The relationship between completing standardised questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-004363
Article Type:	Research
Date Submitted by the Author:	30-Oct-2013
Complete List of Authors:	Holmberg, Christine; Charité Universitätsmedizin Berlin, Institute for Social Medicine, Epidemiology, and Health Economics Karner, Julia; Charité Universitätsmedizin Berlin, Institute for Social Medicine, Epidemiology, and Health Economics Rappenecker, Julia; Charité Universitätsmedizin Berlin, Institute for Social Medicine, Epidemiology, and Health Economics Witt, Claudia; 1Charité University Medical Center, Institute for Social Medicine, Epidemiology and Health Economics; University of Maryland School of Medicine, Center for Integrative Medicine
Primary Subject Heading:	Qualitative research
Secondary Subject Heading:	Geriatric medicine, Communication, Complementary medicine
Keywords:	QUALITATIVE RESEARCH, PAIN MANAGEMENT, GERIATRIC MEDICINE

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
R…RELEVANCE OF STUDY QUESTION	
Is the research question interesting? *Yes.* Is the research question relevant to clinical practice, public health, or policy? *Yes, the research question is relevant to clinical practice and public health. The research question addresses one of the major pillars of clinical trials regarding problems associated with filling out questionnaires by the elderly, a population in which more health research is necessary. This is detailed in the "introduction" section.*	Research question explicitly stated Research question justified and linked to the existing knowledge base (empirical research, theory, policy)
A…APPROPRIATENESS OF QUALITATIVE METHOD	
Is qualitative methodology the best approach for the study aims? • Interviews: experience, perceptions, behaviour, practice, process *The chosen qualitative methodology, interviews, provided an effective way to gain deeper insight as to how participants translate (or not translate) experience into validated questionnaires. This is stated in the methods section.*	Study design described and justified i.e., why was a particular method (e.g., interviews) chosen?
T…TRANSPARENCY OF PROCEDURES *Sampling*	
Is the sampling strategy appropriate? *To create the random sample a data manager designated each QIBANE participant with a number and then chose random numbers using SPSS. Then a ranking list was created that randomly selected QIBANE participants from the group with improvement, with worsening, or with no change from baseline to follow-up assessment. This approach was necessary because we aimed to include the different experiences possible in QIBANE that we captured in the RCT also in the interview study. Thus, this was in line with a maximum variation sampling.*	
Are the participants selected the most appropriate to provide access to the type of knowledge sought by the study? *The goal of the study was to understand how elderly women transfer their experiences onto validated study instruments used in an RCT. Thus, the population of the RCT was the most appropriate population to sample from.*	Criteria for selecting the study sample justified and explained • *theoretical:* based on preconceived or

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
	emergent theory • *purposive:* diversity of opinion • *volunteer:* feasibility, hard-to-reach groups
Recruitment	
Was recruitment conducted using appropriate methods? *A member of the RCT research team called randomly chosen participants from the RCT and asked for permission to conduct an interview with them. The interview was conducted at the home of the interviewee. This location was chosen to make participation easy for the elderly and because it has been shown that interviews conducted at places in which interviewees feel at home help create an atmosphere that facilitates interviewing. We explain this in the "data collection" section of the paper.*	Details of how recruitment was conducted and by whom.
Is the sampling strategy appropriate?	
In the results section under "sample" we describe that 6 people who were asked to participate refused for fear of fraud.	Details of who chose not to participate and why-
Data collection	
Was collection of data systematic and comprehensive? *Yes. This is described under data collection.*	Method(s) outlined and examples given (e.g., interview questions)
Are characteristics of the study group and setting clear? *The QIBANE RCT, in which this study was nested, mostly consisted of female participants (95%). Thus, the female sample in the interview study is a reflection of the RCT population which in turn is a reflection of the larger proportion of females in this age group overall.* *The qualitative study group was elderly (mean age: 76 ± 8 years) and female with previous neck pain.* *The interview setting was at the participants' home.*	Study group and setting clearly described

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
Why and when was data collection stopped, and is this reasonable? *Data collection ended after 10 interviews from each intervention group (20 total) were conducted. The sample size of 20 was chosen based on other qualitative interview studies that are nested within RCTs (detailed in "study design" section).*	End of data collection justified and described
Role of researchers	
Are the researcher(s) appropriate? *All researchers were well trained in qualitative interviewing and qualitative content analysis. The research team consisted of MDs, epidemiologists, and an anthropologist. This range of disciplines allowed for a broad view onto the topic and onto the analysis of the materials. These different views that were brought to the materials and discussed in regular team meetings were able to highlight the different assumptions everyone brought to the materials and to ensure a rigorous analysis of materials.* How might they bias (good and bad) the conduct of the study and results? *The senior researcher of the project (CW) was the PI of the entire study. Similarly, those who called QIBANE participants to participate in the interview portion of the study had worked on the RCT. This could bias the interviewing in that interviewees would maybe not freely discuss problems they had with the RCT. Also the interviewer was an MD which may have led to participants (elderly women) be too respectful to present negative views or personal experiences. We carefully scrutinized the interview materials for such cues and found that interviewees indeed were not openly opposing the interviewer but did so clearly in small statements. All interviewees talked freely about their difficulties and about the nonsense they thought was asked in the questionnaires. Similarly, all interviews were listened to immediately after the interview took place by CH to pick up on any problems that may arise due to inadequate questioning and each interview was then discussed by CH and the interviewer. To detect additional biases the interviewer wrote an interview protocol after each interview in which such things were recorded as when the interviewer felt awkward asking a question or what dynamics developed during the interview. We talk about this in the method and in the discussion section of the paper.*	Do the researchers occupy dual roles (clinician and researcher)? Are the ethics of this discussed? Do the researcher(s) critically examine their own influence on the formulation of the research question, data collection, and interpretation?

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
Ethics	
Was informed consent sought and granted? *The RCT was approved by the appropriate ethics review board (EA1/265/05). Participants gave written and oral consent to the intervention study. The interview study was part of the intervention study and selected participants were invited by phone to participate in an interview on their experiences with the RCT. The interviews took place at their home and they were asked to provide additional oral consent for a home visit. The consent process was documented in the case report forms We detail this in the paper in the "Methods" section under "data collection" and "data analysis."*	Informed consent process explicitly and clearly detailed
Were participants' anonymity and confidentiality ensured? *All interview materials was pseudonomysed. The process is detailed in the informed consent and was discussed with interview participants prior to the interview.*	Anonymity and confidentiality discussed
Was approval from an appropriate ethics committee received? *Yes. The intervention study was approved by the ethics review board of the Charité Universitätsmedizin Berlin (EA1/265/05). This is stated in the Methods section under "study design"*	Ethics approval cited
S..SOUNDNESS OF INTERPRETIVE APPROACH *Analysis*	
Is the type of analysis appropriate for the type of study? • The analysis was conducted using content analysis and comparing filled out questionnaires with the results of the content analysis. Since the aim of the study was to identify how RCT participants fill out validated questionnaires this was the most appropriate method of analysis. Codes for the content analysis were developed inductively to capture all themes the interviewees brought up. Are the interpretations clearly presented and adequately supported by the evidence?	Analytic approach described in depth and justified *Indicators of quality:* Description of how themes were derived from the data (inductive or deductive) Evidence of alternative explanations being sought Analysis and presentation of negative or deviant cases
Are quotes used and are these appropriate and effective? *All themes that were detected through the analysis were presented in the results and are supported by a quote. Quotes*	Description of the basis on which quotes were chosen Semi-quantification

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
were selected to reflect the identified theme (strategy to fill out questionnaire). All themes are semi-quantified to show how often it approximately appeared.	when appropriate Illumination of context and/or meaning, richly detailed
Was trustworthiness/reliability of the data and interpretations checked? *As a first step the qualitative interview materials were read and then analyzed independently by two researchers using content analysis.[1-3] This allowed focusing on the interview passages in which the questionnaires were discussed. The coding scheme was developed based on the interview material by two of the authors (JK and JR) and then refined by the research team (all authors). In addition, coding and results were regularly presented and discussed in a qualitative working group. The goal of the presentation in the working group was to ensure that materials and results were consistent with each other and to broaden the perspectives on the materials and ensure intersubjectivity of results. After analysis of interviews, we compared the quantitative questionnaires that had been completed by the interviewees in the RCT with* interview *results to identify strategies of how they were completed. This is detailed in the method section of the paper.*	Method of reliability check described and justified e.g., was an audit trail, triangulation, or member checking employed? Did an independent analyst review data and contest themes? How were disagreements resolved?
Discussion and presentation	
Are findings sufficiently grounded in a theoretical or conceptual framework? Is adequate account taken of previous knowledge and how the findings add? *Yes, both are addressed in the discussion of the paper.*	Findings presented with reference to existing theoretical and empirical literature, and how they contribute
Are the limitations thoughtfully considered? *Yes. We do this at the end of the discussion section.*	Strengths and limitations explicitly described and discussed

REFERENCES

1. Morse JM, A. Tashakkori, C. Teddlie. Principles of mixed methods and multimethod research design. In: Tashakkori A, C. Teddlie editor. Handbook of mixed methods in social and behavioral research. Thousand Oaks, CA: Sage 2003:189-208.
2. Flick U. Qualitative Forschung - ein Handbuch. Reinbek bei Hamburg: rowohlts enzyklopädie 2007.

3. Mayring P. Einführung in die Qualitative Sozialforschung. Weinheim und Basel: Beltz Studium 2002.

1. TITLE:

The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study

2. CORRESPONDING AUTHOR:

Christine Holmberg

Institute for Social Medicine, Epidemiology, and Health Economics

Charité Universitätsmedizin Berlin

Luisenstrasse 57

Berlin 10117 Germany

Email: christine.holmberg@charite.de

Phone: +49 30 450 529 192

Fax: +49 30 450 529 917

2. AUTHORS:

Christine Holmberg,[a,b] Dr. phil., M.P.H., (christine.holmberg@charite.de), Julia Karner,[a] (julia.karner@charite.de), Dr. med., Julia Rappenecke,[a] (julia.rappenecker@hotmail.com), Dr. med., & Claudia Witt,[a,c] (Claudia,witt@charite.de), Professor, Dr. med.

[a] Institute for Social Medicine, Epidemiology, and Health Economics, Charité Universitätsmedizin Berlin, Berlin, Germany

[b] Berlin School of Public Health, Charité Universitätsmedizin Berlin, Berlin, Germany

[c] University of Maryland School of Medicine, Center for Integrative Medicine, Baltimore, USA

4. KEYWORDS

validation, study instruments, clinical trial participants, elderly, experience, quantification

5. WORD COUNT

4192

ABSTRACT

Objectives:

In order to improve clinical study developments for elderly populations, we aim to understand how elderly participants in research studies transfer their experiences into validated, standardized study measurement instruments. Particularly, we analyzed how women (mean 78 ± 8 years of age) who participated in an RCT cognised the study instruments that were used to evaluate the outcome of the intervention.

Setting:

The interview study was situated in an elderly community in Berlin, Germany. The elderly have the option of living on their own or of assisted-living in this community.

Participants:

The sample for the interview study was selected from an RCT. The sample of the RCT was mostly female (95%) and on average 76 years of age (SD=±8 years). From this sample, a purposive sampling list was created for the qualitative interview study based on the outcome scoring in the RCT. 20 participants of the RCT were included in the interview study. All interview participants were female.

Outcomes:

We asked patients about their experiences of completing questionnaires in the RCT. Interviews were analysed thematically and then compared to the questionnaires.

Results:

Interviewees had difficulties translating complex experiences into a single value on a scale and understanding the relationship of the questionnaires with the study aims. Interviewees thought it was important for the trial that their actual experiences were understood by trial organizers. This information was not transferrable by means of the questionnaires. To rectify these difficulties, interviewees used strategies such as adding notes, adding response categories, or skipping an item.

Conclusion:

Participants in the RCT understood their importance for the study and wanted to convey their personal experiences as best as possible. This led to strategies that resulted in "missing data". To improve data collection in elderly populations, educational materials addressing the differential logics should be developed and tested.

STRENGTHS AND LIMITATIONS OF THE STUDY

- This qualitative study gives insight into how elderly women think about and fill out validated study instruments.
- Interviewed women used satisficing strategies to complete questionnaires and making notes to convey their experiences to study personnel to ensure that "good" information was collected in the study.
- This differential logic led to strategies of completing questionnaires that produce missing data
- Increasing elderly participants' understanding of research improves data collection
- Data collection was conducted by clinical research staff. This may have influenced participants' ease to be honest and critical of their experience with the questionnaires. In addition, findings should be tested in other elderly study populations.

INTRODUCTION

There are many factors that are crucial to the success of clinical trials, including validated study instruments. An adequate assessment of the study endpoint is a crucial aspect of clinical trials; for this validated questionnaires are considered good assessment tool for the aim. The utilized instruments should be able to measure the same constructs across individuals. Some strategies of completing questionnaires that hinder adequate assessment such as putting a mark on the mean value of a scale are well known. Much effort has been devoted to the design of study instruments to discourage such behavior.

The analysis of inherently subjective experiences such as chronic pain has been shown to be difficult with existing measurements. Diagnosing chronic pain poses problems to researchers and clinicians.[1-3] Pain is a subjective experience that cannot be directly measured[4-7] This makes an objective evaluation of pain extremely difficult. The problem lies in trying to quantify a subjective and complex experience in such a way that it can be reproduced.[8] A series of pain scales based on self-evaluation have been developed.[9] There exist one-dimensional pain scales to capture pain intensity and multi-dimensional ones to capture intensity and duration of pain. In clinical practice and research, one-dimensional scales are usually used,[10] as multi-dimensional scales are more complex to administer.[11] A commonly used one-dimensional pain scale is the visual analogue scale (VAS). The VAS has shown good validity and reliability and is easy to use.[12] Overall self-reported pain scales are considered the gold standard for pain assessment.[13] However, pain is a fluctuating experience which one-dimensional self-report pain scales cannot capture adequately.[4, 6, 9, 14, 15] They continue to be in high demand in clinical and research practice because chronic pain presents a severe problem to all health care systems.[1, 8, 16, 17] Chronic pain is particularly prevalent in female elderly patients[18] arising from the musculoskeletal system in particular such as neck or back pain.[19-21] Thus, known shortcomings in pain assessment tools may make pain management in the elderly difficult.[2, 22-26] For instrument use, the most important performance criteria are the validity and reliability of instruments. Validity corresponds to the question of how well an instrument measures what it intends to measure, such as pain intensity.[27] For example, the

validity for pain scales is based on tests and retests as well as their comparability with other scales.[14, 28] This means that if an instrument is used repeatedly and achieves the same results throughout or gives similar results to an instrument that has already been validated then its results are considered valid. However, good results for these performance criteria do not necessarily suggest that they can depict complex subjective experience.

In an RCT that compared the effects of Qigong and exercise therapy on neck pain in the elderly, no effect on pain intensity could be detected.[29] Three groups were compared, a Qigong group, an exercise therapy group, and a waiting list group. No difference between groups was found for the primary (VAS) and the secondary endpoints (Neck, Pain and Disability Scale depression, based on a common depression, health-related quality of life, sleep quality, and satisfaction with the therapies). However, patients were highly satisfied with the interventions and some even chose to continue the interventions at their own expenditure. Thus, we were interested to understand how participants transferred their observations and experiences into the study measurement instruments.

