropriate. (Reviewer 2)

We have aimed to modify our original “program theory” in light of the findings of the current study, and it is intended to highlight that our main finding was identifying the role of the mechanism of supervisor’s orientation in shaping how the activities were implemented and received. Through observing the pattern of difference in the implementation of supervision in HP1 compared to 2-5, we realized that we had assumed that the supervisors (and ANs) would have a shared orientation and priorities about what constitutes performance but this was not the case. We have tried to address the reviewer’s concerns through some modifications in the language of the sub-section Revisiting the program theory, and integrating reference to Figures 1 and 2 which illustrate the influence of the different mechanisms on supervision (see pp. 18-19).

1. Title: The title “more than a checklist” is misleading. The study did not have anything to do with checklists and I suggest revision. (Reviewer 2)

In the sub-section Revisiting the program theory we also highlight the rationale behind the title of the article (which is reiterated in the Conclusions). The title lifts up the uniqueness of this article’s focus on the mechanisms behind the activities of supportive supervision, while other studies of supportive supervision cited in the introduction evaluate implementation by determining whether or not the designated activities have been completed.
Interpretation of results (Discussion)

Effectiveness of humanized support and other influences on performance

My main concern with the manuscript is that the authors seem to strongly favor the humanized support as more effective, but study does not seem to be powered to answer the specific question of which managerial style is more effective (or to correlate managerial style with outcomes). The systematic evaluation and classification of managerial styles in this manuscript provides important insights that could be used to evaluate programs in other resource-limited settings with similar task shifting to mid-level providers. I believe that with some changes, the manuscript can be reframed to focus on the findings while tampering the authors’ enthusiasm for the humanized approach, which I also share, but is not strongly supported by the data presented, except for the comments of one AN. Perhaps the authors can mention how their findings could lay a foundation for a larger outcomes study. (Reviewer 1)

18. The authors might note that only one of the five supervisors used the humanized support approach. How might this temper the conclusions drawn? (Reviewer 2)

It was not our intention to give the impression that this study could provide the basis for drawing conclusions about the effectiveness of the “humanized support” approach. The case study design and realist evaluation were used to focus our analysis on the processes that connect activities to outcomes, which are typically invisible in studies designed to evaluate effectiveness. This design was chosen in order to deepen understanding of the social nature of the intervention of supervision, and how its implementation is influenced by context and the interpretations of the supervisor and health worker. This justification was presented in the 1st paragraph of the sub-section of Study Design in Methodology. We have also tried to clarify the contribution of this study to understanding of rural health care supervision and provide better explanation of how its results should be interpreted in the 6th paragraph of the Discussion. The following text in the last paragraph of the sub-section Methodological Considerations was added specifically to provide a more nuanced view of how understanding one case of “humanized support” can provide a base for further study that contributes to strengthening supportive supervision through connecting its principles to local frames of understanding:

Based on one case, this study cannot offer conclusions regarding the effectiveness of supervision through the humanized support approach. However, the findings from this case provided evidence of a locally-generated orientation that deviates from the dominant institutional norms and reflects principles of supportive supervision. Findings across cases also indicated that the idea of supporting ANs through “accompaniment” is common among supervisors in this setting. These findings can serve as a starting point for further exploration of supervisors’ interpretation of their role and how their managerial style is generated in order to identify aspects that can be fortified to promote implementation of more supportive supervision in this context, or similar contexts.
Multi-factorial causality

1. Table 1/methodology
   a. 3 high performing sites and 2 lower performing sites were selected for the study. Only HP1 (high performing) had a “humanized” managerial approach. What do the authors attribute the success of HP2 and 3? The more hierarchical managerial style seems to be working there. Was there any clear distinction in management style between the 3 high performing posts compared to HP4 and HP5?

   b. HP5 seems to be an outlier with worse physical conditions, lower staffing (only HP with 1 AN with less experience) and a fairly disperse population covered. How do these factors influence its performance (compared to managerial style)?

2. ...There should also be a comment on how the general competence of the ANs could impact the performance of each site regardless of managerial style. The authors do acknowledge that other contextual factors were not fully evaluated (though some are presented in Table 1) (Reviewer 1)

The selection of cases based on performance scores was intended to provide the opportunity to observe the operation of supervision in a variety of settings, and not to explicitly connect supervision activities to an explanation of the scores. However, we can understand how the reader is attuned to the distinction between the groups of good and poor performance and may be expecting some explanation for the difference to arise from the analysis of cases. As the reviewer brings up, competence, physical conditions along with many other intersecting factors may be contributing to variation in these scores as well. In the 6th paragraph of the Discussion, we have acknowledged the difficulty of understanding supervision’s contribution to performance outcomes without a more complete picture of other factors that shape performance:

   However, supervision is one of many factors influencing the motivation and performance of health workers and differences in supervision alone did not account for variation in health posts’ measured performance levels. Attention to other factors shaping performance, such as health worker competence and the community’s care-seeking behavior, would provide further insight into how performance levels were reached. In order to better understand supervision’s contribution, there is a need to follow the process further to understand how the impact of supervisors’ support on AN motivation translates to performance outcomes, and how supervision interacts with other interventions to support health workers, including community participation.

Limitations

2. The authors should acknowledge the study’s limitations based on the small size of the study/personnel evaluated. (Reviewer 1)

As mentioned in response to Reviewer 2’s question regarding how the number of cases was decided (see Data Collection Process p. 4 of this letter), the n of 5 cases was appropriate within
the logic of case study design. However, our understanding of the phenomenon in these cases could have been improved through more extensive contact. This has been acknowledged in the following text in the 3rd paragraph of Methodological Considerations:

Our understanding of the inter-relationship between supervision and motivation could have been strengthened through repeated contacts, interviews with other personnel in the health posts and district management, and extended observation of practice, if time and resources allowed.

-------------------------------

We thank you again for your efforts to help us to improve the presentation of our findings that we believe can contribute to identifying context-appropriate paths to strengthening supportive supervision of health workers serving vulnerable rural populations.

Sincerely,
Alison Hernandez, Anna-Karin Hurtig, Kjerstin Dahlblom and Miguel San Sebastian