BMJ Open

How Novice, Skilled and Advanced Clinical Researchers Include Variables in a Case Report Form for Clinical Research:A Qualitative Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016760
Article Type:	Research
Date Submitted by the Author:	14-Mar-2017
Complete List of Authors:	Chu, Hongling; Peking University Third Hospital, Research Center of Clinical Epidemiology Zeng, Lin; Peking University Third Hospital, Research Center of Clinical Epidemiology Fetters, Micheal; University of Michigan Department of Family Medicine Li, Nan; Peking University Third Hospital, Research Center of Clinical Epidemiology Tao, Liyuan; Peking University Third Hospital, Research Center of Clinical Epidemiology Shi, Yanyan; Peking University Third Hospital, Research Center of Clinical Epidemiology Zhang, Hua; Peking University Third Hospital, Research Center of Clinical Epidemiology Wang, Xiaoxiao; Peking University Third Hospital, Research Center of Clinical Epidemiology Li, Fengwei; Beijing Aerospace General Hospital zhao, yiming; Peking University Third Hospital, Research Center of Clinical Epidemiology
Primary Subject Heading:	Qualitative research
Secondary Subject Heading:	Research methods
Keywords:	QUALITATIVE RESEARCH, clinical study, data collection, research design

How Novice, Skilled and Advanced Clinical Researchers Include Variables in a Case Report Form for Clinical Research: A Qualitative Study

Chu Hongling, Zeng Lin, Michael D Fetters, Li Nan, Tao Liyuan, Shi Yanyan, Zhang Hua, Wang Xiaoxiao, Li Fengwei, Zhao Yiming

Corresponding author: Zhao Yiming, Peking University Third Hospital, Research Center of Clinical Epidemiology, 49#, Huayuanbei Road, Haidian District Beijing, P.R.China 100191. yimingzhao115@163.com, +86-010-82266577, +86-010-82265732.

Chu Hongling, Zeng Lin, Li Nan, Tao Liyuan, Shi Yanyan, Zhang Hua, Wang Xiaoxiao, Zhao Yiming: Peking University Third Hospital, Research Center of Clinical Epidemiology, Beijing, China.

Michael D Fetters: University of Michigan Department of Family Medicine, Ann Arbor, MI, US.

Li Fengwei: Beijing Aerospace General Hospital, Beijing, China.

Abstract

Objectives: Despite varying degrees in research training, most academic clinicians are expected to conduct clinical research. The objective of this research was to understand how clinical researchers of different skill levels include variables in a case report form for their clinical research.

Setting: The setting for this research was a major academic institution in Beijing, China.

Participants: The target population was 17 clinical researchers with three levels of experience, namely, limited clinical research experience, clinicians with rich clinical

research experience and clinical research experts.

Methods: Using a qualitative approach, we conducted 13 individual interviews (face–to-face) and one group interview (n=4) with clinical researchers from June to September 2016. We used maximum variation sampling to identify 17 researchers with three levels of research experience: 8 clinicians with limited clinical research experience, 5 clinicians with rich clinical research experience, and 4 experts in clinical research. The researchers had diverse hospital-based medical specialties and or specialization in clinical research.

Results: Our analysis yields a typology of three processes developing a case report form (CRF) that varies according to research experience level. Novice clinician researchers often have an incomplete protocol or none at all, and conduct data collection and publication based on a general framework. Experienced clinician researchers include variables in the case report form based on previous experience with attention to including domains or items at risk for omission, and by eliminating unnecessary variables. Expert researchers consider comprehensively in advance data collection and implementation needs and plan accordingly.

Conclusion: These results illustrate increasing levels of sophistication in research planning that reflect increasing levels of selecting variables in case report form. These findings suggest that novice and intermediate-level researchers could benefit by emulating the comprehensive planning procedures such as those utilized by expert clinical researchers.