This analysis aims to understand how women (mean 78 ± 8 years of age) who participated in an RCT cognised the study instruments that were used to evaluate the outcome of the intervention.

METHODS

Study Design

We conducted a qualitative interview study based on an RCT that aimed to better understand the results of the RCT.[29] The RCT was approved by the appropriate ethics review board (EA1/265/05). Participants gave written and oral consent to the intervention study. The interview study was part of the intervention study and selected participants were invited by phone to participate in an interview on their experiences with the RCT. The interviews took place at their home and they were asked to provide additional oral consent for a home visit. The consent process was documented in the case report forms. The RCT included 117 patients with chronic neck pain that were randomized to a Qigong group, an exercise therapy

group, or to a waiting list group. At three different time points, all three groups completed four validated questionnaires: the VAS, the Neck Pain and Disability Scale (NPDS),[30] the Short-Form.36-Questionnaire (SF-36),[31, 32] and a common depression scale (ADS).[33] The NPDS is a specific evaluation instrument for neck pain that has shown to be valid and reliable to measure neck pain[34-36] and to detect clinically relevant changes in neck pain.[37] It consists of 20 items that assess intensity of pain using neck problems as well as emotional and cognitive influences on work and everyday life[38]. The ADS assesses length and adverse effects of depressive symptoms, bodily problems and negative thought patterns. It is the German version of the Center for Epidemiological Studies Depression Scale (CES-D).[39] This instrument is recommended for use with chronic pain patients.[40]

We developed a guideline for semi-structured in-depth interviews that included questions on interventions and study instruments asking about difficulties the patients may have had in completing the questionnaires and what was important for them in relation to the study interventions. Prior to the study the interview guideline was tested and used in practice interviews with older patients with neck pain to ensure that the questions functioned well and the information was received as intended by the study aims.

Recruitment

A systematic sample was selected from the participants of the RCT. The RCT participants were ranked according to the difference between baseline and follow-up in the VAS. In each group a ranking with those at the beginning who showed the biggest difference between follow-up and baseline was developed. The RCT participants were systematically telephoned according to the ranking and asked to participate in the in-depth interview study. Those who called participants had previously conducted the RCT and were known to participants. Participants in the RCT were mostly female (95%) which led to a ranking list that was predominantly female. Recruitment ended after the first ten RCT participants from the Qigong and another ten from the exercise therapy group had agreed to participate. A sample size of twenty participants was chosen based on the experiences of other qualitative studies that were nested in RCTs.[41, 42]

Data collection

Interviews were conducted in the homes of the participants to ensure that participants felt comfortable and were willing to speak openly.[43] Interviewers had previously organized the RCT and were well-known to the interviewees. Interviews were conducted at the home of the interviewees to accommodate study participants and to create a relaxing atmosphere for the interviewee.[43] To help their memory, interview participants received blank sample questionnaires similar to the ones they had filled out during their RCT participation. While an interview guideline was used for the interview, it was used in a flexible manner to give space for themes that were important to the interviewees.[44, 45] After each interview, a protocol was written by the interviewer to capture the atmosphere of the interview. Interviews were digitally recorded and transcribed. The text documents were then entered into software program ATLAS.ti for coding and analysis.

Data Analysis

Because interviewers were not involved in data analysis, the interview protocols provided the contextual information for the research team to situate the interview, its dynamics and its content. Analysis of the study was multi–layered. As a first step the qualitative interview materials were read and then analysed independently by two researchers using content analysis.[44-46] This allowed focusing on the interview passages in which the questionnaires were discussed. The coding scheme was developed based on the interview material by two of the authors (JK and JR) and then refined by the research team (all authors). In addition, coding and results were regularly presented and discussed in a qualitative working group. The goal of the presentation in the working group was to ensure that materials and results were consistent with each other and to broaden the perspectives on the materials and ensure intersubjectivity of results. After analysis of interviews, we compared the quantitative questionnaires that had been completed by the interviewees in the RCT with interview results to identify strategies of how they were completed.

RESULTS

Sample Description

Of those who were called and invited to participate in the interview study six declined a home visit due to fear of fraud. The remaining twenty people agreed to participate in the interviews. Table 1 shows the changes the interviewees had indicated on the validated scales during the RCT. Eleven of the interviewees indicated a wish to continue the therapy even though they had not experienced an improvement of pain according to the validated instruments.

Table 1: Changes in measurements between baseline and primary endpoint of the interviewees.

Questionnaire[a]	**Improvement** (number of patients)	**Worsening** (number of patients)	**Missing Data** (number of patients)
VAS[b]	9	11	0
NPDS[c]	13	6	1
SF-36[d]	11	8	1
ADS[e]	5	8	7

[a]One participant had no change in the NPDS

[b]**VAS**: Visual Analogue Scale, [c]**NPDS**: Neck Pain and Disability Scale, [d]**SF-36** (mcs = mental component score: Mental component summary scale of the Short-Form-36-Questionnaire), [e]**ADS**: Common Depression Scale

All interviewees were female with an average age of seventy-six years of age. They had an age range of 67-85 years. On average they had experienced pain for fifteen years. All interviewees lived in residencies for seniors in Berlin.

Experiences completing the questionnaires

Many of the interviewees were dissatisfied either with the questionnaires that they had to complete or the strategies they used to complete it. They complained about the difficulties of expressing complex experiences in the standardised terms the questionnaire asked of them.

> "Questionnaires are always terrible because you never can express by checking a box what one wants to say." [QG2/241]

> "If I make this movement, it hurts here. If I make that movement, it hurts there. Now the pain is gone. Now I look at you and I don't experience any pain. Now you tell me, do I have pain or do I not have pain? You tell me!" [QG2/318]

Some women were also concerned about the type of questions that were asked of them; questions related to their mental state as in the ADS and partially in the NPDS were especially disconcerting to some interviewees. Some interviewees were concerned that study staff may not adequately interpret their answers in the questionnaires because they were not able to precisely express what they felt through the questionnaires.

> "In these questions one often has potential answers that partially fit and partially do not fit, so that one would say, 'yes, that is how it is, but….' (…) and since there is no possibility for the opposite, the whole answer isn't right." [QG10/011]

None of the interviewees felt that their experiences with pain or with living as an elderly person could be adequately described with the questionnaires that were given to them. There were particularly two difficulties discussed by the participants. These were translating complex experiences into a single value on a scale and understanding the relationship between the questionnaires and the study aims. Participants used different strategies to deal with these problems when they completed the scales. These were mainly additional notes, placing the mark in the middle of a scale, adding answer categories, or skipping an item. The women used these strategies because they felt that the scales could not capture their experiences. At the same time at least some felt indebted to the study since it gave them free exercise classes and they wanted to attend to the questionnaires in the best possible manner. Thus they added to the questionnaires the information they found pertinent.

Specificing standardised answers

Adding notes

Adding notes was a common strategy among the interviewees. Of the interviewees, 15 added information on some item to clarify what the value on the scale they marked signified. For example, one participant added to her answer for the item, "frequency of physical activity," the time frame, "*30 or 60 minutes!"* and the circumstances of the exercise, "*with partner or by myself,*" [PN7]. The same participant added to her answer to the item, "frequency of falls," "*in snow*." She had indicated that she had fallen once. Lastly in the NPDS the patient wanted to specify her pain and added "*in the lumbar spine and in the knees*." Another participant added a note to the value she selected on the VAS, "*I exercise daily. This is the only way I can remain relatively painless*," [PN6].

Others substituted selecting a value on the scale by adding handwritten notes to the response options. For example, one participant [QG5] added verbal signifiers to the scale on the NPDS such as "*seldom*," "*satisfied*," or "*little*." In the interviews, this particular participant complained about the questionnaires. Another participant [QG6] specified one question in the NPDS in the interview. Instead of putting a mark next to the question "does the pain hinder you with activities such as eating, dressing, or hygiene?" the participant responded by writing "*dressing*."

Similarly, where items asked for specific time frames, participants sometimes chose to change the time frame in order to meaningfully answer the questions. They noted on the side the time frame they are referring to in the answer. For example, for a question that asked for a judgment of the last three months, one respondent wrote, *"[t]his has been in the last six months*," [QG10]. The theme addressed in the question seemed more important to the interviewees than the requested time frame.

Selecting parts of an item

Another strategy to specify general questions was to underline parts of a question to highlight to what exactly the answer referred to. For example, one participant [PN7] underlined "kneeling" in an item of the

SF-36 that stated “to bend forward, kneeling.” Another such example comes from an either/or question in the NPDS. Two of the participants [PN1] [QG4] marked one of the two given possibilities in the item “How difficult is it for you to look up or down?” Underlining was also used in questions that required a response along a scale. Several of the interviewees simply underlined one of the top or bottom values on the scale instead of marking a point along the scale.

Being a study participant

Some of the women had a clear understanding of the reciprocal relationship with the staff of the RCT; the women received interventions in exchange for completing the questionnaires. However, this required that they seriously considered their responsibility and wanted to complete the questionnaire adequately. Further, the questionnaires used in the study left some women feeling uneasy and unhappy with their contribution.

> “I was actually glad when I was done, just like school work that I had to do and I did very thoroughly. But I was not satisfied with my work and also not with the questions! So, I wasn’t– but I have experienced such feelings with other questionnaires before.” [QG2/265]

> “I hope I have filled out everything correctly. I do not know if I filled them out correctly.” [PN9/163]

One woman called a family member and her family physician to assist her in completing the questionnaire in order to ensure the correctness of the questionnaires.

> “I don’t remember for which question that was. I really did not know what to do with that question. I did not want to do anything wrong, so I called my daughter. She is a teacher and she

also really had to think about it. But I cannot tell you which question that was at the time. I don't know. But the question was phrased very strange." [QG8/207]

Seriously considering their role as study participant was a common theme in the interviewees and was the main reason why interviewees were dissatisfied with the assessment tools used in the RCT. The interviewed women assumed that their precise and exact experiences were of importance to the clinical trial staff and they were very concerned that the staff could not interpret their marks correctly on the assessments. The women made clear in the qualitative interviews that they preferred such an assessment much more than the questionnaires because the interviews enabled them to correctly state their experiences.

"One just could not answer that question clearly. I don't know. I basically followed my feelings– but did you understand my answer? What you really get out of my answer is the question […]. So I had the feeling after I filled out the questionnaire that you cannot learn anything from those answers. I guess I would have to say: I would not trust those questionnaires. But those are your main interest, aren't they?!" [QG6/028]

"The questions [in the questionnaire] do not make sense. I would have thought it better if you would have, just like you are doing now, asked the people directly." [QG4/250]

DISCUSSION

The interviewees in this study considered their role as a study participant important and perceived it as their responsibility to adequately answer the questionnaires. However, this was not easy for them. They had difficulties making nonspecific statements about specific experiences and many thought that their experiences could not be depicted in the questionnaires; many also feared that their answers could be

misunderstood. Several strategies were used to deal with the problem, such as adding notes, marking particular parts of a question, or leaving an item open. Some also asked others to help them complete the questionnaires correctly. Strategies such as adding notes have been called “optimising” strategies.[47] In addition, leaving an item blank or putting the mark in the middle of a scale are called “satisficing” strategies, suggesting that questions are answered cursorily.

Satisficing strategies are more common when study participants do not understand why certain questions are asked.[48] This is what the interviewees described especially in relation to the ADS. In the rationality of the researchers, it was necessary to use the ADS because an association between depression and chronic pain has been found before. However, the association between mental state and pain was one that alienated research participants from the study. The age group of the interviewed women may be one in which depression and other psychiatric diseases have a strong stigma associated with it. In addition, women in that age group may still belong to a generation that had learned that one had to be strong and go about one’s business without complaints. Such attitudes may make it difficult to admit psychological problems as well as difficulties with chronic pain more generally. The conflict between the logic of the researchers and that of the study participants was obvious throughout the interview results. Researchers need standardized questionnaires of intra- and inter-individual comparisons and a particular kind of objectivity.[49] This contrasts with the participants’ sense of personal experience. Questionnaires are developed to deduce complex experiences for statistical analysis. For our interviewees this reduction in fact meant that it was more difficult to answer the questionnaires and some of the interviewees felt frustrated by the inability of the questionnaires to capture their experiences.

The importance interviewees thought their particular experience had for the trial in fact contradicted the researchers’ efforts. To adequately present their experiences, interviewees tried to manipulate the questionnaires. In addition to adding notes, women in our sample marked different points of a scale to describe their experiences. While this was an optimizing strategy for the women, researchers consider

such items as “missing data” or “unscorable data.” Thus the effect was the opposite of what the women had intended and in fact undermines the validity of study results.

The conflict lies in a classic problem: questionnaires by default oversimplify complex experiences. The way these are reduced reflect interests of the researchers’ more than the patients.’[50] In the process of such reduction research subjects in fact become objects who shall produce data that is acceptable to the researchers.[51] What could be tactics to potentially untangle these two different logics that clash in clinical trial participation, specifically in completing questionnaires?

Warms et al. analysed the strategy of adding notes more closely and found that questionnaires were seen as a means of communication of study participants with study researchers.[52] This corresponds to our study findings of the importance the interviewees assigned to their trial participation. As such they assumed their individual experiences were of importance. When the communication tool is not perceived as a good one, study participants may react with frustration.[53] Again, this may have direct consequences on study results as it may lead to satisficing strategies in completing the questionnaires.[54] However, if one takes adding notes seriously as a participant’s wish to directly communicate with the researchers of the study, one can develop solutions that may not undermine the efforts of the researchers. Nesting qualitative interviews into clinical trials may facilitate such communications and help to respect participants’ perspectives and give them voice to communicate with researchers.[55, 56]

In our interview study only women participated. The RCT in which this study was nested had mostly female participants (95%). Thus, the female sample in the interview study is a reflection of the RCT population which in turn is a reflection of the larger proportion of females in this age group overall. Regardless, it is likely that men may not have been as eager to adequately depict their personal experiences in the questionnaires or would not have taken their responsibility as study participant as important as have the interviewees in the study. Similarly, since we only interviewed 20 of the 117 RCT

participants, it is conceivable that those who agreed to participate in the interview study took their role as study participant particularly serious. However, since we had created a ranking list with which we began recruitment and only six refused because they feared fraud, it would be surprising that we found those that were extraordinarily eager. Their seriousness about study participation could be a reflection of the values of a particular generation and age group. Finally, this group of women had experienced pain for a very long time and was open enough to try treatments that they did not know before. This may make this group of women especially thankful for providing options to treat their long-lasting pain. It is therefore possible that these findings are particular to an elderly population. Considering that there is a need for more medical research in elderly populations, it seems important to carefully evaluate the types of questionnaires used in such populations and to consider ways to explain the importance of standardised answers for clinical trials research.

While the strategies that were used by the women in this study in completing their questionnaires have been described in the literature, no study has yet described the relationship between perceptions research participants have about their role and the ways they complete their questionnaires. Overall, participants were frustrated with the questionnaires used, all of which are well validated questionnaires that are commonly used in research. To improve knowledge production in medicine it may be important to address these differential understandings of the ways in which clinical trial participants are of importance.

