

## 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

<b>TITLE (PROVISIONAL)</b>	Let's Talk Numbers – A Qualitative Study of Community-dwelling US Adults to Understand the Role of Numeracy in the Management of Heart Failure
<b>AUTHORS</b>	Sterling, Madeline; Silva, Ariel; Robbins, Laura; Dargar, Savira; Schapira, Marilyn; Safford, Monika

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Candace McNaughton Vanderbilt University Medical Center, Nashville, Tennessee, USA
<b>REVIEW RETURNED</b>	16-Apr-2018

<b>GENERAL COMMENTS</b>	<p>The authors address an important topic that has implications for the approach to treating and educating patients with heart failure. Their framework may fill a gap in our understanding of the patient-provider relationship for patients with HF, although it will need to be tested in future work.</p> <p>The authors conducted a thoughtful qualitative study, paying attention to important details and following accepted processes. My questions and suggestions are minor and easily addressable.</p> <p>Abstract: Although it may be true, the last sentence of the abstract should be reworded to more accurately reflect the findings of this study. The study was not designed to compare the relative importance of individual patient numeracy to the role of other factors, please rephrase, e.g., something along the lines of “These findings suggest that HF-specific training of caregivers may play an important role in successful HF self-care. Future work addressing HF self-care should include HF-specific training of caregivers, and more work is needed to better understand the intricacies of the relationships between HF patients and their caregivers.”</p> <p>“ER” and “emergency room” should be changed to “emergency department” on pages 9 and 17, since this is the preferred term.</p> <p>Despite the focused on caregivers, interviews were conducted among patients and some conclusions related to caregivers extrapolate from patient report. The authors describe a gap between self-reported numeracy and use/understanding of numbers, which raises questions about the accuracy of patient comments regarding caregivers. The conclusions centered on caregivers appear to overstate the data available.</p> <p>Page 18 line 24, please change “is often” to “may be variable”, given the sampling approach.</p>
-------------------------	---

	Framework: Is the relationship between knowledge and communication uni-directional? Numeracy likely plays a role in the quality of communication. Many of the factors listed as impacting knowledge would also seem to impact patient-provider communication. Have you all discussed these questions/issues in development of the framework?
--	--

<b>REVIEWER</b>	Rocco Palumbo University of Salerno, Italy
<b>REVIEW RETURNED</b>	04-May-2018

<b>GENERAL COMMENTS</b>	<p>Thank you very much for the opportunity to read your interesting work. This paper tackles an important issue and adds some intriguing insights to push forward our understanding about the appropriate management of HF. However, after carefully reading your article, some concerns appeared in my mind. I will summarize my suggestions to improve the quality of your paper in the following lines:</p> <p>1) I found that the introduction was not able to "set the hook" and to catch the readers' attention. On the one hand, you emphasize what is the main topic of your research; on the other hand, you do little to stress the originality of your contribution and the relevance of your study. Therefore, I suggest to improve the introductory section: in particular, you should better state the specific purpose of your research and point out how it adds something new to the scientific knowledge.</p> <p>2) You argue that a "one to one semi-structured" approach was used to interview patients. However, the development of the manuscript suggests that a de-structured approach was used. Maybe, it would be worth providing more details on the research strategy and design, in an attempt to minimize the risks of readers' misunderstanding. Besides, you should better justify the criteria which were used to build your sample. You properly state that a purposive sampling approach was used to select patients; however, I am wondering if purposive sampling generated biases in the collection and interpretation of the findings, in light of the small number of people included. Also, it is hard to maintain that you achieved data saturation after 17 interviews if you definitively used a purposive sampling. Probably, more details are needed about this issue. Lastly, yet importantly, you should clarify how codes were built: did you autonomously conceived it? Were they drawn on the scientific literature? Did you use a mixed approach? Since your findings are based on such coding, you should be more clear about it. After reading the manuscript, I think that you autonomously developed the coding design; however, you should clearly depict the approach you used in your research. Please, consider that some information included in the paper is redundant, since it is reported twice (e.g. the brief description of the pilot study with 5 participants): a careful review of the paper will allow to save some space, which could be used to address some shortcomings of this research.</p> <p>3) In my opinion, the description of the study sample (and, therefore, Table 1) will perform better if anticipated in the "methods" section. It is not clear how did you synthesized the initial 501 codes in 13 categories, which were then aggregated in 5 key themes. More attention and details should be spent on this issue. In the current version of the manuscript, the findings suffer from an excessively fragmented organization of the text. To deal with this problem, consider to avoid the use of subsections to describe the various key</p>
-------------------------	--