Keywords: qualitative research; clinical study; data collection, research design

Article summary

Strengths and limitations of this study

The study particularly informs how clinical researchers select variables in a CRF in their clinical research. In previous research, this question has been only marginally explored relative to the structure and surface, although it can play an important role in improving quality of clinical research. Furthermore, based on reports of clinical research experts, these findings illustrate an overarching and effective process for

determining variables for inclusion in a CRF. In addition, the PICO framework for thinking about the study represents novel thinking in research design raised by these experts. Additionally, the complementary ideas of "few to many" and "many to few" for refining the variables in CRF represents innovative thinking for CRF design.

A limitation of our study concerns the potential selection bias of our sample. Clinician participants are from a premier academic hospital in Beijing, and these participants might have greater knowledge, cognitive skills, and opportunities for research than clinicians in many other hospitals. We both acknowledge the potential for variations on the findings, while also acknowledging the results are genuine and representative of the participants' experiences. Despite variations that exist in practice, we believe the lessons learned based on this study are robust ideas.

These results illustrate increasing levels of sophistication in research planning that reflect increasing levels of selecting variables in case report form. These findings suggest that novice and intermediate-level researchers could benefit by emulating the comprehensive planning procedures such as those utilized by expert clinical researchers.

INTRODUCTION

Conducting clinical research is a multistage process. It can be divided into three stage: the first stage is top-level design, the second stage is protocol design and implementation, the third stage is conducting data analysis, interpretation and writing of the paper[1]. During the first two stages, researchers should consider what variables are important, and how to collect these data.

Case report form (CRF) is an instrument to structure and facilitate collection of data for clinical research[2]. Most CRFs are customized to collect data specific to a particular clinical study protocol. Case report form development represents a significant part of the clinical trial process and can impact study success. A well-designed CRF is required for database construction, data accuracy, data query/cleaning, CRF completion and statistical analysis.

In our institution, to facilitate greater clinical research efficiency, our lab became interested in how to develop an approach to build a CRF. Our interest extended to both the structure and variables of the CRF. The structure has been the topic of concern among some researchers[3-11]. In a published paper, it was emphasized that the design of CRF needs the cooperation and efforts of each member of the study group[7]. In another published related paper, it was noted that the CRF should keep privacy for participants, include a tracked page or modules, and some other things on forms[10]. However, the question that which variables should be selected in CRF has received less consideration but plays a crucial role in the quality of clinical research.

The goal of this qualitative research study was to explore how clinical researchers select variables in a CRF and was part of a larger mixed-methods project to develop an approach for clinical researchers to build systematic variables for clinical research. These findings could provide an approach for choosing variables in a CRF by clinical researchers. This can serve as a reference to other researchers, and lay a foundation for further inquiry into what variables to include in a CRF.

METHOD

Qualitative inquiry is an approach particularly useful when little is known about the phenomenon under study[12]. As little is known about how clinicians determine variables to include in a CRF we deemed a qualitative approach as useful.

Setting:

The setting for this research was a major academic institution in Beijing, China.

Study Population:

The target population was clinical researchers with three levels of experience, namely, limited clinical research experience, clinicians with rich clinical research experience and clinical research experts.

Data collection instrument:

We developed a semi-structured interview guide based on group discussion and a preliminary pilot study in 2 clinical researchers. The primary interview question generating data for this study was, “What process do you use to design a case report

form for clinical research?" Key probes were, "What are the difficulties and challenges encountered when you designed case report form for your project?" and "What is your previous experience joining a clinical research project?" The overall interview guide was designed for a parent study looking comprehensively at clinical researchers approach to data collection and thus had other questions in the interview guide.

Recruitment:

Individuals targeted for enrollment were contacted by email by a research assistant. We used maximum variation sampling to identify with different degrees of research experience and different medical specialties in the hospital and clinical research institute.

The inclusion criteria were:

a. Meet one of the following conditions: 1) Clinicians with limited clinical research experience who were defined as rarely participating in clinical research, a criterion operationalized as no experience to one experience designing a CRF for clinical research; 2) clinicians with rich clinical research experience defined as researchers with experience participating in several clinical research studies and experience designing three to nine data collection reforms or case report forms for clinical research. 3) Clinical research experts defined as researchers with experience in clinical research for five years or more, and participation directly in 10 or more clinical research projects.
b. Being able to express themselves with well-articulated stories and to deeply reflect on their stories;
c. Being willing to participate in the study.