In this study we showed that a clear discrepancy existed between the logic of the researchers and the logic of RCT participants. Interviewees thought it was important for the trial that their actual experiences were understood by trial organizers. These were not transferrable by means of the provided questionnaires, so they added their experiences by hand to the questionnaires. However, the statistical analysis of RCT data needs this reduction of experience in order to produce results.[11] Study participants are a crucial component of clinical trials research as they are necessary for data production, but these data necessarily

are reductionist and aim to generate data that is comparable, which means numerical. Individual experiences need to be reworked to fit such criteria as comparability and objectivity. While participants in the RCT understood their importance for the study they developed strategies to convey their personal experiences that undermined the aims of the study. To improve data collection, increased effort may have to be invested in educating about the ways "experiences" need to be translated into comparative, standardised information to be able to use them for clinical trials research.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or non-profit sectors.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

ABBREVIATIONS

CES-D: Center for Epidemiological Studies Depression Scale

NPDS: Neck Pain and Disability Scale

QIBANE: Qigong and exercise therapy for elderly patients with chronic neck pain

RCT: randomized controlled trial

SF-36 Short-Form-36-Questionnaire

VAS: Visual Analogue Scale

REFERENCES

1. Breivik H, Collett B, Ventafridda V, et al. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *Eur J Pain* 2006;**10**(4):287-333.

2. Gagliese L, Melzack R. Chronic pain in elderly people. *Pain* 1997;**70**(1):3-14.

3. Miller LL, Talerico KA. Pain in older adults. *Annu Rev Nurs Res* 2002;**20**:63-88.

4. Rosier EM, Iadarola MJ, Coghill RC. Reproducibility of pain measurement and pain perception. *Pain* 2002;**98**(1-2):205-16.

5. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs* 2005;**14**(7):798-804.

6. Wood BM, Nicholas MK, Blyth F, et al. Assessing pain in older people with persistent pain: the NRS is valid but only provides part of the picture. *J Pain* 2010;**11**(12):1259-66.

7. Katz J, Melzack R. Measurement of pain. *Surg Clin North Am* 1999;**79**(2):231-52.

8. Nickel R, Raspe HH. [Chronic pain: epidemiology and health care utilization]. *Nervenarzt* 2001;**72**(12):897-906.

9. Broderick JE, Stone AA, Calvanese P, et al. Recalled pain ratings: a complex and poorly defined task. *J Pain* 2006;**7**(2):142-9.

10. Eccleston C, Crombez G, Aldrich S, et al. Attention and somatic awareness in chronic pain. *Pain* 1997;**72**(1-2):209-15.

11. Litcher-Kelly L, Martino SA, Broderick JE, et al. A systematic review of measures used to assess chronic musculoskeletal pain in clinical and randomized controlled clinical trials. *J Pain* 2007;**8**(12):906-13.

12. Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. *Acad Emerg Med* 2001;**8**(12):1153-7.

13. Miaskowski C, Bair M. Principles of Analgesic Use in the Treatment of Acute Pain and Cancer Pain: American Pain Society 2008.

14. Ferrell BA. Pain management. *Clin Geriatr Med* 2000;**16**(4):853-74.

15. Williams ACD, Davies HTO, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. *Pain* 2000;**85**(3):457-63.

16. Blyth FM, March LM, Brnabic AJ, et al. Chronic pain in Australia: a prevalence study. *Pain* 2001;**89**(2-3):127-34.

17. Zimmermann M. [Chronic pain. Epidemiology and management in Germany]. *Orthopade* 2004;**33**(5):508-14.

18. Lansinger B, Larsson E, Persson LC, et al. Qigong and exercise therapy in patients with long-term neck pain: a prospective randomized trial. *Spine (Phila Pa 1976)* 2007;**32**(22):2415-22. Epub 2007/12/20.

19. Gran JT. The epidemiology of chronic generalized musculoskeletal pain. *Best Practice & Research Clinical Rheumatology* 2003;**17**(4):547-61.

20. Gunzelmann T, Schumacher J, Brahler E. [The prevalence of pain in the elderly German population: results of population-based studies with the Giessen Subjective Complaints List (Giessener Beschwerdebogen GBB)]. *Schmerz* 2002;**16**(4):249-54.

21. Kohlmann T. [Musculoskeletal pain in the population]. *Schmerz* 2003;**17**(6):405-11.

22. Herr KA, Garand L. Assessment and measurement of pain in older adults. *Clin Geriatr Med* 2001;**17**(3):457-78, vi.

23. Schuler M, Hestermann M, Hauer K, et al. [Problems recognizing pain in geriatrics]. *Schmerz* 2002;**16**(3):171-8.

24. Herr K. Pain in the older adult: an imperative across all health care settings. *Pain Manag Nurs* 2010;**11**(2 Suppl):S1-10.

25. Strohbuecker B, Mayer H, Evers GC, et al. Pain prevalence in hospitalized patients in a German university teaching hospital. *J Pain Symptom Manage* 2005;**29**(5):498-506.

26. Gibson SJ, Lussier D. Prevalence and relevance of pain in older persons. *Pain Med* 2012;**13 Suppl 2**:S23-S6.

27. Kollbaum A. Die Gütekriterien der quantitativen Forschung. Flensburg: Universität Flensburg 2009.

28. Kaasalainen S, Crook J. A comparison of pain-assessment tools for use with elderly long-term-care residents. *Can J Nurs Res* 2003;**35**(4):58-71.

29. Von Trott P, Wiedemann AM, Ludtke R, et al. Qigong and exercise therapy for elderly patients with chronic neck pain (QIBANE): a randomized controlled study. *J Pain* 2009;**10**(5):501-8.

30. Wheeler AH, Goolkasian P, Baird AC, et al. Development of the Neck Pain and Disability Scale. Item analysis, face, and criterion-related validity. *Spine* 1999;**24**(13):1290-4.

31. Bullinger M. [Assessment of health related quality of life with the SF-36 Health Survey]. *Rehabilitation (Stuttgart)* 1996;**35**(3):XVII-XXVII.

32. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;**30**(6):473-83.

33. Hautzinger M. Allgemeine Depressions Skala. Weinheim: Beltz 1993.

34. Goolkasian P. Neck Pain and Disability Scale: a critical evaluation. *Expert Review of Pharmacoeconomics & Outcomes Research* 2003;**3**(4):379-82.

35. Salo P, Ylinen J, Kautiainen H, et al. Reliability and validity of the Finnish version of the neck disability index and the modified neck pain and disability scale. *Spine* 1976;**35**(5):552-6.

36. Wu S, Ma C, Mai M, et al. Translation and validation study of Chinese versions of the neck disability index and the neck pain and disability scale. *Spine* 1976;**35**(16):1575-9.

37. Wlodyka-Demaille S, Poiraudeau S, Catanzariti JF, et al. The ability to change of three questionnaires for neck pain. *Joint Bone Spine* 2004;**71**(4):317-26.

38. Goolkasian P, Wheeler AH, Gretz SS. The neck pain and disability scale: test-retest reliability and construct validity. *Clin J Pain* 2002;**18**(4):245-50.

39. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement* 1977;**1**(3):385-401.

40. Gunzelmann T. Ist Altern messbar? In: Oswald WD, Gatterer G, Fleischmann U, editors. Gerontopsychologie: Springer Vienna; 2008:59-77.

41. Campbell R, Quilty B, Dieppe P. Discrepancies between patients' assessments of outcome: qualitative study nested within a randomised controlled trial. *BMJ* 2003;**326**(7383):252-3. Epub 2003/02/01.

42. Daley AJ, Copeland RJ, Wright NP, et al. 'I can actually exercise if I want to; it isn't as hard as I thought': a qualitative study of the experiences and views of obese adolescents participating in an exercise therapy intervention. *Journal of health psychology* 2008;**13**(6):810-9. Epub 2008/08/14.

43. Manderson L, Bennett E, Andajani-Sutjahjo S. The social dynamics of the interview: age, class, and gender. *Qual Health Res* 2006;**16**(10):1317-34. Epub 2006/11/03.

44. Flick U. Qualitative Forschung - ein Handbuch. Reinbek bei Hamburg: rowohlts enzyklopädie 2007.

45. Mayring P. Einführung in die Qualitative Sozialforschung. Weinheim und Basel: Beltz Studium 2002.

46. Morse JM, A. Tashakkori, C. Teddlie. Principles of mixed methods and multimethod research design. In: Tashakkori A, C. Teddlie editor. Handbook of mixed methods in social and behavioral research. Thousand Oaks, CA: Sage 2003:189-208.

47. Krosnick JA. Survey research. *Annu Rev Psychol* 1999;**50**:537-67.

48. Holen JC, Hjermstad MJ, Loge JH, et al. Pain assessment tools: is the content appropriate for use in palliative care? *J Pain Symptom Manage* 2006;**32**(6):567-80.

49. Basler HD. [The Marburg questionnaire on habitual health findings--a study on patients with chronic pain]. *Schmerz* 1999;**13**(6):385-91.

50. Girtler R. Methoden der qualitativen Sozialforschung. Wien: Böhlau 1984.

51. Lamnek S. Gruppendiskussion. Theorie und Praxis. Weinheim 1998: Psychologie Verlag Union; 1998.

52. Warms CA, Marshall HM, Hoffman AJ, et al. There are a few things you did not ask about my pain: writing on the margins of a survey questionnaire. *Rehabil Nurs* 2005;**30**(6):248-56.

53. Slade SC, Molloy E, Keating JL. 'Listen to me, tell me': a qualitative study of partnership in care for people with non-specific chronic low back pain. *Clin Rehabil* 2009;**23**(3):270-80.

54. Meyer T, Deck R, Raspe H. [Validity of patient self report data in rehabilitation research: identifying circumstances of the patients' completion of questionnaires]. *Rehabilitation (Stuttgart)* 2006;**45**(2):118-27.

55. Verhoef MJ, Casebeer AL, Hilsden RJ. Assessing efficacy of complementary medicine: adding qualitative research methods to the "Gold Standard". *J Altern Complement Med* 2002;**8**(3):275-81.

56. Paterson C. Seeking the patient's perspective: a qualitative assessment of EuroQol, COOP-WONCA charts and MYMOP. *Qual Life Res* 2004;**13**(5):871-81.

Clinical Trial Participants' Experiences of Completing Questionnaires: A Qualitative Study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-004363.R1
Article Type:	Research
Date Submitted by the Author:	10-Jan-2014
Complete List of Authors:	Holmberg, Christine; Charité Universitätsmedizin Berlin, Institute for Social Medicine, Epidemiology, and Health Economics Karner, Julia; Charité Universitätsmedizin Berlin, Institute for Social Medicine, Epidemiology, and Health Economics Rappenecker, Julia; Charité Universitätsmedizin Berlin, Institute for Social Medicine, Epidemiology, and Health Economics Witt, Claudia; 1Charité University Medical Center, Institute for Social Medicine, Epidemiology and Health Economics; University of Maryland School of Medicine, Center for Integrative Medicine
Primary Subject Heading:	Qualitative research
Secondary Subject Heading:	Geriatric medicine, Communication, Complementary medicine
Keywords:	QUALITATIVE RESEARCH, PAIN MANAGEMENT, GERIATRIC MEDICINE

1. **TITLE:**

Clinical Trial Participants' Experiences of Completing Questionnaires: A Qualitative Study

2. **CORRESPONDING AUTHOR:**

Christine Holmberg

Institute for Social Medicine, Epidemiology, and Health Economics

Charité Universitätsmedizin Berlin

Luisenstrasse 57

Berlin 10117 Germany

Email: christine.holmberg@charite.de

Phone: +49 30 450 529 192

Fax: +49 30 450 529 917

3. **AUTHORS:**

Christine Holmberg,[a,b] Dr. phil., M.P.H., (christine.holmberg@charite.de), Julia Karner,[a] (julia.karner@charite.de), Dr. med., Julia Rappenecker,[a] (julia.rappenecker@hotmail.com), Dr. med., & Claudia M Witt,[a,c] (Claudia,witt@charite.de), Professor, Dr. med.

[a] Institute for Social Medicine, Epidemiology, and Health Economics, Charité Universitätsmedizin Berlin, Berlin, Germany

[b] Berlin School of Public Health, Charité Universitätsmedizin Berlin, Berlin, Germany

[c] Centre for Complementary and Integrative Medicine, University Hospital Zurich, Zurich, Switzerland

4. **KEYWORDS**

validation, study instruments, clinical trial participants, elderly, experience, quantification

5. **WORD COUNT: 5,116 (including abstract)**

ABSTRACT

Objectives:

To improve clinical study developments for elderly populations, we aim to understand how they transfer their experiences into validated, standardised self-completed study measurement instruments. We analysed how women (mean 78 ± 8 years of age) participating in an RCT cognised study instruments used to evaluate outcomes of the intervention.

Setting:

The interview study was nested in an RCT on chronic neck pain (Trial registration: ISRCTN77108101807) using common measurement instruments situated in an elderly community in Berlin, Germany comprised of units for independent and assisted-living options.

Participants:

The sample (n=20 women) was selected from the RCT sample (n=117, 95% women, mean age 76 (SD±8) years). Interview participants were selected using a purposive sampling list based on the RCT outcomes . Outcomes:

We asked participants about their experiences completing the RCT questionnaires. Interviews were analysed thematically, then compared to the questionnaires.

Results:

Interviewees had difficulties translating complex experiences into a single value on a scale and understanding the relationship of the questionnaires to study aims. Interviewees considered important for the trial that their actual experiences were understood by trial organisers. This information was not transferrable by means of the questionnaires. To rectify these difficulties, interviewees used strategies such as adding notes, adding response categories, or skipping an item.

Conclusion:

Elderly interview participants understood the importance of completing questionnaires for trial success. This led to strategies of completing the questionnaires that resulted in "missing" or ambiguous data. To improve data collection in elderly populations, educational materials addressing the differential logics should be developed and tested. Pilot testing validated instruments using cognitive interviews may be particularly important in such populations. Finally, when the target of an

intervention is subjective experience it seems important to create a methodby which participants can convey their personal experiences. These could be nested qualitative studies.

STRENGTHS AND LIMITATIONS OF THE STUDY

• This qualitative study gives insight into how elderly women think about and fill out validated study instruments.

• Interviewed women used satisficing strategies to complete questionnaires and made notes to convey their experiences to study personnel to ensure that "good" information was collected in the study.

• This differential logic led to strategies of completing questionnaires that produced missing data.

• Increasing elderly participants' understanding of research improves data collection.

• Data collection was conducted by clinical research staff. This may have influenced participants' ease to be honest and critical of their experiences with the questionnaires. In addition, findings should be tested in other elderly study populations.

INTRODUCTION

There are many factors that are crucial to the success of clinical trials, including validated study instruments. An adequate assessment of the study endpoint is a crucial aspect of clinical trials; for this validated questionnaires are considered one assessment tool for this purpose. The utilised instruments should be able to measure the same constructs consistently and accurately across individuals. There are some well-known questionnaire completion strategies such as marking the midpoint of a scale that prevent an accurate assessment of outcomes. Much effort has been devoted to the design of study instruments to discourage such behaviour.