	<p>themes which emerged from the interviews' analysis. Rather, such sub-themes could emerge in the section which is devoted to the development of the theoretical framework.</p> <p>4) I really appreciate the theoretical framework you developed as a result of your analysis. However, it is poorly described in the current version of your manuscript. More information should be provided about the conceptual building blocks of such a theoretical framework, as well as on its ability to shed light on the effective use of numbers in HF self-care.</p> <p>5) In my own opinion, the discussion section is the weakest part of your paper. A major effort should be performed here to make this work publishable. Firstly, you should critically discuss the main by-product of your research, that is to say the conceptual framework developed at the end of section 3. Secondly, you throw a spotlight on some interesting and relevant issues (such as the influence of socio-demographic factors on patients' numeracy and self-care; the role played by informal caregivers in supporting patients; and patients' self-efficacy in dealing with HF); however, you do not dig into your findings to fully illuminate these issues and to push forward our understanding about them.</p> <p>6) Lastly, yet importantly, conclusions do not focus on the (conceptual and practical) implications of this study. Rather, they merely summarize the points which are debated in the discussions section. Obviously, this undermines the quality of the manuscript. I hope that these comments may help.</p>
--	---

### VERSION 1 – AUTHOR RESPONSE

Response to Reviewer #1:

1. The authors address an important topic that has implications for the approach to treating and educating patients with heart failure. Their framework may fill a gap in our understanding of the patient-provider relationship for patients with HF, although it will need to be tested in future work. The authors conducted a thoughtful qualitative study, paying attention to important details and following accepted processes. My questions and suggestions are minor and easily addressable.

We thank the reviewer for this feedback and appreciate her thoughtful review.

2. Abstract: Although it may be true, the last sentence of the abstract should be reworded to more accurately reflect the findings of this study. The study was not designed to compare the relative importance of individual patient numeracy to the role of other factors, please rephrase, e.g., something along the lines of “These findings suggest that HF-specific training of caregivers may play an important role in successful HF self-care. Future work addressing HF self-care should include HF-specific training of caregivers, and more work is needed to better understand the intricacies of the relationships between HF patients and their caregivers.

The reviewer makes an excellent point and we have changed the abstract to reflect this suggestion.

The abstract conclusion now reads: “Effective communication of numerical concepts which pertain to HF self-care is highly variable. Many HF patients lack basic understanding and numeracy skills required for adequate self-care. As such, patients rely on caregivers who may lack HF training. HF specific training of caregivers and research that seeks to elucidate the intricacies of the patient-caregiver relationship in the context of health numeracy and HF self-care, are warranted.

3. “ER” and “emergency room” should be changed to “emergency department” on pages 9 and 17, since this is the preferred term.

We have changed ER and emergency room to emergency department on pages 9 and 17.

4. Despite the focus on caregivers, interviews were conducted among patients and some conclusions related to caregivers extrapolate from patient report. The authors describe a gap between self-reported numeracy and use/understanding of numbers, which raises questions about the accuracy of patient comments regarding caregivers. The conclusions centered on caregivers appear to overstate the data available.

We thank the reviewer for this observation and agree, the interviews were conducted with patients only. We have modified the discussion section and the conclusion such that we do not overstate the available data (inappropriately extrapolate to caregivers).