Data collection

Two research assistants conducted interviews from June to September in 2016. The research assistants (RAs) conducted first conducted one group interview with clinicians who had limited experience in clinical research. As gathering busy

clinicians for a group interview proved difficult, the RAs changed to a face–to-face semi-structured in-depth interviews with nine clinicians and four clinical research experts,. The same questions were posed in the same order to all the participants, whether the interviews were performed individually or in the group interview. One question was added for the four clinical research experts, that is, "what did you encounter for when you directed other clinical researchers' project". The interviews lasted between (25-40) minutes and were conducted in a location that was quiet without interruption, either the participant's office or the interviewee's conference room. All interviews were audio-recorded and transcribed verbatim into Chinese.

Data analysis

The transcribed data were analyzed in Chinese using thematic analysis with an inductive approach. Two researchers (CH and Zl) independently coded and analyzed the transcripts by: selecting the units of analysis, making sense of the transcribed data, developing codes, categorizing the data and abstracting[13]. The analysis focused on text from the four primary questions noted above but also used related information from other interview guide questions and context-specific language probes. The analysts discussed and reached an agreement on the coding and categorization after reviewing one interview. The two researchers independently coded the remaining transcripts. Differences were minimal. All research team members agreed with the final results. After constructing the models, the team confirmed the findings by checking the results with the interviews transcript.

Below, we present the findings of the study by illustrating the three identified models and illustrative quotes from the interviews. The researchers who conducted the data collection and analysis translated the quotes selected for this article into English. Original quotes are available on request.

Ethical considerations

Each participant was informed about the study procedures and was free to withdraw from the research. All participants provided consent to participate in the

study including audiotaping and transcription of the interviews by providing written informed consent. Potentially identifying information was removed from each transcript and each interviewee was assigned a unique identification number to protect his/her anonymity. The institutional review board (IRB) at Peking University Third Hospital approved the study.

RESULTS

The 17 clinical researchers had diverse background, those are otolaryngology, pharmacy, endocrinology, orthopedics, anesthesia, radiology, neurologist, cardiology, nephrology, Hematology, thoracic surgery, epidemiology and biostatistics, clinical epidemiology, clinical research data management, clinical research methodology. (Table1)

Table 1 Demographics of the participants

Participants level	Clinicians with limited clinical research experience	Clinicians with rich clinical research experience	Clinical research experts
Number	8	5	4
Gender			
Male (N)	2	2	3
Female (N)	6	3	1
Average age (mean± SD)	26.63±3.50	35.40±2.07	43.75±12.45
Department	Otolaryngology, Pharmacy, Endocrinology, Orthopedics, Anesthesia, Radiology, Neurologist, Cardiology.	Nephrology, Otolaryngology, Thoracic surgery, Hematology	Epidemiology Or Clinical research methodology
Experience in clinical research (projects)	Participating in clinical research work for 1-3 years	Being involved in 5-12 clinical research projects	Working in clinical research for 11-20 years
Data collection	4 semi-structured in-depth interviews, and 1 group interview for 4 clinicians	Semi-structured in-depth interviews	Semi-structured in-depth interviews

A typology of how clinical researchers of different skill levels include variables in

CRF for clinical research

Based on our analysis, we developed a typology of how clinical researchers of different skill levels select variables in a CRF for clinical research. These three models are illustrated as Figures 1, 2, and 3. The findings are supported by quoted comments from the research participants. The models are illustrated using a flow chart showing the different process for each of three levels of the participants, e.g., novice clinicians with limited clinical research experience, intermediate-level clinicians with rich clinical research experience, and clinical research experts.

a. The novice clinical researcher's approach to selecting variables in a CRF for clinical research.

Novice clinicians with limited clinical research experience described a multi-faceted approach to selecting variables for inclusion in their CRFs. When they planned the variables to include in their CRFs, most had no clearly defined comprehensive approach.