The gold standard to assess subjective study endpoints are valid and reliable instruments. Validity corresponds to the question of how well an instrument measures what it intends to measure, such as pain intensity.[1] Reliability is established through tests and retests and validity through the comparability of a scale with other scales.[2 3] This means that if an instrument is used repeatedly and achieves the same results throughout or gives similar results to an instrument that has already been validated then its results are considered valid and reliable. For fluctuating, subjective experiences, such as pain, reliability and validity of scales only depicts part of the picture. [4-7] The experience of pain is influenced by context, meaning, emotional aspects, expectations, attitudes and beliefs associated with pain.[5] These aspects make it difficult to know what dimensions pain scales capture.Indeed while commonly used one-dimensional pain rating scales, such as the reliable Visual Analogue Scale (VAS)[8] are considered the gold standard for pain assessment,[9] have been validated in various populations, including elderly populations,[10-16] and are more often used in clinical practice and research[17] it remains unclear what the meaning of the information that such one-dimensional pain scales deliver represents.[18 19] Thus, diagnosing chronic pain poses problems to researchers and clinicians, despite existing validated instruments.[20-22]

As the example of one-dimensional pain scales show, adequate results for commonly used performance criteria such as validity and reliability do not necessarily suggest that they suffice for depicting complex subjective experiences.

In an RCT that compared the effects of Qigong and exercise therapy on neck pain in the elderly, no effect on pain intensity could be detected.[23] Three groups were compared: a Qigong group, an exercise therapy group, and a waiting list group. No difference between groups was found for the primary (VAS) and the secondary endpoints (Neck, Pain and Disability Scale based on a common depression, health-related quality of life, sleep quality, and satisfaction with the therapies). However, almost all study participants indicated that they would recommend the therapy to others and some even chose to continue the interventions at their own expenditure.[23] Thus, we were interested to understand how participants transferred their observations and experiences into the study measurement instruments.

The analysis aimed to understand how women (mean 78 ± 8 years of age) who participated in an RCT cognised the study instruments that were used to evaluate the primary and secondary endpoint outcomes of the intervention.

METHODS

Study Design

We conducted a qualitative study nested within an RCT to better understand the RCT results.[23] The trial was conducted by the Institute for Social Medicine, Epidemiology, and Health Economics at the Charité Universitätsmedizin Berlin and received ethics clearance by the appropriate ethics review board (EA1/265/05). Participants gave written and oral consent to participate in the RCT. Participants selected for the interview were invited by phone to participate in an interview on their experiences with the RCT. The interviews took place at the participants' homes and they were asked to provide additional oral consent for a home visit. The consent process was documented in the case report forms. The RCT included 117 patients with chronic neck pain that were randomised to a Qigong group, an exercise therapy group, or to a waiting list group. At three different time points, all three groups completed four validated questionnaires: the VAS, the Neck Pain and Disability Scale (NPDS),[24] the Short-Form.36-Questionnaire (SF-36),[25 26] and a common depression scale (ADS).[27] The NPDS is a specific evaluation instrument for neck pain that has shown to be valid and

reliable to measure neck pain[28-30] and to detect clinically relevant changes in neck pain.[31] It consists of 20 items that assess intensity of pain using neck problems as well as emotional and cognitive influences on work and everyday life[32]. The ADS assesses length and adverse effects of depressive symptoms, bodily problems and negative thought patterns. It is the German version of the Center for Epidemiological Studies Depression Scale (CES-D).[33] This instrument is recommended for use with chronic pain patients.[34] These instruments are the standard tools for these diagnoses. However, they are not satisfactorily validated for the age group under study.[23]

We developed a semi-structured interview guide that included questions related to the intervention and study instruments, more specifically asking about difficulties the patients may have had in completing the questionnaires and what was important for them in their experiences related to the study interventions. Prior to conducting the interviews, the interview guide was piloted in mock interviews with older patients with neck pain to ensure that the questions functioned well and the information was received as intended by the study aims.

Recruitment

In order to achieve a diverse selection of interview participants from the quantitative study (QIBANE) for the interview study, sampling was based on the results of the primary endpoint of the study. We wanted to ensure that the interview sample reflected the entire range of responses to the primary endpoint, which was decrease in neck pain as measured by the Visual Analogue Scale (VAS).[35] In addition, secondary endpoints such as the Neck Pain and Disability Scale (NPDS)[36] and the quality of life questionnaire SF-36 [37] were considered as secondary criteria for sample diversity. Thus, we created different groups of QIBANE participants: one group comprised of QIBANE participants who had indicated an improvement of symptoms between baseline and follow-up assessments, one group who had showed a worsening in symptoms, and a group of those that had no change between baseline assessment and three month follow-up. In each group, a ranking was established that started with the individuals with the largest differences between both assessment points. Once the rankings were established, participants were called until ten participants from the Qigong group and ten participants from the exercise therapy group agreed to a qualitative interview. Interviewee recruiters who called

participants had previously conducted the RCT and were known to participants. Participants in the RCT were mostly female (95%) which led to a list of potential interview participants that was predominantly female. Recruitment ended after the first ten RCT participants from the Qigong and another ten from the exercise therapy group had agreed to participate. A sample size of twenty participants was chosen based on the experiences of other qualitative studies that were nested in RCTs.[38 39]

Data collection

Interviews were conducted in the homes of the participants to ensure that participants felt comfortable and were willing to speak openly.[40] Interviewers had previously organised the RCT and were well-known to the interviewees. Interviews were conducted at the home of the interviewees to accommodate study participants and to create a relaxing atmosphere for the interviewee.[40] To help their memory, interview participants received blank sample questionnaires similar to the ones they had filled out during their RCT participation. While an interview guide was prepared for the interview, it was used in a flexible manner to allow for discussion that was important to the interviewees.[41 42] After each interview, the interviewer completed a standard protocol developed by Miles and Hubermann[43] to capture the atmosphere, setting and main themes of the interview. Interviews were digitally recorded and transcribed. The text documents were then entered into software programme ATLAS.ti for coding and analysis.

Data Analysis

Because interviewers were not involved in data analysis, the interview protocols provided the contextual information for the research team to situate the interview, its dynamics and content. Analysis of the study was multi–layered. As a first step, the qualitative interview materials were read by all researchers and analysed independently by JK and JR using content analysis according to Mayring.[41 42 44] This allowed focusing the analysis on the interview passages in which the questionnaires were discussed. The coding scheme was developed based on the emerging themes from the interview material by two of the authors (JK and JR) and then refined by the research team (all authors). In addition, coding and results were regularly presented and discussed in a qualitative

working group. The goal of the presentation to the working group was to ensure that materials and results were consistent with each other and to broaden the perspectives on the materials and ensure intersubjectivity of results. After analysis of interviews, we compared the quantitative questionnaires that had been completed by the interviewees in the RCT with interview results to identify strategies of how study instruments were completed.

RESULTS

Sample Description

Of those who were called and invited to participate in the interview study six declined a home visit due to fear of fraud. A short time prior to the recruitment for the interview study there had been some robberies in the senior residency and there was heightened awareness with regards to possible scam calls. The remaining twenty people agreed to participate in the interviews. Table 1 shows the changes the interviewees had indicated on the validated scales during the RCT. Eleven of the interviewees indicated a wish to continue the therapy even though they had not experienced an improvement of pain according to the validated instruments.

Table 1: Changes in measurements between baseline and primary endpoint of the interviewees.

Questionnaire[a]	**Improvement** (number of patients)	**Worsening** (number of patients)	**Missing Data** (number of patients)
VAS[b]	9	11	0
NPDS[c]	13	6	1
SF-36[d]	11	8	1
ADS[e]	5	8	7

[a]One participant had no change in the NPDS

[b]**VAS**: Visual Analogue Scale, [c]**NPDS**: Neck Pain and Disability Scale, [d]**SF-36** (mcs = mental component score: Mental component summary scale of the Short-Form-36-Questionnaire), [e]**ADS**: Common Depression Scale

All interviewees were female with an average age of seventy-six years of age. They had an age range of 67-85 years. On average they had experienced pain for fifteen years. All interviewees lived in residencies for seniors in Berlin.

Experiences completing the questionnaires

Many of the interviewees were dissatisfied either with the questionnaires and scales that they had to complete or the strategies they used to complete them. They complained about the difficulties of expressing complex experiences in the standardised terms the questionnaire asked of them.

> "Questionnaires are always terrible because you never can express by checking a box what one wants to say." [QG2/241]

> "If I make this movement, it hurts here. If I make that movement, it hurts there. Now the pain is gone. Now I look at you and I don't experience any pain. Now you tell me, do I have pain or do I not have pain? You tell me!" [QG2/318]

Some women were also concerned about the type of questions that were asked of them; questions related to their mental state as asked on the ADS and partially in the NPDS were especially disconcerting to some interviewees. Some interviewees were concerned that study staff may not adequately interpret their answers in the questionnaires because they were not able to precisely express on them how they felt.

> "In these questions one often has potential answers that partially fit and partially do not fit, so that one would say, 'yes, that is how it is, but….' (…) and since there is no possibility for the opposite, the whole answer isn't right." [QG10/011]

None of the interviewees felt that their experiences with pain or with living as an elderly person could be adequately described based on responses to the questionnaires that were administered to them. Particularly translating complex experiences into a single response on a scale was a challenge for the women. Participants used different strategies to deal with these problems when they completed the

scales. These were mainly additional notes, placing the mark in the middle of a scale, adding answer categories, or skipping an item. The women used these strategies because they felt that the scales could not capture their individual experiences. At the same time at least some felt indebted to the study since it gave them free exercise classes and they wanted to attend to the questionnaires in the best possible manner. Thus they added to the questionnaires the information they found pertinent.

Specificing standardised answers

Adding notes

Adding notes was a common strategy amongst the interviewees. Of the interviewees, 15 added information on an item to clarify what the value on the scale they marked signified. For example, one participant added to her answer for the item, "frequency of physical activity," the time frame, "*30 or 60 minutes!"* and the circumstances of the exercise, "*with partner or by myself,*" [PN7]. The same participant added to her answer to the item, "frequency of falls," "*in snow*." She had indicated that she had fallen once. Lastly on the NPDS the patient wanted to specify her pain and added "*in the lumbar spine and in the knees*." Another participant added a note to the value she selected on the VAS, "*I exercise daily. This is the only way I can remain relatively painless*," [PN6].

Others added handwritten notes to the response options instead of selecting a response on the scale. For example, one participant [QG5] added verbal signifiers to the scale on the NPDS such as "*seldom*," "*satisfied*," or "*little*." In the interviews, this particular participant complained about the questionnaires. Another participant [QG6] specified one question in the NPDS in the interview. Instead of putting a mark next to the question "does the pain hinder you with activities such as eating, dressing, or hygiene?" the participant responded by writing "*dressing*."

Similarly, where items asked for specific time frames, participants sometimes chose to change the time frame in order to meaningfully answer the questions. They noted on the side the time frame they referred to in the answer. For example, for a question that asked for a judgement of the last three months, one respondent wrote, *"[t]his has been in the last six months*," [QG10]. The theme addressed in the question seemed more important to the interviewees than the requested time frame.

Selecting parts of an item

Another strategy to respond to the questionnaire and specify general questions was to underline parts of a question to highlight what exactly the answer referred to. For example, one participant [PN7] underlined "kneeling" in an item of the SF-36 that stated "to bend forward, kneeling." Another such example comes from an either/or question in the NPDS. Two of the participants [PN1] [QG4] marked one of the two given possibilities in the item, "How difficult is it for you to look up or down?" Underlining was also used in questions that required a response along a scale. Several of the interviewees simply underlined one of the top or bottom values on the scale instead of marking a point along the scale.

Being a study participant

Some of the women had a clear understanding of the reciprocal relationship with the staff of the RCT; the women received interventions in exchange for completing the questionnaires. However, this required that they seriously considered their responsibility and wanted to complete the questionnaire adequately and accurately. Further, the questionnaires used in the study left some women feeling uneasy and unhappy with their contribution.

> "I was actually glad when I was done, just like school work that I had to do and I did very thoroughly. But I was not satisfied with my work and also not with the questions! So, I wasn't– but I have experienced such feelings with other questionnaires before." [QG2/265]

> "I hope I have filled out everything correctly. I do not know if I filled them out correctly." [PN9/163]

One woman called a family member and her family physician to assist her in completing the questionnaire in order to ensure the correctness of the questionnaires.

> "I don't remember for which question that was. I really did not know what to do with that question. I did not want to do anything wrong, so I called my daughter. She is a teacher and she also really had to think about it. But I cannot tell you which question that was at the time. I don't know. But the question was phrased very strange." [QG8/207]

Seriously considering their role as study participant was a common theme in the interviews and was the main reason why interviewees were dissatisfied with the assessment tools used in the RCT. The interviewed women assumed that their precise and exact experiences were of importance to the clinical trial staff and they were very concerned that the staff could not interpret their marks correctly on the assessments. The women made clear in the qualitative interviews that they preferred such an assessment much more than the questionnaires because the interviews enabled them to correctly state their experiences.

> "One just could not answer that question clearly. I don't know. I basically followed my feelings– but did you understand my answer? What you really get out of my answer is the question […]. So I had the feeling after I filled out the questionnaire that you cannot learn anything from those answers. I guess I would have to say: I would not trust those questionnaires. But those are your main interest, aren't they?!" [QG6/028]

> "The questions [in the questionnaire] do not make sense. I would have thought it better if you would have, just like you are doing now, asked the people directly." [QG4/250]

DISCUSSION

The interviewees in this study considered their role as a study participant important and perceived it as their responsibility to answer the questionnaires as accurately as they could to depict their experiences with chronic pain. However, they reported that this task was not easy for them. They had

difficulties making nonspecific statements about specific experiences and many thought that their experiences could not be depicted in the questionnaires; many also feared that their answers could be misunderstood. Several strategies were used by respondents to deal with the problem, such as adding notes, marking particular parts of a question, or leaving an item open. Some also asked others to help them complete the questionnaires correctly. Strategies such as adding notes have been called "optimising" strategies.[45] In addition, leaving an item blank or putting the mark in the middle of a scale are called "satisficing" strategies, suggesting that questions are answered cursorily.

Satisficing strategies are more common when study participants do not understand why certain questions are asked.[46] This is what the interviewees described especially in relation to the ADS. In the rationality of the researchers it was necessary to use the ADS for the RCT because an association between depression and chronic pain has been found before and therefore needed to be controlled for in the RCT. However, examining the association between mental state and pain was one that alienated research participants from the study. They did not consider pain and mental state as related to each other. The age group of the interviewed women may be one in which depression and other psychiatric diseases have a strong stigma associated with them. In addition, women in the study population may belong to a generation that had learned that one had to be strong and go about one's business without complaints. Such attitudes may make it difficult to admit psychological problems as well as difficulties with chronic pain more generally. The conflict between the logic of quantitative research and that of the study participants was obvious throughout the interview results. Medical research needs standardised questionnaires of intra- and inter-individual comparisons and a particular kind of objectivity.[47] It depends on de-contextualising personal experience in order to make the experience comparable and transferrable independent of time and place. This contrasts with the participants' sense of personal experience. Participants aimed to describe a precise and specific personal experience that aimed at being as accurate as possible. Questionnaires are developed to deduce complex experiences for statistical analysis. For our interviewees this reduction in fact meant that it was more difficult to answer the questionnaires and some of the interviewees felt frustrated by their inability to give an exact depiction of their experience through their answers to the questionnaires.