5. Page 18 line 24, please change “is often” to “may be variable”, given the sampling approach.

We have changed often to “may be variable” on page 18.

6. Framework: Is the relationship between knowledge and communication uni-directional? Numeracy likely plays a role in the quality of communication. Many of the factors listed as impacting knowledge would also seem to impact patient-provider communication. Have you all discussed these questions/issues in development of the framework?

We thank the reviewer for these excellent points about our theoretical framework. Our data, as well as participant and researcher feedback enabled us to construct the framework. We also looked to existing frameworks on numeracy to ensure that we had not left out any important factors. The reviewer is correct, the relationship between knowledge and patient-provider communication is bi-directional, as is evidenced by our clinical experiences with HF patients and also by a previous general framework of health numeracy by Schapira et al (our co-author).<sup>1</sup> We have thus changed the solid, unidirectional arrow between the patient-provider communication box and the knowledge, understanding and application of numerical concepts box to a dashed, bi-directional arrow. In addition, the reviewer has made us further reflect on some of the details of the figure. We had intended bi-directional relationships to be represented by dashed blue lines, however we now feel it would be clearer if these are dashed bi-directional blue arrows. Finally, we added an additional dashed bi-directional blue arrow between the patient-provider communication box and the blue box directly beneath it. We did so because there is a bi-directional relationship between patient-provider communication and many of these factors (patient-provider trust, language, sensory deficits, etc). Finally, in order to more clearly define the categories of factors that may mediate the pathways described in the framework, we consolidated some categories previously listed. We have changed the text in the Results section as well as the figure legend to reflect these changes. Below is the revised framework:

#### Reviewer # 2

1. Thank you very much for the opportunity to read your interesting work. This paper tackles an important issue and adds some intriguing insights to push forward our understanding about the appropriate management of HF. However, after carefully reading your article, some concerns

appeared in my mind. I will summarize my suggestions to improve the quality of your paper in the following lines:

We thank the reviewer for all of his feedback and appreciate his comprehensive review of our work. We have revised the manuscript significantly and think that overall, it is much stronger, due to this reviewer's comments.

2. I found that the introduction was not able to "set the hook" and to catch the readers' attention. On the one hand, you emphasize what is the main topic of your research; on the other hand, you do little to stress the originality of your contribution and the relevance of your study. Therefore, I suggest to improve the introductory section: in particular, you should better state the specific purpose of your research and point out how it adds something new to the scientific knowledge.

We appreciate the reviewer's comment and have addressed them by adding more context to the third paragraph of the introduction. Namely, how our study fills an important gap in the literature. Starting from line 16, the last paragraph of the introduction now reads:

"...However, despite the degree to which numerical skills are needed for adequate HF self-care, to date only two studies have examined numeracy in the context of HF and they have done so by investigating the association between numeracy and the risk of readmissions among adults hospitalized for HF.<sup>2,3</sup> Although informative, the results of these studies were conflicting, and neither investigated the role of numeracy in the management of HF or from the patients' perspective. Herein we addressed this gap by examining how numerical information is viewed and used among community-dwelling adults with HF."

3. You argue that a "one to one semi-structured" approach was used to interview patients. However, the development of the manuscript suggests that a de-structured approach was used. Maybe, it would be worth providing more details on the research strategy and design, in an attempt to minimize the risks of readers' misunderstanding. Besides, you should better justify the criteria which were used to build your sample. You properly state that a purposive sampling approach was used to select patients; however, I am wondering if purposive sampling generated biases in the collection and interpretation of the findings, in light of the small number of people included. Also, it is hard to maintain that you achieved data saturation after 17 interviews if you definitively used a purposive sampling. Probably, more details are needed about this issue. Lastly, yet importantly, you should clarify how codes were built: did you autonomously conceived it? Were they drawn on the scientific literature? Did you use a mixed approach? Since your findings are based on such coding, you should be clearer about it. After reading the manuscript, I think that you autonomously developed the coding design; however, you should clearly depict the approach you used in your research. Please, consider that some information included in the paper is redundant, since it is reported twice (e.g. the brief description of the pilot study with 5 participants): a careful review of the paper will allow to save some space, which could be used to address some shortcomings of this research.