Finding the template from similar research and imitating it. Most novice clinical researchers mentioned that they would find a similar data collection form for reference and modify it according to the needs of their own study. As one endocrinologist noted, "First, I search on the Internet to find a template from similar research, then I modify the template based on my research". One otolaryngology clinician opined, "I ask for a data collection template from other experienced researchers in my department. I then imitate the template to make my own data collection template, and modify it in places according my research." An Orthopedics clinician reflected his opinion that most novice clinical researchers are doing like this, "I will imitate another data collection template, then copy what is applicable to my research, and modify those variables that are not applicable. I think most of us are doing like this, because it is easier for us."

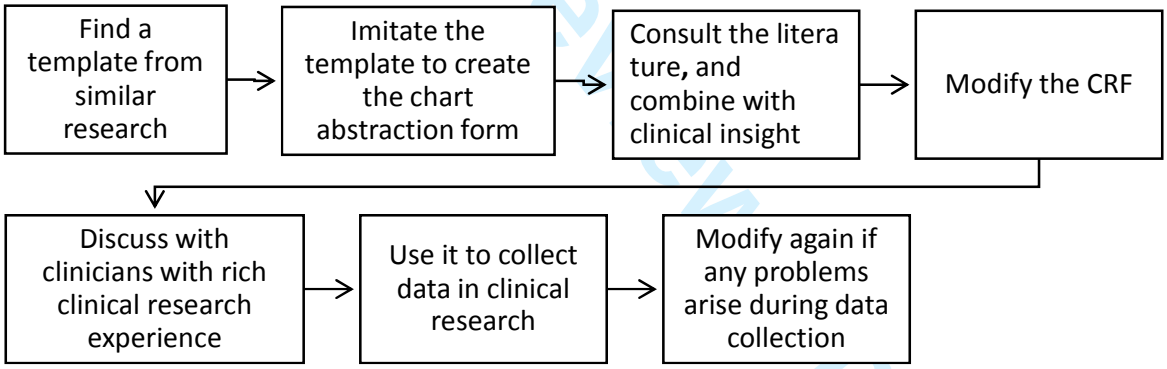
Discussing with clinicians with rich clinical research experience. Some novice clinical researchers reported discussing the CRF with experienced clinical researchers. One participant recounted, "I discussed the CRF with clinicians with rich

experience in clinical research. They deleted useless items directly."

Modifying again if any problems are found during use. Some novice clinical researchers directly used the data collection form without a pilot study, and modified it only when they found problems during use. An anesthesia clinician said, "When the CRF is used to collect data in a clinical research project, some variables are found not to be applicable or absent. In this case, I always think about it and then modify/add the variables." Another cardiology clinician also remarked, "When I use the CRF to collect data from the first few participants, I find some variables can't get data at all. So I realize that these variables also can't get data from other participants, and these variables can contribute little to my research. Then I delete them."

The following flow chart was refined from the above in-depth interviews with four clinicians and a group interview with another four clinicians.

Figure 1. Novice clinical researchers' approach to developing a CRF for clinical research



b. The experienced clinical researcher's approach to selecting variables in a CRF for clinical research.

Experienced clinical researcher also described a multi-faceted approach about how they include variables in a CRF. Most of them have a clear purpose and protocol firstly, and think about the importance, feasibility and statistics of these variables according to their experience in clinical research.

Confirming the purpose and protocol. These clinical researchers first concern themselves with confirming the purpose and developing a protocol before designing

the CRF. For example, a nephrologist said, “First, you should confirm the purpose of your study, disease features or prognosis, follow-up study or cross-sectional study.”