The extra effort interviewees went through to document their particular experiences contradicted the researchers' efforts to obtain quantitative data that is comparable across time and place. To adequately present their experiences, interviewees manipulated the questionnaires where they found it necessary for a more accurate description of their experiences. In addition to adding notes, women in our sample marked different points of a scale to describe their experiences. While this was an optimising strategy for the women, researchers consider such items as "missing data" or "unscorable data." Thus the effect was the opposite of what the women had intended and in fact these strategies could undermine the validity of study results. The interviewed women aimed at optimising their data to give a full picture of their experiences and in some instances produced data that was then not interpretable anymore from a statistical point of view, e.g. two marks on one scale. However, the range of using these strategies and the amount of missing data in the overall study (5% across all measurements and time points) is comparable to other RCTs. To minimize such faulty data, it is important to know how the elderly may understand the significance of their study participation in order to intervene and improve data collection in this age group.

The conflict lies in a classic problem: questionnaires by default oversimplify complex experiences. The way these are reduced reflect interests of the researchers more than the patients.[48] In the process of such reduction, research subjects in fact become objects that produce data that is acceptable to the researchers.[49] What are the implications to potentially untangle these two different logics that clash in clinical trial participation, specifically in completing questionnaires?

Warms et al. analysed the strategy of adding notes more closely and found that questionnaires were seen as a means that study participants communicated with study researchers.[50] This corresponds to our findings of the importance interviewees assigned to trial participation. As such they assumed their individual experiences were of importance. When the communication tool is not perceived as a good one, study participants may react with frustration.[51] Again, this may have direct consequences on study results as it may lead to satisficing strategies in completing the questionnaires.[52] However, these strategies are not a sign that study participants do not want to comply with study requirements.

On the contrary, our study participants developed these strategies precisely because they knew how important accurate data is for an RCT to be successful. Thus, understanding adding notes as a participant's wish to directly communicate with the researchers of the study and as participants' correct understanding of the importance of completing these questionnaires, helps develop solutions that may not undermine the efforts of the research. Nesting qualitative components such as interviews into clinical trials may facilitate such communications and help to respect participants' perspectives and give them voice to communicate with researchers. [53 54] In addition, measurements have been developed that assess participants' experiences of study participation.[55-57] If the study endpoint consists of a subjective experience that needs to be assessed in a standardized manner, it may be necessary to address that "accurate" has a particular meaning in research that may differ from how study participants consider "accurate" and explain the importance of sticking to provided instructions. It may be useful to develop such a standard leaflet explaining the need of standardization.

In our interview study only women participated. The RCT in which this study was nested had mostly female participants (95%). Thus, the female sample in the interview study is a reflection of the RCT population which in turn is a reflection of the larger proportion of females in this age group overall. Regardless, it is likely that men may not have been as eager to adequately depict their personal experiences in the questionnaires or would not have taken their responsibility as study participant as important as have the interviewees in the study. Similarly, since we only interviewed 20 of the 117 RCT participants, it is conceivable that those who agreed to participate in the interview study took their role as study participant seriously. However, since we had created a ranking list with which we began recruitment and only six refused because they feared fraud, it would be surprising that we found those that were extraordinarily eager. Their seriousness about study participation could be a reflection of the values of a particular generation and age group. Finally, this group of women had experienced pain for a very long time and was open enough to try treatments that they had not tried before, which may make this group of women especially thankful for providing options to treat their long-lasting pain. It is therefore possible that these findings are particular to an elderly population.

Considering that there is a need for more medical research in elderly populations, it seems important to carefully evaluate the types of questionnaires used in such populations and to consider ways to explain the importance of standardised answers for clinical trials research. The need to use cognitive interviewing to improve questionnaires has been voiced before[58] and the findings of this study underline the importance of such pilot testing before instruments are used in specific populations.

While the strategies that were used by the women in this study in completing their questionnaires have been described in the literature[45 46 52], no study has yet described the relationship between perceptions research participants have about their role and the ways they complete their questionnaires. Overall, participants were frustrated with the questionnaires used, all of which are standards for diagnosis that are commonly used in research. To improve knowledge production in medicine it may be important to address these differential understandings of the ways in which clinical trial participants are of importance.

In this study we showed that a clear discrepancy existed between the logic of quantitative research and the logic of RCT participants. Interviewees thought it was important for the trial that their actual experiences were understood by trial organisers. These were not transferrable by means of the provided questionnaires, so they added their experiences by hand to the questionnaires. However, the statistical analysis of RCT data needs this reduction of experience in order to produce results.[59] Study participants are a crucial component of clinical trials research as they are necessary for data production, but these data necessarily are reductionist and aim to generate data that is comparable and quantitative in nature. Individual experiences need to be reworked to fit such criteria as comparability and objectivity. Interviewees who had participated in QIBANE knew of their importance for the trial. Consequently they seriously considered their task of filling out questionnaires and tried to provide the best possible information. However, it was exactly this effort that in some cases led to strategies to convey their personal experience as best as possible, that undermined the aims of the study to get complete data. To improve data collection, increased effort may have to be invested in educating

about the ways “experiences” need to be translated into comparative, standardised information to be able to use them for clinical trials research and what “accurate” filling out of questionnaires means from a research perspective. Similarly, additional venues to the regularly used validated instruments that measure subjective and fluctuating experiences should be implemented to enable research participants to voice their experiences. These could include group discussions or interviews. Integrating qualitative and quantitative components such as implementation and process evaluation in addition to interviews can provide essential information that can improve research with this unique and growing population.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or non-profit sectors.

CONTRIBUTORSHIP STATEMENT

Christine Holmberg and Claudia Witt have developed the study design and have supervised the analysis of the study. Julia Rappenecker and Julia Karner have analyzed the materials. Christine Holmberg has written the manuscript and Claudia Witt, Julia Rappencker, and Julia Karner have given substantial input throughout the development and writing of the paper.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

DATA SHARING STATEMENT

The interview transcripts are available to show proof of the paper. However, they would only be available for legal purposes. They are confidential and can only be given access to in case of legal requirements.

ABBREVIATIONS

CES-D: Center for Epidemiological Studies Depression Scale

NPDS: Neck Pain and Disability Scale

QIBANE: Qigong and exercise therapy for elderly patients with chronic neck pain

RCT: Randomised controlled trial

SF-36 Short-Form-36-Questionnaire

VAS: Visual Analogue Scale

REFERENCES

1. Kollbaum A. Die Gütekriterien der quantitativen Forschung. Flensburg: Universität Flensburg, 2009.
2. Ferrell BA. Pain management. Clin Geriatr Med 2000;**16**(4):853-74
3. Kaasalainen S, Crook J. A comparison of pain-assessment tools for use with elderly long-term-care residents. Can J Nurs Res 2003;**35**(4):58-71
4. Rosier EM, Iadarola MJ, Coghill RC. Reproducibility of pain measurement and pain perception. Pain. 2002;**98**(1-2):205-16
5. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. Journal of clinical nursing 2005;**14**(7):798-804 doi: 10.1111/j.1365-2702.2005.01121.x[published Online First: Epub Date]|.
6. Wood BM, Nicholas MK, Blyth F, et al. Assessing pain in older people with persistent pain: the NRS is valid but only provides part of the picture. The journal of pain : official journal of the American Pain Society 2010;**11**(12):1259-66
7. Katz J, Melzack R. Measurement of pain. Surg Clin North Am 1999;**79**(2):231-52
8. Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. Acad Emerg Med 2001;**8**(12):1153-57
9. Miaskowski C, Bair M. Principles of Analgesic Use in the Treatment of Acute Pain and Cancer Pain: American Pain Society, 2008.
10. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. Pain 1986;**27**(1):117-26
11. E F. Pain Assessment for Older Adults. In: M W, ed. Try This: Best Practices in Nursing Care to Older Adults: Hartford Institute for Geriatric Nursing, 2000.

12. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. Journal of pain and symptom management 2011;**41**(6):1073-93 doi: 10.1016/j.jpainsymman.2010.08.016[published Online First: Epub Date]|.

13. Peters ML, Patijn J, Lame I. Pain assessment in younger and older pain patients: psychometric properties and patient preference of five commonly used measures of pain intensity. Pain medicine 2007;**8**(7):601-10 doi: 10.1111/j.1526-4637.2007.00311.x[published Online First: Epub Date]|.

14. Summers S. Evidence-based practice part 2: reliability and validity of selected acute pain instruments. Journal of perianesthesia nursing : official journal of the American Society of PeriAnesthesia Nurses / American Society of PeriAnesthesia Nurses 2001;**16**(1):35-40 doi: 10.1053/jpan.2001.20657[published Online First: Epub Date]|.

15. Herr KA, Spratt K, Mobily PR, et al. Pain intensity assessment in older adults: use of experimental pain to compare psychometric properties and usability of selected pain scales with younger adults. The Clinical journal of pain 2004;**20**(4):207-19

16. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. Pain 2011;**152**(10):2399-404 doi: 10.1016/j.pain.2011.07.005[published Online First: Epub Date]|.

17. Eccleston C, Crombez G, Aldrich S, et al. Attention and somatic awareness in chronic pain. Pain 1997;**72**(1-2):209-15

18. Broderick JE, Stone AA, Calvanese P, et al. Recalled pain ratings: a complex and poorly defined task. The journal of pain : official journal of the American Pain Society 2006;**7**(2):142-9 doi: 10.1016/j.jpain.2005.09.012[published Online First: Epub Date]|.

19. de CWAC, Davies HT, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. Pain 2000;**85**(3):457-63

20. Breivik H, Collett B, Ventafridda V, et al. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006;**10**(4):287-333

21. Gagliese L, Melzack R. Chronic pain in elderly people. Pain. 1997;**70**(1):3-14

22. Miller LL, Talerico KA. Pain in older adults. Annu Rev Nurs Res 2002;**20:63-88.**:63-88

23. Von Trott P, Wiedemann AM, Ludtke R, et al. Qigong and exercise therapy for elderly patients with chronic neck pain (QIBANE): a randomized controlled study. The journal of pain : official journal of the American Pain Society 2009;**10**(5):501-08

24. Wheeler AH, Goolkasian P, Baird AC, et al. Development of the Neck Pain and Disability Scale. Item analysis, face, and criterion-related validity. Spine. 1999;**24**(13):1290-94

25. Bullinger M. [Assessment of health related quality of life with the SF-36 Health Survey]. Rehabilitation (Stuttg) 1996;**35**(3):XVII-XXVII

26. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;**30**(6):473-83

27. Hautzinger M. Allgemeine Depressions Skala. Weinheim: Beltz, 1993.

28. Goolkasian P. Neck Pain and Disability Scale: a critical evaluation. Expert Rev Pharmacoecon Outcomes Res 2003;**3**(4):379-82

29. Salo P, Ylinen J, Kautiainen H, et al. Reliability and validity of the finnish version of the neck disability index and the modified neck pain and disability scale. Spine 1976;**2010 Mar 1;35**(5):552-56

30. Wu S, Ma C, Mai M, et al. Translation and validation study of Chinese versions of the neck disability index and the neck pain and disability scale. Spine 1976;**2010 Jul 15;35**(16):1575-79

31. Wlodyka-Demaille S, Poiraudeau S, Catanzariti JF, et al. The ability to change of three questionnaires for neck pain. Joint Bone Spine. 2004;**71**(4):317-26

32. Goolkasian P, Wheeler AH, Gretz SS. The neck pain and disability scale: test-retest reliability and construct validity. The Clinical journal of pain 2002;**18**(4):245-50

33. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Appl Psychol Meas 1977;**1**(3):385-401 doi: 10.1177/014662167700100306[published Online First: Epub Date]|.

34. Oswald WD, Gatterer G, Fleischmann U, et al. Ist Altern messbar? Gerontopsychologie: Springer Vienna, 2008:59-77.

35. Lara-Munoz C, De Leon SP, Feinstein AR, et al. Comparison of three rating scales for measuring subjective phenomena in clinical research. I. Use of experimentally controlled auditory stimuli. Archives of medical research 2004;**35**(1):43-8 doi: 10.1016/j.arcmed.2003.07.007[published Online First: Epub Date]|.

36. Scherer M, Blozik E, Himmel W, et al. Psychometric properties of a German version of the neck pain and disability scale. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society 2008;**17**(7):922-9 doi: 10.1007/s00586-008-0677-y[published Online First: Epub Date]|.

37. Morfeld M, Bullinger M, Nantke J, et al. [The version 2.0 of the SF-36 Health Survey: results of a population-representative study]. Sozial- und Praventivmedizin 2005;**50**(5):292-300

38. Campbell R, Quilty B, Dieppe P. Discrepancies between patients' assessments of outcome: qualitative study nested within a randomised controlled trial. BMJ 2003;**326**(7383):252-3

39. Daley AJ, Copeland RJ, Wright NP, et al. 'I can actually exercise if I want to; it isn't as hard as I thought': a qualitative study of the experiences and views of obese adolescents participating in an exercise therapy intervention. Journal of health psychology 2008;**13**(6):810-9 doi: 10.1177/1359105308093865[published Online First: Epub Date]|.

40. Manderson L, Bennett E, Andajani-Sutjahjo S. The social dynamics of the interview: age, class, and gender. Qual. Health Res. 2006;**16**(10):1317-34 doi: 10.1177/1049732306294512[published Online First: Epub Date]|.

41. Flick U. Qualitative Forschung - ein Handbuch. Reinbek bei Hamburg: rowohlts enzyklopädie, 2007.

42. Mayring P. *Einführung in die Qualitative Sozialforschung*. Weinheim und Basel: Beltz Studium, 2002.

43. Miles M HM. *Qualitative Data Analysis. An Expanded Sourcebook*. London: Sage Publications, 1994.

44. Morse JM, A. Tashakkori, C. Teddlie. Principles of mixed methods and multimethod research design. In: Tashakkori A, C. Teddlie ed. Handbook of mixed methods in social and behavioral research. Thousand Oaks, CA: Sage 2003:189-208.

45. Krosnick JA. Survey research. Annu Rev Psychol 1999;**50**:537-67

46. Holen JC, Hjermstad MJ, Loge JH, et al. Pain assessment tools: is the content appropriate for use in palliative care? Journal of pain and symptom management 2006;**32**(6):567-80

47. Basler HD. [The Marburg questionnaire on habitual health findings--a study on patients with chronic pain]. Schmerz. 1999;**13**(6):385-91

48. Girtler R. Methoden der qualitativen Sozialforschung. Wien: Böhlau, 1984.

49. Lamnek S. Gruppendiskussion. Theorie und Praxis. Weinheim 1998: Psychologie Verlag Union, 1998.

50. Warms CA, Marshall HM, Hoffman AJ, et al. There are a few things you did not ask about my pain: writing on the margins of a survey questionnaire. Rehabil Nurs 2005;**30**(6):248-56

51. Slade SC, Molloy E, Keating JL. 'Listen to me, tell me': a qualitative study of partnership in care for people with non-specific chronic low back pain. Clin Rehabil 2009;**23**(3):270-80

52. Meyer T, Deck R, Raspe H. [Validity of patient self report data in rehabilitation research: identifying circumstances of the patients' completion of questionnaires]. Rehabilitation (Stuttg). 2006;**45**(2):118-27

53. Verhoef MJ, Casebeer AL, Hilsden RJ. Assessing efficacy of complementary medicine: adding qualitative research methods to the "Gold Standard". J Altern Complement Med 2002;**8**(3):275-81

54. Paterson C. Seeking the patient's perspective: a qualitative assessment of EuroQol, COOP-WONCA charts and MYMOP. Qual Life Res 2004;**13**(5):871-81

55. Kost RG, Lee LM, Yessis J, et al. Assessing participant-centered outcomes to improve clinical research. The New England journal of medicine 2013;**369**(23):2179-81 doi: 10.1056/NEJMp1311461[published Online First: Epub Date]|.