We thank the reviewer for the many points he raises regarding methodology. We attempt to address them one by one:

a. In this qualitative study, we interviewed 30 patients (one-on-one) using a semi-structured topic guide. The topic guide was developed from our understanding of the existing literature on numeracy in other chronic diseases as well as our experience taking care of HF patients. We have added this information to the Data Collection section of the methods. We have also included the topic guide as a supplemental file so our methods are more transparent for the reader (Appendix 1).

b. We thank the reviewer for his comments about our approach to the research strategy and design. We used a semi-structured approach to interviewing, which meant that we used a standard interview guide, which included probes to elicit clarification and additional detail from participants (Appendix 1). This method allowed us to outline the topics that we wished to cover, but also allowed the interviewees responses to set the flow and tone of the interview. This approach is the most commonly used type of interview in qualitative studies which pertain to self-care. The semi-structured guide provided the interviewer with a clear set of instructions such that consistency occurred across all 30 interviews, and also allowed for reliable, comparable qualitative data. One interviewer did all 30 interviews and was trained in qualitative interviewing. Finally, we also included open open-ended questions which allowed for participants to share new ways of seeing and understanding numeracy in HF self-care.<sup>4</sup> We have added 2-3 sentences to the Methods section to clarify our approach.

c. We used purposive sampling to achieve a balanced sample with respect to gender, race/ethnicity, and duration of HF (years), since prior qualitative studies have shown variation across these characteristics with respect to self-care.<sup>1,5,6</sup> Purposive sampling is a technique widely used in qualitative research for the identification and selection of information-rich informants that cover the range of variation.<sup>7</sup> Purposive sampling has been used widely in qualitative research which pertains to patient care, including studies published in BMJ Open.<sup>8,9</sup> We agree with the reviewer that the sample is not fully generalizable. As in many qualitative research studies, the sample size is small and random sampling of a target population was not conducted. The goal of purposive sampling, however, was to elicit a range of experiences that reflect a sample diverse in gender, race/ethnicity and duration of HF, which we believe we successfully accomplished.

d. While we agree with the reviewer, that the sample may not be fully generalizable, we do not feel this is because of purposive sampling. If anything, the sample is more diverse than if we had just used convenience or snow ball sampling, as most qualitative studies do. Our findings may lack generalizability since we interviewed patients from one, urban clinic in New York, NY and thus may not be applicable to HF patients who are cared for in other settings (non-teaching hospitals, global setting, or in rural areas). To more fully explain our sampling approach to the readership, we have added detail to the second paragraph of the Participant Selection and Study Setting section of the Methods section and to the limitation section in the Discussion.

e. We agree with the reviewer that we could provide more information on how we reached data saturation. By team consensus, we did indeed reach data saturation at the 17th interview, which we defined as the point at which no new themes emerged.<sup>10,11</sup> Work by Morse et al (1995) highlights that saturation in qualitative research occurs quickly when a cohesive sample is obtained and if purposive sampling is used.<sup>11</sup> Thus, we were not surprised when at the 17th interview we hit saturation. One of the guiding methodological papers on data saturation by Guest et al (2006) found that in a study where purposive sampling was used and 60 one-on-one interviews were conducted, data saturation occurred at the 12th interview.

f. We thank the reviewer for his comments about the process in which we coded our data and developed categories and subsequent themes. Because numeracy has not been examined in the context of HF, we took an inductive analytic approach, meaning that we used detailed readings of raw data to derive open codes, categories, and themes, and ultimately a model through interpretations made from the raw data by an evaluator or researcher.<sup>12</sup> Thus, he is correct, we autonomously developed the codes from the data. This general inductive approach is commonly used in health and social science research and evaluation which use grounded theory,<sup>13</sup> as we did. Our coding process can be explained and visualized by the following table, taken from Thomas et al (2006):