Conducting a literature search for references and related materials. Most of the experienced clinical researchers would review the literature and related materials to consider the variables for inclusion in the CRF. “*I think searching the related literature and other references is important. You can find many variables mentioned in other’s research. You also can list all the related items for reference.*”

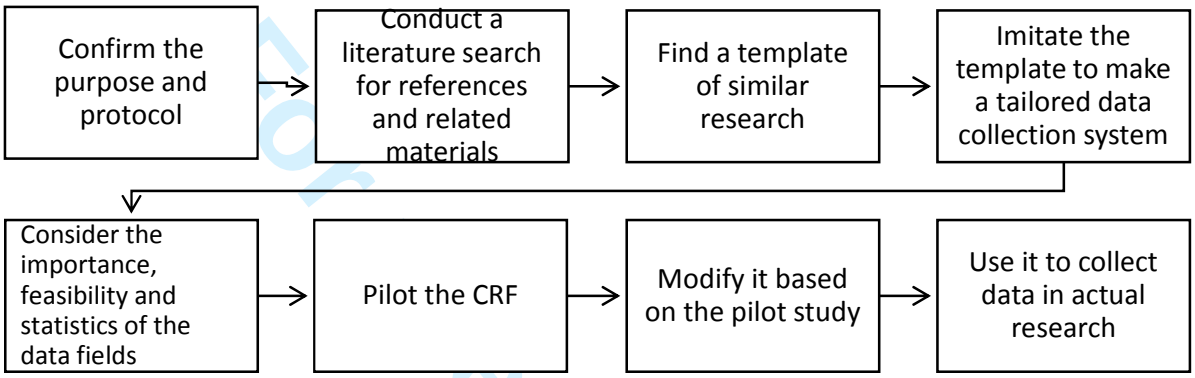
Finding a template from similar research and intimating it. Experienced clinical researchers also find a template to imitate for their own study. However, they consider the variables according to the research protocol, literature and their previous experience. One hematologist explained, “(I) search materials first, then find a template from the literature or similar research previously conducted in my department. Then I imitate the template to make CRF for my project, and optimize and perfect the CRF according to the research protocol.”

Consider the importance, feasibility and statistics of fields. The experienced clinical researchers note that the importance, feasibility and statistical analysis of collected variables should be considered before deciding whether to collect the data in the first place. One hematologist said that, “Even if we do the optimizing and perfecting, we can also find many problems when we use it to collect data in clinical research. We can just modify it again and again. In the final version, many variables are deleted that are not easy to collect, or not very clear, or have little relationship with primary outcome.”

Conducting pilot study and modify the CRF. These experienced researchers also note that piloting the CRF before using to collect data is very important. If any problems are found in the pilot study, the problem can be resolved in a timely way. A thoracic surgeon shared, “A pilot study was conducted after the draft CRF was completed. We then recruited some patients to complete it. We found that some items or scale could not be completed. For example, the tinnitus handicap inventory is too complex for patients. And it is not very important for the final evaluation about the treatment, so we deleted it in the follow-up visits.”

The flow chart depicted in Figure 2 illustrates the overall steps employed by these experienced clinicians based on their interviews.

Figure 2. Experienced clinical researcher's approach to selecting variables in a CRF for clinical research.



c. The clinical research experts' approach to selecting variables in a CRF for clinical research.

Clinical research experts described their multi-faceted approach to identifying variables for inclusion in a CRF or how they help other researchers to design a CRF. They would consider comprehensively about the study, and raised some different and crucial views.