56. Kost RG, Lee LM, Yessis J, et al. Assessing research participants' perceptions of their clinical research experiences. Clinical and translational science 2011;**4**(6):403-13 doi: 10.1111/j.1752-8062.2011.00349.x[published Online First: Epub Date]|.

57. Yessis JL, Kost RG, Lee LM et al. Development of a research participants' perception survey to improve clinical research. Clinical and translational science 2012;**5**(6):452-60 doi: 10.1111/j.1752-8062.2012.00443.x[published Online First: Epub Date]|.

58. Willis GB. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*: Sage Publlications, Inc, 2005.

59. Litcher-Kelly L, Martino SA, Broderick JE, et al. A systematic review of measures used to assess chronic musculoskeletal pain in clinical and randomized controlled clinical trials. The journal of pain : official journal of the American Pain Society 2007;**8**(12):906-13

1. TITLE:

Clinical Trial Participants' Experiences of Completing Questionnaires: A Qualitative Study

2. CORRESPONDING AUTHOR:

Christine Holmberg

Institute for Social Medicine, Epidemiology, and Health Economics

Charité Universitätsmedizin Berlin

Luisenstrasse 57

Berlin 10117 Germany

Email: christine.holmberg@charite.de

Phone: +49 30 450 529 192

Fax: +49 30 450 529 917

3. AUTHORS:

Christine Holmberg,[a,b] Dr. phil., M.P.H., (christine.holmberg@charite.de), Julia Karner,[a] (julia.karner@charite.de), Dr. med., Julia Rappenecker,[a] (julia.rappenecker@hotmail.com), Dr. med., & Claudia M Witt,[a,c] (Claudia,witt@charite.de), Professor, Dr. med.

[a] Institute for Social Medicine, Epidemiology, and Health Economics, Charité Universitätsmedizin Berlin, Berlin, Germany

[b] Berlin School of Public Health, Charité Universitätsmedizin Berlin, Berlin, Germany

[c] Centre for Complementary and Integrative Medicine, University Hospital Zurich, Zurich, Switzerland

4. KEYWORDS

validation, study instruments, clinical trial participants, elderly, experience, quantification

5. WORD COUNT: 5,116 (including abstract)

ABSTRACT

Objectives:

To improve clinical study developments for elderly populations, we aim to understand how they transfer their experiences into validated, standardised self-completed study measurement instruments. We analysed how women (mean 78 ± 8 years of age) participating in an RCT cognised study instruments used to evaluate outcomes of the intervention.

Setting:

The interview study was nested in an RCT on chronic neck pain (Trial registration: ISRCTN77108101807) using common measurement instruments situated in an elderly community in Berlin, Germany comprised of units for independent and assisted-living options.

Participants:

The sample (n=20 women) was selected from the RCT sample (n=117, 95% women, mean age 76 (SD±8) years). Interview participants were selected using a purposive sampling list based on the RCT outcomes . Outcomes:

We asked participants about their experiences completing the RCT questionnaires. Interviews were analysed thematically, then compared to the questionnaires.

Results:

Interviewees had difficulties translating complex experiences into a single value on a scale and understanding the relationship of the questionnaires to study aims. Interviewees considered important for the trial that their actual experiences were understood by trial organisers. This information was not transferrable by means of the questionnaires. To rectify these difficulties, interviewees used strategies such as adding notes, adding response categories, or skipping an item.

Conclusion:

Elderly interview participants understood the importance of completing questionnaires for trial success. This led to strategies of completing the questionnaires that resulted in "missing" or ambiguous data. To improve data collection in elderly populations, educational materials addressing the differential logics should be developed and tested. Pilot testing validated instruments using cognitive interviews may be particularly important in such populations. Finally, when the target of an

intervention is subjective experience it seems important to create a methodby which participants can convey their personal experiences. These could be nested qualitative studies.

STRENGTHS AND LIMITATIONS OF THE STUDY

• This qualitative study gives insight into how elderly women think about and fill out validated study instruments.

• Interviewed women used satisficing strategies to complete questionnaires and made notes to convey their experiences to study personnel to ensure that "good" information was collected in the study.

• This differential logic led to strategies of completing questionnaires that produced missing data.

• Increasing elderly participants' understanding of research improves data collection.

• Data collection was conducted by clinical research staff. This may have influenced participants' ease to be honest and critical of their experiences with the questionnaires. In addition, findings should be tested in other elderly study populations.

INTRODUCTION

There are many factors that are crucial to the success of clinical trials, including validated study instruments. An adequate assessment of the study endpoint is a crucial aspect of clinical trials; for this validated questionnaires are considered one assessment tool for this purpose. The utilised instruments should be able to measure the same constructs consistently and accurately across individuals. There are some well-known questionnaire completion strategies such as marking the midpoint of a scale that prevent an accurate assessment of outcomes. Much effort has been devoted to the design of study instruments to discourage such behaviour.

The gold standard to assess subjective study endpoints are valid and reliable instruments. Validity corresponds to the question of how well an instrument measures what it intends to measure, such as pain intensity.[1] Reliability is established through tests and retests and validity through the comparability of a scale with other scales.[2 3] This means that if an instrument is used repeatedly and achieves the same results throughout or gives similar results to an instrument that has already been validated then its results are considered valid and reliable. For fluctuating, subjective experiences, such as pain, reliability and validity of scales only depicts part of the picture. [4-7] The experience of pain is influenced by context, meaning, emotional aspects, expectations, attitudes and beliefs associated with pain.[5] These aspects make it difficult to know what dimensions pain scales capture.Indeed while commonly used one-dimensional pain rating scales, such as the reliable Visual Analogue Scale (VAS)[8] are considered the gold standard for pain assessment,[9] have been validated in various populations, including elderly populations,[10-16] and are more often used in clinical practice and research[17] it remains unclear what the meaning of the information that such one-dimensional pain scales deliver represents.[18 19] Thus, diagnosing chronic pain poses problems to researchers and clinicians, despite existing validated instruments.[20-22]

As the example of one-dimensional pain scales show, adequate results for commonly used performance criteria such as validity and reliability do not necessarily suggest that they suffice for depicting complex subjective experiences.

In an RCT that compared the effects of Qigong and exercise therapy on neck pain in the elderly, no effect on pain intensity could be detected.[23] Three groups were compared: a Qigong group, an exercise therapy group, and a waiting list group. No difference between groups was found for the primary (VAS) and the secondary endpoints (Neck, Pain and Disability Scale based on a common depression, health-related quality of life, sleep quality, and satisfaction with the therapies). However, almost all study participants indicated that they would recommend the therapy to others and some even chose to continue the interventions at their own expenditure.[23] Thus, we were interested to understand how participants transferred their observations and experiences into the study measurement instruments.

The analysis aimed to understand how women (mean 78 ± 8 years of age) who participated in an RCT cognised the study instruments that were used to evaluate the primary and secondary endpoint outcomes of the intervention.

METHODS

Study Design

We conducted a qualitative study nested within an RCT to better understand the RCT results.[23] The trial was conducted by the Institute for Social Medicine, Epidemiology, and Health Economics at the Charité Universitätsmedizin Berlin and received ethics clearance by the appropriate ethics review board (EA1/265/05). Participants gave written and oral consent to participate in the RCT. Participants selected for the interview were invited by phone to participate in an interview on their experiences with the RCT. The interviews took place at the participants' homes and they were asked to provide additional oral consent for a home visit. The consent process was documented in the case report forms. The RCT included 117 patients with chronic neck pain that were randomised to a Qigong group, an exercise therapy group, or to a waiting list group. At three different time points, all three groups completed four validated questionnaires: the VAS, the Neck Pain and Disability Scale (NPDS),[24] the Short-Form.36-Questionnaire (SF-36),[25 26] and a common depression scale (ADS).[27] The NPDS is a specific evaluation instrument for neck pain that has shown to be valid and

reliable to measure neck pain[28-30] and to detect clinically relevant changes in neck pain.[31] It consists of 20 items that assess intensity of pain using neck problems as well as emotional and cognitive influences on work and everyday life[32]. The ADS assesses length and adverse effects of depressive symptoms, bodily problems and negative thought patterns. It is the German version of the Center for Epidemiological Studies Depression Scale (CES-D).[33] This instrument is recommended for use with chronic pain patients.[34] These instruments are the standard tools for these diagnoses. However, they are not satisfactorily validated for the age group under study.[23]

We developed a semi-structured interview guide that included questions related to the intervention and study instruments, more specifically asking about difficulties the patients may have had in completing the questionnaires and what was important for them in their experiences related to the study interventions. Prior to conducting the interviews, the interview guide was piloted in mock interviews with older patients with neck pain to ensure that the questions functioned well and the information was received as intended by the study aims.

Recruitment

In order to achieve a diverse selection of interview participants from the quantitative study (QIBANE) for the interview study, sampling was based on the results of the primary endpoint of the study. We wanted to ensure that the interview sample reflected the entire range of responses to the primary endpoint, which was decrease in neck pain as measured by the Visual Analogue Scale (VAS).[35] In addition, secondary endpoints such as the Neck Pain and Disability Scale (NPDS)[36] and the quality of life questionnaire SF-36 [37] were considered as secondary criteria for sample diversity. Thus, we created different groups of QIBANE participants: one group comprised of QIBANE participants who had indicated an improvement of symptoms between baseline and follow-up assessments, one group who had showed a worsening in symptoms, and a group of those that had no change between baseline assessment and three month follow-up. In each group, a ranking was established that started with the individuals with the largest differences between both assessment points. Once the rankings were established, participants were called until ten participants from the Qigong group and ten participants from the exercise therapy group agreed to a qualitative interview. Interviewee recruiters who called

participants had previously conducted the RCT and were known to participants. Participants in the RCT were mostly female (95%) which led to a list of potential interview participants that was predominantly female. Recruitment ended after the first ten RCT participants from the Qigong and another ten from the exercise therapy group had agreed to participate. A sample size of twenty participants was chosen based on the experiences of other qualitative studies that were nested in RCTs.[38 39]

Data collection

Interviews were conducted in the homes of the participants to ensure that participants felt comfortable and were willing to speak openly.[40] Interviewers had previously organised the RCT and were well-known to the interviewees. Interviews were conducted at the home of the interviewees to accommodate study participants and to create a relaxing atmosphere for the interviewee.[40] To help their memory, interview participants received blank sample questionnaires similar to the ones they had filled out during their RCT participation. While an interview guide was prepared for the interview, it was used in a flexible manner to allow for discussion that was important to the interviewees.[41 42] After each interview, the interviewer completed a standard protocol developed by Miles and Hubermann[43] to capture the atmosphere, setting and main themes of the interview. Interviews were digitally recorded and transcribed. The text documents were then entered into software programme ATLAS.ti for coding and analysis.

Data Analysis

Because interviewers were not involved in data analysis, the interview protocols provided the contextual information for the research team to situate the interview, its dynamics and content. Analysis of the study was multi–layered. As a first step, the qualitative interview materials were read by all researchers and analysed independently by JK and JR using content analysis according to Mayring.[41 42 44] This allowed focusing the analysis on the interview passages in which the questionnaires were discussed. The coding scheme was developed based on the emerging themes from the interview material by two of the authors (JK and JR) and then refined by the research team (all authors). In addition, coding and results were regularly presented and discussed in a qualitative

working group. The goal of the presentation to the working group was to ensure that materials and results were consistent with each other and to broaden the perspectives on the materials and ensure intersubjectivity of results. After analysis of interviews, we compared the quantitative questionnaires that had been completed by the interviewees in the RCT with interview results to identify strategies of how study instruments were completed.

RESULTS

Sample Description

Of those who were called and invited to participate in the interview study six declined a home visit due to fear of fraud. A short time prior to the recruitment for the interview study there had been some robberies in the senior residency and there was heightened awareness with regards to possible scam calls. The remaining twenty people agreed to participate in the interviews. Table 1 shows the changes the interviewees had indicated on the validated scales during the RCT. Eleven of the interviewees indicated a wish to continue the therapy even though they had not experienced an improvement of pain according to the validated instruments.

Table 1: Changes in measurements between baseline and primary endpoint of the interviewees.

Questionnaire[a]	**Improvement** (number of patients)	**Worsening** (number of patients)	**Missing Data** (number of patients)
VAS[b]	9	11	0
NPDS[c]	13	6	1
SF-36[d]	11	8	1
ADS[e]	5	8	7

[a]One participant had no change in the NPDS

[b]**VAS**: Visual Analogue Scale, [c]**NPDS**: Neck Pain and Disability Scale, [d]**SF-36** (mcs = mental component score: Mental component summary scale of the Short-Form-36-Questionnaire), [e]**ADS**: Common Depression Scale

All interviewees were female with an average age of seventy-six years of age. They had an age range of 67-85 years. On average they had experienced pain for fifteen years. All interviewees lived in residencies for seniors in Berlin.

Experiences completing the questionnaires

Many of the interviewees were dissatisfied either with the questionnaires and scales that they had to complete or the strategies they used to complete them. They complained about the difficulties of expressing complex experiences in the standardised terms the questionnaire asked of them.

> "Questionnaires are always terrible because you never can express by checking a box what one wants to say." [QG2/241]

> "If I make this movement, it hurts here. If I make that movement, it hurts there. Now the pain is gone. Now I look at you and I don't experience any pain. Now you tell me, do I have pain or do I not have pain? You tell me!" [QG2/318]

Some women were also concerned about the type of questions that were asked of them; questions related to their mental state as asked on the ADS and partially in the NPDS were especially disconcerting to some interviewees. Some interviewees were concerned that study staff may not adequately interpret their answers in the questionnaires because they were not able to precisely express on them how they felt.

> "In these questions one often has potential answers that partially fit and partially do not fit, so that one would say, 'yes, that is how it is, but….' (…) and since there is no possibility for the opposite, the whole answer isn't right." [QG10/011]

None of the interviewees felt that their experiences with pain or with living as an elderly person could be adequately described based on responses to the questionnaires that were administered to them. Particularly translating complex experiences into a single response on a scale was a challenge for the women. Participants used different strategies to deal with these problems when they completed the

scales. These were mainly additional notes, placing the mark in the middle of a scale, adding answer categories, or skipping an item. The women used these strategies because they felt that the scales could not capture their individual experiences. At the same time at least some felt indebted to the study since it gave them free exercise classes and they wanted to attend to the questionnaires in the best possible manner. Thus they added to the questionnaires the information they found pertinent.

Specificing standardised answers

Adding notes

Adding notes was a common strategy amongst the interviewees. Of the interviewees, 15 added information on an item to clarify what the value on the scale they marked signified. For example, one participant added to her answer for the item, "frequency of physical activity," the time frame, "*30 or 60 minutes!"* and the circumstances of the exercise, "*with partner or by myself,*" [PN7]. The same participant added to her answer to the item, "frequency of falls," "*in snow*." She had indicated that she had fallen once. Lastly on the NPDS the patient wanted to specify her pain and added "*in the lumbar spine and in the knees*." Another participant added a note to the value she selected on the VAS, "*I exercise daily. This is the only way I can remain relatively painless*," [PN6].