We have added more detail to the Data Analysis sub-section of the Methods section such that there is more information and transparency on our coding approach.

g. To address redundancy, we removed the 5-patient pilot from the Data Collection section of the methods. This information, as correctly pointed out by the reviewer, is discussed in the Quality Assurance and Patient Involvement section of the Methods.

a. In my opinion, the description of the study sample (and, therefore, Table 1) will perform better if anticipated in the "methods" section. It is not clear how did you synthesized the initial 501 codes in 13 categories, which were then aggregated in 5 key themes. More attention and details should be spent on this issue. In the current version of the manuscript, the findings suffer from an excessively fragmented organization of the text. To deal with this problem, consider to avoid the use of subsections to describe the various key themes which emerged from the interviews' analysis. Rather, such sub-themes could emerge in the section which is devoted to the development of the theoretical framework.

We thank the reviewer for these very helpful comments and have reworked the Methods section in several ways to address them.

We have refined the Data Analysis section, as well as the first sentence of the Result section, both of which speak to the coding process. The 501 codes were the result of the two independent coders review of the initial transcripts. A third independent reviewer reviewed both lists of codes (many of which were duplicative), as well as the transcripts, and consolidated them into a uniform codebook consisting of 77 unique codes. The two coders then independently applied these codes to the initial transcripts and then subsequent transcripts. During this process, the two coders met to revise the codebook, removing some codes while adding others. Next, common codes were compared using dimensions and properties. These were collapsed into broader categories (13) which then evolved into even broader themes (5). All three reviewers participated in this process, which was achieved by consensus. A fourth reviewer reviewed the final list of themes and read the transcripts to ensure that the themes reflected the patient perspective. Quotations from each theme were anonymized and selected for presentation. This iterative analytic process has been used widely in the field,<sup>14</sup> including authors who have published in BMJ Open,<sup>15,16</sup> and we have revised the text to provide sufficient detail of this process.

With respect to our presentation of the five themes and their sub-themes (what we consider to be categories), we appreciate the reviewers' comments. We have significantly revised the presentation of the Results section by removing the sub-themes (categories) altogether and creating a unified text (with quotations) for each of the five themes. We feel that this approach dramatically decreases the fragmentation that the reviewer had trouble with. We noticed that many other studies in BMJ conform to this presentation, and thus feel like this revised approach will be compatible with the journals' readership.

b. I really appreciate the theoretical framework you developed as a result of your analysis. However, it is poorly described in the current version of your manuscript. More information should be provided about the conceptual building blocks of such a theoretical framework, as well as on its ability to shed light on the effective use of numbers in HF self-care.

We thank the reviewer for this comment and have addressed it in 4 main ways. First, we have added a sub-section in Methods section (entitled 'Theoretic Framework') which describes how it was developed from the data and our analytic process. Second, we have revised Figure 1, per Reviewer #1's suggestions (see above), to make it visually clearer. Third, we have added additional explanation of the framework to a sub-section of the Results section. Finally, we have woven into the discussion (paragraphs 3, 4, 5, and 6) how the framework explains our findings and how it could be used to guide future research and potential interventions.

Please see the Methods, Results, and Discussion sections for all of these changes.

c. In my own opinion, the discussion section is the weakest part of your paper. A major effort should be performed here to make this work publishable. Firstly, you should critically discuss the main by-product of your research, that is to say the conceptual framework developed at the end of section 3. Secondly, you throw a spotlight on some interesting and relevant issues (such as the influence of socio-demographic factors on patients' numeracy and self-care; the role played by informal caregivers in supporting patients; and patients' self-efficacy in dealing with HF); however, you do not dig into your findings to fully illuminate these issues and to push forward our understanding about them.

We acknowledge the reviewer's comments and have addressed them in several ways, which we think significantly strengthen the discussion section and the paper over all.