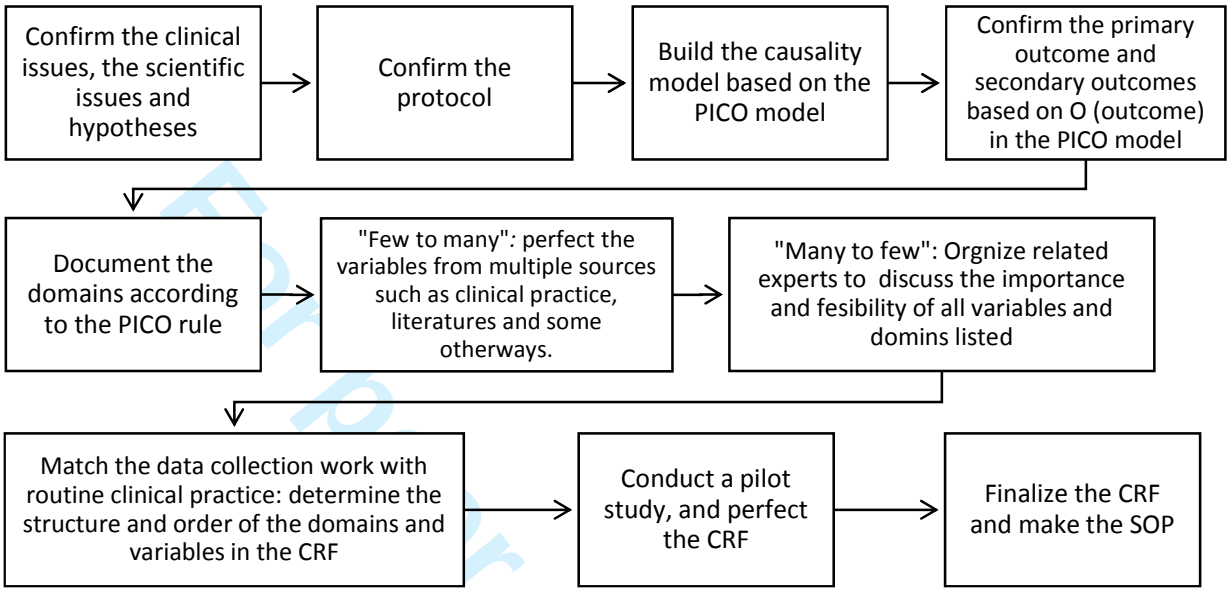
Confirming the clinical and scientific issues, hypotheses, and research protocol. A clear research question and hypotheses are essential before you consider which variables should be included in the CRF. One expert working in the research center of epidemiology stated, "First, you complete the top-level design of the clinical research to confirm the clinical issues, scientific issues and hypotheses, and then confirm the research purpose and protocol."

PICO model is considered to build the causality model. The PICO Model is a format used to help define the clinical research question. P stands for Patient /Population/problem; I stands for Intervention/Prognostic factor /Exposure; C stands for Comparison; O stands for Outcome sought to measure or achieve[14]. The PICO

model as commonly used in evidence-based medicine, can be used to build a causality model and clearly confirm the related variables in the CRF. Another clinical research expert working in the epidemiology research center explained, “In clinical research, you need to confirm the PICO, and build the causality model based on the PICO model. Then confirm the primary outcome and secondary outcome based on O (outcome) in PICO model.” “You need to document the domains based on the PICO model. These domains include but are not limited to population characteristics, intervention or exposure and related outcome, and related confounding factors affecting the causality between intervention and outcome.”

What's more, “few to many” and “many to few” are used to list and screen domains and variables in the research. Many clinical researchers worry about omitting key variables in the CRF, so they often list as many variables as possible. However, many variables are not necessary, and added burden of collecting these data, may even lower the quality of key variables. One expert in clinical research institute emphasized that, “'From less to more', you need to list variables from all domains as much as possible through multiple way such as clinical practice experience and literature. ‘Many to few’, we should organize related experts including clinicians, statisticians, data manager and experts on clinical research design to discuss the necessity and feasibility of all the domains and variables. It is should be considered in terms of accuracy, sensitivity and difficulty of detection and collection, statistics, ethical problems, cost, quality control and so on.” The flow chart depicted in Figure 3 illustrates the overall steps used by these 4 clinical research experts.

Figure 3. The expert clinical researcher's approach to selecting variables in a CRF for clinical research



DISCUSSION

This qualitative study explored the different process of selecting variables in a CRF for clinical research among 3 different levels of clinical researchers. The main findings from this study are that novice, experienced and expert clinical researchers have three progressively sophisticated approaches for selecting variables in a CRF for clinical research. Novice clinician researchers in this study first find a template and then modifying it according to their own study, reviewing it with clinicians with rich clinical research experience, and then completing the CRF to use to collect data. These researchers report operating with a complete protocol, and collect data intending to write an article based on a general structure in mind. If the data were collected according to this process, it would be short of thinking about the whole protocol, study purpose, statistics and so on, which might be a bad quality, in other words, “garbage in garbage out”.