Others added handwritten notes to the response options instead of selecting a response on the scale. For example, one participant [QG5] added verbal signifiers to the scale on the NPDS such as "*seldom*," "*satisfied*," or "*little*." In the interviews, this particular participant complained about the questionnaires. Another participant [QG6] specified one question in the NPDS in the interview. Instead of putting a mark next to the question "does the pain hinder you with activities such as eating, dressing, or hygiene?" the participant responded by writing "*dressing*."

Similarly, where items asked for specific time frames, participants sometimes chose to change the time frame in order to meaningfully answer the questions. They noted on the side the time frame they referred to in the answer. For example, for a question that asked for a judgement of the last three months, one respondent wrote, *"[t]his has been in the last six months*," [QG10]. The theme addressed in the question seemed more important to the interviewees than the requested time frame.

Selecting parts of an item

Another strategy to respond to the questionnaire and specify general questions was to underline parts of a question to highlight what exactly the answer referred to. For example, one participant [PN7] underlined "kneeling" in an item of the SF-36 that stated "to bend forward, kneeling." Another such example comes from an either/or question in the NPDS. Two of the participants [PN1] [QG4] marked one of the two given possibilities in the item, "How difficult is it for you to look up or down?" Underlining was also used in questions that required a response along a scale. Several of the interviewees simply underlined one of the top or bottom values on the scale instead of marking a point along the scale.

Being a study participant

Some of the women had a clear understanding of the reciprocal relationship with the staff of the RCT; the women received interventions in exchange for completing the questionnaires. However, this required that they seriously considered their responsibility and wanted to complete the questionnaire adequately and accurately. Further, the questionnaires used in the study left some women feeling uneasy and unhappy with their contribution.

> "I was actually glad when I was done, just like school work that I had to do and I did very thoroughly. But I was not satisfied with my work and also not with the questions! So, I wasn't– but I have experienced such feelings with other questionnaires before." [QG2/265]

> "I hope I have filled out everything correctly. I do not know if I filled them out correctly." [PN9/163]

One woman called a family member and her family physician to assist her in completing the questionnaire in order to ensure the correctness of the questionnaires.

> "I don't remember for which question that was. I really did not know what to do with that question. I did not want to do anything wrong, so I called my daughter. She is a teacher and she also really had to think about it. But I cannot tell you which question that was at the time. I don't know. But the question was phrased very strange." [QG8/207]

Seriously considering their role as study participant was a common theme in the interviews and was the main reason why interviewees were dissatisfied with the assessment tools used in the RCT. The interviewed women assumed that their precise and exact experiences were of importance to the clinical trial staff and they were very concerned that the staff could not interpret their marks correctly on the assessments. The women made clear in the qualitative interviews that they preferred such an assessment much more than the questionnaires because the interviews enabled them to correctly state their experiences.

> "One just could not answer that question clearly. I don't know. I basically followed my feelings– but did you understand my answer? What you really get out of my answer is the question […]. So I had the feeling after I filled out the questionnaire that you cannot learn anything from those answers. I guess I would have to say: I would not trust those questionnaires. But those are your main interest, aren't they?!" [QG6/028]

> "The questions [in the questionnaire] do not make sense. I would have thought it better if you would have, just like you are doing now, asked the people directly." [QG4/250]

DISCUSSION

The interviewees in this study considered their role as a study participant important and perceived it as their responsibility to answer the questionnaires as accurately as they could to depict their experiences with chronic pain. However, they reported that this task was not easy for them. They had

difficulties making nonspecific statements about specific experiences and many thought that their experiences could not be depicted in the questionnaires; many also feared that their answers could be misunderstood. Several strategies were used by respondents to deal with the problem, such as adding notes, marking particular parts of a question, or leaving an item open. Some also asked others to help them complete the questionnaires correctly. Strategies such as adding notes have been called "optimising" strategies.[45] In addition, leaving an item blank or putting the mark in the middle of a scale are called "satisficing" strategies, suggesting that questions are answered cursorily.

Satisficing strategies are more common when study participants do not understand why certain questions are asked.[46] This is what the interviewees described especially in relation to the ADS. In the rationality of the researchers it was necessary to use the ADS for the RCT because an association between depression and chronic pain has been found before and therefore needed to be controlled for in the RCT. However, examining the association between mental state and pain was one that alienated research participants from the study. They did not consider pain and mental state as related to each other. The age group of the interviewed women may be one in which depression and other psychiatric diseases have a strong stigma associated with them. In addition, women in the study population may belong to a generation that had learned that one had to be strong and go about one's business without complaints. Such attitudes may make it difficult to admit psychological problems as well as difficulties with chronic pain more generally. The conflict between the logic of quantitative research and that of the study participants was obvious throughout the interview results. Medical research needs standardised questionnaires of intra- and inter-individual comparisons and a particular kind of objectivity.[47] It depends on de-contextualising personal experience in order to make the experience comparable and transferrable independent of time and place. This contrasts with the participants' sense of personal experience. Participants aimed to describe a precise and specific personal experience that aimed at being as accurate as possible. Questionnaires are developed to deduce complex experiences for statistical analysis. For our interviewees this reduction in fact meant that it was more difficult to answer the questionnaires and some of the interviewees felt frustrated by their inability to give an exact depiction of their experience through their answers to the questionnaires.

The extra effort interviewees went through to document their particular experiences contradicted the researchers' efforts to obtain quantitative data that is comparable across time and place. To adequately present their experiences, interviewees manipulated the questionnaires where they found it necessary for a more accurate description of their experiences. In addition to adding notes, women in our sample marked different points of a scale to describe their experiences. While this was an optimising strategy for the women, researchers consider such items as "missing data" or "unscorable data." Thus the effect was the opposite of what the women had intended and in fact these strategies could undermine the validity of study results. The interviewed women aimed at optimising their data to give a full picture of their experiences and in some instances produced data that was then not interpretable anymore from a statistical point of view, e.g. two marks on one scale. However, the range of using these strategies and the amount of missing data in the overall study (5% across all measurements and time points) is comparable to other RCTs. To minimize such faulty data, it is important to know how the elderly may understand the significance of their study participation in order to intervene and improve data collection in this age group.

The conflict lies in a classic problem: questionnaires by default oversimplify complex experiences. The way these are reduced reflect interests of the researchers more than the patients.[48] In the process of such reduction, research subjects in fact become objects that produce data that is acceptable to the researchers.[49] What are the implications to potentially untangle these two different logics that clash in clinical trial participation, specifically in completing questionnaires?

Warms et al. analysed the strategy of adding notes more closely and found that questionnaires were seen as a means that study participants communicated with study researchers.[50] This corresponds to our findings of the importance interviewees assigned to trial participation. As such they assumed their individual experiences were of importance. When the communication tool is not perceived as a good one, study participants may react with frustration.[51] Again, this may have direct consequences on study results as it may lead to satisficing strategies in completing the questionnaires.[52] However, these strategies are not a sign that study participants do not want to comply with study requirements.

On the contrary, our study participants developed these strategies precisely because they knew how important accurate data is for an RCT to be successful. Thus, understanding adding notes as a participant's wish to directly communicate with the researchers of the study and as participants' correct understanding of the importance of completing these questionnaires, helps develop solutions that may not undermine the efforts of the research. Nesting qualitative components such as interviews into clinical trials may facilitate such communications and help to respect participants' perspectives and give them voice to communicate with researchers. [53 54] In addition, measurements have been developed that assess participants' experiences of study participation.[55-57] If the study endpoint consists of a subjective experience that needs to be assessed in a standardized manner, it may be necessary to address that "accurate" has a particular meaning in research that may differ from how study participants consider "accurate" and explain the importance of sticking to provided instructions. It may be useful to develop such a standard leaflet explaining the need of standardization.

In our interview study only women participated. The RCT in which this study was nested had mostly female participants (95%). Thus, the female sample in the interview study is a reflection of the RCT population which in turn is a reflection of the larger proportion of females in this age group overall. Regardless, it is likely that men may not have been as eager to adequately depict their personal experiences in the questionnaires or would not have taken their responsibility as study participant as important as have the interviewees in the study. Similarly, since we only interviewed 20 of the 117 RCT participants, it is conceivable that those who agreed to participate in the interview study took their role as study participant seriously. However, since we had created a ranking list with which we began recruitment and only six refused because they feared fraud, it would be surprising that we found those that were extraordinarily eager. Their seriousness about study participation could be a reflection of the values of a particular generation and age group. Finally, this group of women had experienced pain for a very long time and was open enough to try treatments that they had not tried before, which may make this group of women especially thankful for providing options to treat their long-lasting pain. It is therefore possible that these findings are particular to an elderly population.

Considering that there is a need for more medical research in elderly populations, it seems important to carefully evaluate the types of questionnaires used in such populations and to consider ways to explain the importance of standardised answers for clinical trials research. The need to use cognitive interviewing to improve questionnaires has been voiced before[58] and the findings of this study underline the importance of such pilot testing before instruments are used in specific populations.

While the strategies that were used by the women in this study in completing their questionnaires have been described in the literature[45 46 52], no study has yet described the relationship between perceptions research participants have about their role and the ways they complete their questionnaires. Overall, participants were frustrated with the questionnaires used, all of which are standards for diagnosis that are commonly used in research. To improve knowledge production in medicine it may be important to address these differential understandings of the ways in which clinical trial participants are of importance.

In this study we showed that a clear discrepancy existed between the logic of quantitative research and the logic of RCT participants. Interviewees thought it was important for the trial that their actual experiences were understood by trial organisers. These were not transferrable by means of the provided questionnaires, so they added their experiences by hand to the questionnaires. However, the statistical analysis of RCT data needs this reduction of experience in order to produce results.[59] Study participants are a crucial component of clinical trials research as they are necessary for data production, but these data necessarily are reductionist and aim to generate data that is comparable and quantitative in nature. Individual experiences need to be reworked to fit such criteria as comparability and objectivity. Interviewees who had participated in QIBANE knew of their importance for the trial. Consequently they seriously considered their task of filling out questionnaires and tried to provide the best possible information. However, it was exactly this effort that in some cases led to strategies to convey their personal experience as best as possible, that undermined the aims of the study to get complete data. To improve data collection, increased effort may have to be invested in educating

about the ways "experiences" need to be translated into comparative, standardised information to be able to use them for clinical trials research and what "accurate" filling out of questionnaires means from a research perspective. Similarly, additional venues to the regularly used validated instruments that measure subjective and fluctuating experiences should be implemented to enable research participants to voice their experiences. These could include group discussions or interviews. Integrating qualitative and quantitative components such as implementation and process evaluation in addition to interviews can provide essential information that can improve research with this unique and growing population.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or non-profit sectors.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

ABBREVIATIONS

CES-D: Center for Epidemiological Studies Depression Scale

NPDS: Neck Pain and Disability Scale

QIBANE: Qigong and exercise therapy for elderly patients with chronic neck pain

RCT: Randomised controlled trial

SF-36 Short-Form-36-Questionnaire

VAS: Visual Analogue Scale

REFERENCES

1. Kollbaum A. Die Gütekriterien der quantitativen Forschung. Flensburg: Universität Flensburg, 2009.

2. Ferrell BA. Pain management. Clin Geriatr Med 2000;**16**(4):853-74

3. Kaasalainen S, Crook J. A comparison of pain-assessment tools for use with elderly long-term-care residents. Can J Nurs Res 2003;**35**(4):58-71

4. Rosier EM, Iadarola MJ, Coghill RC. Reproducibility of pain measurement and pain perception. Pain. 2002;**98**(1-2):205-16

5. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. Journal of clinical nursing 2005;**14**(7):798-804 doi: 10.1111/j.1365-2702.2005.01121.x[published Online First: Epub Date]|.

6. Wood BM, Nicholas MK, Blyth F, Asghari A, Gibson S. Assessing pain in older people with persistent pain: the NRS is valid but only provides part of the picture. The journal of pain : official journal of the American Pain Society 2010;**11**(12):1259-66

7. Katz J, Melzack R. Measurement of pain. Surg Clin North Am 1999;**79**(2):231-52

8. Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. Acad Emerg Med 2001;**8**(12):1153-57

9. Miaskowski C, Bair M. Principles of Analgesic Use in the Treatment of Acute Pain and Cancer Pain: American Pain Society, 2008.

10. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. Pain 1986;**27**(1):117-26

11. E F. Pain Assessment for Older Adults. In: M W, ed. Try This: Best Practices in Nursing Care to Older Adults: Hartford Institute for Geriatric Nursing, 2000.

12. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. Journal of pain and symptom management 2011;**41**(6):1073-93 doi: 10.1016/j.jpainsymman.2010.08.016[published Online First: Epub Date]|.

13. Peters ML, Patijn J, Lame I. Pain assessment in younger and older pain patients: psychometric properties and patient preference of five commonly used measures of pain intensity. Pain medicine 2007;**8**(7):601-10 doi: 10.1111/j.1526-4637.2007.00311.x[published Online First: Epub Date]|.

14. Summers S. Evidence-based practice part 2: reliability and validity of selected acute pain instruments. Journal of perianesthesia nursing : official journal of the American Society of PeriAnesthesia Nurses / American Society of PeriAnesthesia Nurses 2001;**16**(1):35-40 doi: 10.1053/jpan.2001.20657[published Online First: Epub Date]|.

15. Herr KA, Spratt K, Mobily PR, Richardson G. Pain intensity assessment in older adults: use of experimental pain to compare psychometric properties and usability of selected pain scales with younger adults. The Clinical journal of pain 2004;**20**(4):207-19

16. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. Pain 2011;**152**(10):2399-404 doi: 10.1016/j.pain.2011.07.005[published Online First: Epub Date]|.

17. Eccleston C, Crombez G, Aldrich S, Stannard C. Attention and somatic awareness in chronic pain. Pain 1997;**72**(1-2):209-15

18. Broderick JE, Stone AA, Calvanese P, Schwartz JE, Turk DC. Recalled pain ratings: a complex and poorly defined task. The journal of pain : official journal of the American Pain Society 2006;**7**(2):142-9 doi: 10.1016/j.jpain.2005.09.012[published Online First: Epub Date]|.