1. We have reworked the entire discussion section and have given special emphasis to how our findings can shed light on future research the design of interventions which pertain to numeracy in HF.
2. We have guided the reader through the implications of the theoretical framework and how it can be used by other investigators in Discussion paragraphs 4, 5, and 6.
3. We also highlight some shortcomings in the limitations section of the discussion, which we hope future studies can address.

d. Lastly, yet importantly, conclusions do not focus on the (conceptual and practical) implications of this study. Rather, they merely summarize the points which are debated in the discussions section.

We agree that the Conclusions could be improved, as also suggested by Reviewer #1, and have done our best to not just summarize our main findings, but also provide implications of our work for both clinicians caring for HF patients and for researchers who wish to build upon the study's findings.

References for this Response Letter (which have also been incorporated into the paper):

1. Schapira MM, Fletcher KE, Ganschow PS, et al. The meaning of numbers in health: exploring health numeracy in a Mexican-American population. *J Gen Intern Med.* 2011;26(7):705-711.
2. McNaughton CD, Collins SP, Kripalani S, et al. Low numeracy is associated with increased odds of 30-day emergency department or hospital recidivism for patients with acute heart failure. *Circ Heart Fail.* 2013;6(1):40-46.
3. Sterling MR, Safford MM, Goggins K, et al. Numeracy, Health Literacy, Cognition, and 30-Day Readmissions among Patients with Heart Failure. *Journal of hospital medicine.* 2018.
4. Stuckey HL. Three types of interviews: Qualitative research methods in social health. *J Soc Health Diabetes.* 2013;1:56-59.
5. Schapira MM, Fletcher KE, Gilligan MA, et al. A framework for health numeracy: how patients use quantitative skills in health care. *J Health Commun.* 2008;13(5):501-517.
6. Abdel-Kader K, Dew MA, Bhatnagar M, et al. Numeracy Skills in CKD: Correlates and Outcomes. *Clin J Am Soc Nephrol.* 2010;5(9):1566-1573.
7. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Adm Policy Ment Health.* 2015;42(5):533-544.

8. Leung AY, Bo A, Hsiao HY, Wang SS, Chi I. Health literacy issues in the care of Chinese American immigrants with diabetes: a qualitative study. *BMJ Open*. 2014;4(11):e005294.
9. Hipwell AE, Sturt J, Lindenmeyer A, et al. Attitudes, access and anguish: a qualitative interview study of staff and patients' experiences of diabetic retinopathy screening. In: *BMJ Open*. Vol 4.2014.
10. Guest G, Bunce, A., Johnson, L. . How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods*. 2006;18(1):59-82.
11. Morse JM. The Significance of Saturation. *Qualitative Health Research*. 1995;5(2):147-149.
12. Thomas DR. A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*. 2006;27(2):237-246.
13. Strauss Ca. *Basics of Qualitative Research*. 2015; <https://us.sagepub.com/en-us/nam/basics-of-qualitative-research/book235578>.
14. Bradley EH, Holmboe ES, Mattera JA, Roumanis SA, Radford MJ, Krumholz HM. A qualitative study of increasing beta-blocker use after myocardial infarction: Why do some hospitals succeed? *Jama*. 2001;285(20):2604-2611.
15. Schwappach DL, Gehring K. 'Saying it without words': a qualitative study of oncology staff's experiences with speaking up about safety concerns. *BMJ Open*. 2014;4(5):e004740.
16. MacKay C, Jaglal SB, Sale J, Badley EM, Davis AM. A qualitative study of the consequences of knee symptoms: 'It's like you're an athlete and you go to a couch potato'. *BMJ Open*. 2014;4(10):e006006.

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Palumbo Rocco University of Salerno, Italy
<b>REVIEW RETURNED</b>	10-Jul-2018
<b>GENERAL COMMENTS</b>	Thank you very much for this revised version of the manuscript. You were effective in addressing all my concerns. Therefore, I have no further suggestions to improve the quality of this research. Well done for this interesting research!