The clinicians with more, but still somewhat limited clinical research experience report trying first to have a more clear purpose and protocol. Similar to novice

researchers, they report looking for a related template, and referring to the structure and domains contained in the template. They report incremental steps well such as referring to the literature, considering clinical practice issues and also thinking about the importance, feasibility and statistics analysis of these items. In addition, they draw lessons from previous studies, conducting pilot studies, and testing the feasibility of data collection in clinical practice. They report this process helps them identify potential problems and modify the data collection template. This process is mainly developed from clinicians' clinical research experience. It should be based on not only rich experience, but also high risk of missing some important domains or variables or excessing unnecessary domains or items.

As for the clinical research experts, before building a data collection template, they report a process of considering the design, confirming the clinical and scientific issues and hypotheses, and developing a research protocol. In addition, they report using the PICO model to help build the domains of CRF. What's more, "few to many" and "many to few" are used to list and screen domains and items in the research. This resonates with the advice of Li et al to collect data so as to be able to just answer the scientific issues or hypothesis, nothing more and nothing less[7].

These results suggest that these clinical researchers have different process of building a CRF that appear to be greatly driven by previous research experience. The process from clinical research experts serves as a valuable example due to its scientific and efficient. It further highlights pitfalls that can be avoided by clinicians who find themselves with little experience, but an expectation for academic productivity.

Based on views from the novice to the expert clinician, the take away messages from this research are:

(a) **Build the CRF to align closely with the research design, and develop and modify the protocol and template according to the feasibility of collecting data in clinical practice**. The risk of designing a CRF after completing the research protocol, is that some variables might not be feasible to collect as part of the clinical enterprise. The consequence of deviating from the protocol is the

risk of compromised research quality. These findings further remind researchers to be certain to modify the CRF when the research protocol is modified.

(b) **Align the CRF with the research protocol.** The variables in the CRF should be consistent with protocol. That is, all domains and variables mentioned in the protocol should be included in CRF, to avoid missing any important domains or variables.

(c) **Develop the CRF to be concise and to the point.** In general, variables only related to the study purpose could be included in the CRF. Unnecessary or redundant variables can require unnecessary time, effort collecting and monitoring. This can distract from limited time and energy of researchers to ensure the quality of key variables in clinical research.

(d) **Refine the template to be readily understood and operational.** In the final template of CRF, every variable should be clear. The structure and order of the CRF should be consistent with research flow and clinical practice.

Strengths, limitations and further research

The strength of this study contribution is a deepened understanding of ommissions that can be made by novice researchers as well as even experienced researchers, and mechanisms to avoid some pitfalls that can occur in the conduct of clinical research. The study particularly informs how clinical researchers select variables in a CRF in their clinical research. In previous research, this question has been only marginally explored relative to the structure and surface, although it can play an important role in improving quality of clinical research.

Furthermore, based on reports of clinical research experts, these findings illustrate an overarching and effective process for determining variables for inclusion in a CRF. Two strategies, the "Top-level Design and concept, and the application of the PICO framework for thinking about the study represents novel thinking in research design raised by these experts. Additionally, the complementary ideas of “few to many” and “many to few” for refining the variables in CRF represents innovative thinking for CRF design.

A limitation of our study concerns the potential selection bias of our sample. Clinician participants are from a premier academic hospital in Beijing, and these participants might have greater knowledge, cognitive skills, and opportunities for research than clinicians in many other hospitals. We both acknowledge the potential for variations on the findings, while also acknowledging the results are genuine and representative of the participants' experiences. Despite variations that exist in practice, we believe the lessons learned based on this study are robust ideas. As noted above, we initially conducted one group interview with clinicians with limited research, but for then for feasibility issues changed to individual interviews. It is plausible that group think may have occurred in the group interview, but there were similar findings.

Finally, further studies are needed to assess whether the approaches and variations described are robust across a wider variety of settings. We are aware of these limits of the study and to further validate the model by applying it to design of CRF for clinical research in different levels of hospitals and clinicians from other settings. None the less, we believe that while there may be additional considerations, these findings are robust enough for any clinicians engaged in clinical research, particularly, novice and moderately experienced researchers.