19. de CWAC, Davies HT, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. Pain 2000;**85**(3):457-63

20. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006;**10**(4):287-333

21. Gagliese L, Melzack R. Chronic pain in elderly people. Pain. 1997;**70**(1):3-14

22. Miller LL, Talerico KA. Pain in older adults. Annu Rev Nurs Res 2002;**20:63-88.**:63-88

23. Von Trott P, Wiedemann AM, Ludtke R, Reishauer A, Willich SN, Witt CM. Qigong and exercise therapy for elderly patients with chronic neck pain (QIBANE): a randomized controlled study. The journal of pain : official journal of the American Pain Society 2009;**10**(5):501-08

24. Wheeler AH, Goolkasian P, Baird AC, Darden BV. Development of the Neck Pain and Disability Scale. Item analysis, face, and criterion-related validity. Spine. 1999;**24**(13):1290-94

25. Bullinger M. [Assessment of health related quality of life with the SF-36 Health Survey]. Rehabilitation (Stuttg) 1996;**35**(3):XVII-XXVII

26. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;**30**(6):473-83

27. Hautzinger M. Allgemeine Depressions Skala. Weinheim: Beltz, 1993.

28. Goolkasian P. Neck Pain and Disability Scale: a critical evaluation. Expert Rev Pharmacoecon Outcomes Res 2003;**3**(4):379-82

29. Salo P, Ylinen J, Kautiainen H, Arkela-Kautiainen M, Hakkinen A. Reliability and validity of the finnish version of the neck disability index and the modified neck pain and disability scale. Spine 1976;**2010 Mar 1;35**(5):552-56

30. Wu S, Ma C, Mai M, Li G. Translation and validation study of Chinese versions of the neck disability index and the neck pain and disability scale. Spine 1976;**2010 Jul 15;35**(16):1575-79

31. Wlodyka-Demaille S, Poiraudeau S, Catanzariti JF, Rannou F, Fermanian J, Revel M. The ability to change of three questionnaires for neck pain. Joint Bone Spine. 2004;**71**(4):317-26

32. Goolkasian P, Wheeler AH, Gretz SS. The neck pain and disability scale: test-retest reliability and construct validity. The Clinical journal of pain 2002;**18**(4):245-50

33. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Appl Psychol Meas 1977;**1**(3):385-401 doi: 10.1177/014662167700100306[published Online First: Epub Date]|.

34. Oswald WD, Gatterer G, Fleischmann U, Gunzelmann T. Ist Altern messbar? Gerontopsychologie: Springer Vienna, 2008:59-77.

35. Lara-Munoz C, De Leon SP, Feinstein AR, Puente A, Wells CK. Comparison of three rating scales for measuring subjective phenomena in clinical research. I. Use of experimentally controlled auditory stimuli. Archives of medical research 2004;**35**(1):43-8 doi: 10.1016/j.arcmed.2003.07.007[published Online First: Epub Date]|.

36. Scherer M, Blozik E, Himmel W, Laptinskaya D, Kochen MM, Herrmann-Lingen C. Psychometric properties of a German version of the neck pain and disability scale. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society 2008;**17**(7):922-9 doi: 10.1007/s00586-008-0677-y[published Online First: Epub Date]|.

37. Morfeld M, Bullinger M, Nantke J, Brahler E. [The version 2.0 of the SF-36 Health Survey: results of a population-representative study]. Sozial- und Praventivmedizin 2005;**50**(5):292-300

38. Campbell R, Quilty B, Dieppe P. Discrepancies between patients' assessments of outcome: qualitative study nested within a randomised controlled trial. BMJ 2003;**326**(7383):252-3

39. Daley AJ, Copeland RJ, Wright NP, Wales JK. 'I can actually exercise if I want to; it isn't as hard as I thought': a qualitative study of the experiences and views of obese adolescents participating in an exercise therapy intervention. Journal of health psychology 2008;**13**(6):810-9 doi: 10.1177/1359105308093865[published Online First: Epub Date]|.

40. Manderson L, Bennett E, Andajani-Sutjahjo S. The social dynamics of the interview: age, class, and gender. Qual. Health Res. 2006;**16**(10):1317-34 doi: 10.1177/1049732306294512[published Online First: Epub Date]|.

41. Flick U. Qualitative Forschung - ein Handbuch. Reinbek bei Hamburg: rowohlts enzyklopädie, 2007.

42. Mayring P. *Einführung in die Qualitative Sozialforschung*. Weinheim und Basel: Beltz Studium, 2002.

43. Miles M HM. *Qualitative Data Analysis. An Expanded Sourcebook*. London: Sage Publications, 1994.

44. Morse JM, A. Tashakkori, C. Teddlie. Principles of mixed methods and multimethod research design. In: Tashakkori A, C. Teddlie ed. Handbook of mixed methods in social and behavioral research. Thousand Oaks, CA: Sage 2003:189-208.

45. Krosnick JA. Survey research. Annu Rev Psychol 1999;**50**:537-67

46. Holen JC, Hjermstad MJ, Loge JH, et al. Pain assessment tools: is the content appropriate for use in palliative care? Journal of pain and symptom management 2006;**32**(6):567-80

47. Basler HD. [The Marburg questionnaire on habitual health findings--a study on patients with chronic pain]. Schmerz. 1999;**13**(6):385-91

48. Girtler R. Methoden der qualitativen Sozialforschung. Wien: Böhlau, 1984.

49. Lamnek S. Gruppendiskussion. Theorie und Praxis. Weinheim 1998: Psychologie Verlag Union, 1998.

50. Warms CA, Marshall HM, Hoffman AJ, Tyler EJ. There are a few things you did not ask about my pain: writing on the margins of a survey questionnaire. Rehabil Nurs 2005;**30**(6):248-56

51. Slade SC, Molloy E, Keating JL. 'Listen to me, tell me': a qualitative study of partnership in care for people with non-specific chronic low back pain. Clin Rehabil 2009;**23**(3):270-80

52. Meyer T, Deck R, Raspe H. [Validity of patient self report data in rehabilitation research: identifying circumstances of the patients' completion of questionnaires]. Rehabilitation (Stuttg). 2006;**45**(2):118-27

53. Verhoef MJ, Casebeer AL, Hilsden RJ. Assessing efficacy of complementary medicine: adding qualitative research methods to the "Gold Standard". J Altern Complement Med 2002;**8**(3):275-81

54. Paterson C. Seeking the patient's perspective: a qualitative assessment of EuroQol, COOP-WONCA charts and MYMOP. Qual Life Res 2004;**13**(5):871-81

55. Kost RG, Lee LM, Yessis J, Wesley RA, Henderson DK, Coller BS. Assessing participant-centered outcomes to improve clinical research. The New England journal of medicine 2013;**369**(23):2179-81 doi: 10.1056/NEJMp1311461[published Online First: Epub Date]|.

56. Kost RG, Lee LM, Yessis J, Coller BS, Henderson DK. Assessing research participants' perceptions of their clinical research experiences. Clinical and translational science 2011;**4**(6):403-13 doi: 10.1111/j.1752-8062.2011.00349.x[published Online First: Epub Date]|.

57. Yessis JL, Kost RG, Lee LM, Coller BS, Henderson DK. Development of a research participants' perception survey to improve clinical research. Clinical and translational science 2012;**5**(6):452-60 doi: 10.1111/j.1752-8062.2012.00443.x[published Online First: Epub Date]|.

58. Willis GB. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*: Sage Publlications, Inc, 2005.

59. Litcher-Kelly L, Martino SA, Broderick JE, Stone AA. A systematic review of measures used to assess chronic musculoskeletal pain in clinical and randomized controlled clinical trials. The journal of pain : official journal of the American Pain Society 2007;**8**(12):906-13

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
R…RELEVANCE OF STUDY QUESTION	
Is the research question interesting? *Yes.* Is the research question relevant to clinical practice, public health, or policy? *Yes, the research question is relevant to clinical practice and public health. The research question addresses one of the major pillars of clinical trials regarding problems associated with filling out questionnaires by the elderly, a population in which more health research is necessary. This is detailed in the "introduction" section.*	Research question explicitly stated Research question justified and linked to the existing knowledge base (empirical research, theory, policy)
A…APPROPRIATENESS OF QUALITATIVE METHOD	
Is qualitative methodology the best approach for the study aims? • Interviews: experience, perceptions, behaviour, practice, process *The chosen qualitative methodology, interviews, provided an effective way to gain deeper insight as to how participants translate (or not translate) experience into validated questionnaires. This is stated in the methods section.*	Study design described and justified i.e., why was a particular method (e.g., interviews) chosen?
T…TRANSPARENCY OF PROCEDURES *Sampling*	
Is the sampling strategy appropriate? *To create the random sample a data manager designated each QIBANE participant with a number and then chose random numbers using SPSS. Then a ranking list was created that randomly selected QIBANE participants from the group with improvement, with worsening, or with no change from baseline to follow-up assessment. This approach was necessary because we aimed to include the different experiences possible in QIBANE that we captured in the RCT also in the interview study. Thus, this was in line with a maximum variation sampling.*	
Are the participants selected the most appropriate to provide access to the type of knowledge sought by the study? *The goal of the study was to understand how elderly women transfer their experiences onto validated study instruments used in an RCT. Thus, the population of the RCT was the most appropriate population to sample from.*	Criteria for selecting the study sample justified and explained • *theoretical:* based on preconceived or

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
	emergent theory • *purposive:* diversity of opinion • *volunteer:* feasibility, hard-to-reach groups
Recruitment	
Was recruitment conducted using appropriate methods? *A member of the RCT research team called randomly chosen participants from the RCT and asked for permission to conduct an interview with them. The interview was conducted at the home of the interviewee. This location was chosen to make participation easy for the elderly and because it has been shown that interviews conducted at places in which interviewees feel at home help create an atmosphere that facilitates interviewing. We explain this in the "data collection" section of the paper.*	Details of how recruitment was conducted and by whom.
Is the sampling strategy appropriate?	
In the results section under "sample" we describe that 6 people who were asked to participate refused for fear of fraud.	Details of who chose not to participate and why-
Data collection	
Was collection of data systematic and comprehensive? *Yes. This is described under data collection.*	Method(s) outlined and examples given (e.g., interview questions)
Are characteristics of the study group and setting clear? *The QIBANE RCT, in which this study was nested, mostly consisted of female participants (95%). Thus, the female sample in the interview study is a reflection of the RCT population which in turn is a reflection of the larger proportion of females in this age group overall.* *The qualitative study group was elderly (mean age: 76 ± 8 years) and female with previous neck pain.* *The interview setting was at the participants' home.*	Study group and setting clearly described

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
Why and when was data collection stopped, and is this reasonable? *Data collection ended after 10 interviews from each intervention group (20 total) were conducted. The sample size of 20 was chosen based on other qualitative interview studies that are nested within RCTs (detailed in "study design" section).*	End of data collection justified and described
Role of researchers	
Are the researcher(s) appropriate? *All researchers were well trained in qualitative interviewing and qualitative content analysis. The research team consisted of MDs, epidemiologists, and an anthropologist. This range of disciplines allowed for a broad view onto the topic and onto the analysis of the materials. These different views that were brought to the materials and discussed in regular team meetings were able to highlight the different assumptions everyone brought to the materials and to ensure a rigorous analysis of materials.* How might they bias (good and bad) the conduct of the study and results? *The senior researcher of the project (CW) was the PI of the entire study. Similarly, those who called QIBANE participants to participate in the interview portion of the study had worked on the RCT. This could bias the interviewing in that interviewees would maybe not freely discuss problems they had with the RCT. Also the interviewer was an MD which may have led to participants (elderly women) be too respectful to present negative views or personal experiences. We carefully scrutinized the interview materials for such cues and found that interviewees indeed were not openly opposing the interviewer but did so clearly in small statements. All interviewees talked freely about their difficulties and about the nonsense they thought was asked in the questionnaires. Similarly, all interviews were listened to immediately after the interview took place by CH to pick up on any problems that may arise due to inadequate questioning and each interview was then discussed by CH and the interviewer. To detect additional biases the interviewer wrote an interview protocol after each interview in which such things were recorded as when the interviewer felt awkward asking a question or what dynamics developed during the interview. We talk about this in the method and in the discussion section of the paper.*	Do the researchers occupy dual roles (clinician and researcher)? Are the ethics of this discussed? Do the researcher(s) critically examine their own influence on the formulation of the research question, data collection, and interpretation?

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
Ethics	
Was informed consent sought and granted? *The RCT was approved by the appropriate ethics review board (EA1/265/05). Participants gave written and oral consent to the intervention study. The interview study was part of the intervention study and selected participants were invited by phone to participate in an interview on their experiences with the RCT. The interviews took place at their home and they were asked to provide additional oral consent for a home visit. The consent process was documented in the case report forms We detail this in the paper in the "Methods" section under "data collection" and "data analysis."*	Informed consent process explicitly and clearly detailed
Were participants' anonymity and confidentiality ensured? *All interview materials was pseudonomysed. The process is detailed in the informed consent and was discussed with interview participants prior to the interview.*	Anonymity and confidentiality discussed
Was approval from an appropriate ethics committee received? *Yes. The intervention study was approved by the ethics review board of the Charité Universitätsmedizin Berlin (EA1/265/05). This is stated in the Methods section under "study design"*	Ethics approval cited
S..SOUNDNESS OF INTERPRETIVE APPROACH *Analysis*	
Is the type of analysis appropriate for the type of study? • The analysis was conducted using content analysis and comparing filled out questionnaires with the results of the content analysis. Since the aim of the study was to identify how RCT participants fill out validated questionnaires this was the most appropriate method of analysis. Codes for the content analysis were developed inductively to capture all themes the interviewees brought up. Are the interpretations clearly presented and adequately supported by the evidence?	Analytic approach described in depth and justified *Indicators of quality:* Description of how themes were derived from the data (inductive or deductive) Evidence of alternative explanations being sought Analysis and presentation of negative or deviant cases
Are quotes used and are these appropriate and effective? *All themes that were detected through the analysis were presented in the results and are supported by a quote. Quotes*	Description of the basis on which quotes were chosen Semi-quantification

RATS CHECKLIST for the manuscript, ***The relationship between completing standardized questionnaires and perceptions of being a study participant: varying logics of study participants and researchers: A qualitative study***	**THIS SHOULD BE INCLUDED IN THE MANUSCRIPT**
were selected to reflect the identified theme (strategy to fill out questionnaire). All themes are semi-quantified to show how often it approximately appeared.	when appropriate Illumination of context and/or meaning, richly detailed
Was trustworthiness/reliability of the data and interpretations checked? *As a first step the qualitative interview materials were read and then analyzed independently by two researchers using content analysis.[1-3] This allowed focusing on the interview passages in which the questionnaires were discussed. The coding scheme was developed based on the interview material by two of the authors (JK and JR) and then refined by the research team (all authors). In addition, coding and results were regularly presented and discussed in a qualitative working group. The goal of the presentation in the working group was to ensure that materials and results were consistent with each other and to broaden the perspectives on the materials and ensure intersubjectivity of results. After analysis of interviews, we compared the quantitative questionnaires that had been completed by the interviewees in the RCT with* interview *results to identify strategies of how they were completed. This is detailed in the method section of the paper.*	Method of reliability check described and justified e.g., was an audit trail, triangulation, or member checking employed? Did an independent analyst review data and contest themes? How were disagreements resolved?
Discussion and presentation	
Are findings sufficiently grounded in a theoretical or conceptual framework? Is adequate account taken of previous knowledge and how the findings add? *Yes, both are addressed in the discussion of the paper.*	Findings presented with reference to existing theoretical and empirical literature, and how they contribute
Are the limitations thoughtfully considered? *Yes. We do this at the end of the discussion section.*	Strengths and limitations explicitly described and discussed

REFERENCES

1. Morse JM, A. Tashakkori, C. Teddlie. Principles of mixed methods and multimethod research design. In: Tashakkori A, C. Teddlie editor. Handbook of mixed methods in social and behavioral research. Thousand Oaks, CA: Sage 2003:189-208.
2. Flick U. Qualitative Forschung - ein Handbuch. Reinbek bei Hamburg: rowohlts enzyklopädie 2007.

3. Mayring P. Einführung in die Qualitative Sozialforschung. Weinheim und Basel: Beltz Studium 2002.