Acknowledgements: We sincerely appreciate all the participants who made this research possible.

Funding: This study was funded by grant 2013BAI09B14 from the Ministry of Science and Technology of the People's Republic of China. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Contributors: Chu Hongling and Zeng Lin conducted the interviews and all analyses, wrote the first draft of the manuscript and rewrote new drafts based on input from co-authors. Michael Fetters provided extensive input on the structuring of the manuscript, and revisions of the manuscript drafts. Zhao Yiming designed the research project, planned the analyses and provided input and revision of manuscript drafts. All authors read and approved the final manuscript.

Competing interests: None to declare.

Participant consent. Obtained written consent.

Ethics approval: The Peking University Third Hospital Ethics Review Board approved the study.

Provenance and peer review: Not commissioned; externally peer reviewed.

Data sharing statement: Original quotes and audio data of interviews are available on request by emailing the corresponding author. The presented data were anonymized and the risk of identification is low.

REFERENCE

References:

1. Zhao YM, Ding J, Zeng L. Top level design on clinical research. National Medical Journal of China. 2010;90(14):942-946. Chinese.

2. Nahm M, Shepherd J, Buzenberg A, Rostami R, Corcoran A, McCall J, et al. Design and implementation of an institutional case report form library. *Clin Trials* 2011;8(1):94-102.

3. Cao AP, Su BH, Ding DY, et al. The case report form in clinical research. Chinese Journal of Health Statistics. 2000(03):61-64. Chinese.

4. Fu, H.J. and X.X. Luo, Key factors in design of case report form. Yao Xue Xue Bao, 2015. 50(11): p. 1452-5. Chinese.
2015(11):1452-1455.

5. Li QN, Lu F, An FH, Gao R. The design of case report form and analysis of common related problem. Chin J Clin Pharmacol Ther 2013(08):901-906. Chinese.

6. Li ST. The design of case report form and its importance in drug clinical trials. Chines Journal of New Drugs 2004(01):53-55. Chinese.

7. Li XY, Wen ZH, Tang XC, Liang WX. The principle and process of case report form design in clinical trials. Traditional Chinese Drug Research &Clinical Pharmacology 2013(02):206-209. Chinese.

8. Lou H. The design and application of case report form in evidence-based traditional Chinese

medicine research. Shandong Traditional Chinese Medicine University, 2015. Chinese.

9. Lv XY, Zhang ZL, Ai YK, et al. Talking about the design of case report form from the point of data management. Modernization of Traditional Chinese Medicine and Materia-World Science and Technology 2014(03):614-617. Chinese.

10. Wan X, Yang H, Liu JP. The design of case report form in clinical research. Journal of Traditional Chinese Medicine 2007(10):885-887. Chinese.

11. Wang YH, Jiao YZ. The design of case report form and analysis of related problem in traditional Chinese medicine research. China Journal of Traditional Chinese Medicine and Pharmacy 2005(10):620-623. Chinese.

12. Meyer J. Qualitative research in health care. Using qualitative methods in health related action research. *BMJ* 2000;320(7228):178-81.

13. Elo S, Kyngas H. The qualitative content analysis process. *J Adv Nurs* 2008;62(1):107-15.

14. Kleijnen J, Chalmers I. How to practice and teach evidence-based medicine: role of the Cochrane Collaboration. *Acta Anaesthesiol Scand Suppl* 1997;111:231-3.

BMJ Open

How Novice, Skilled and Advanced Clinical Researchers Include Variables in a Case Report Form for Clinical Research:A Qualitative Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016760.R1
Article Type:	Research
Date Submitted by the Author:	05-Jul-2017
Complete List of Authors:	Chu, Hongling; Peking University Third Hospital, Research Center of Clinical Epidemiology Zeng, Lin; Peking University Third Hospital, Research Center of Clinical Epidemiology Fetters, Micheal; University of Michigan Department of Family Medicine Li, Nan; Peking University Third Hospital, Research Center of Clinical Epidemiology Tao, Liyuan; Peking University Third Hospital, Research Center of Clinical Epidemiology Shi, Yanyan; Peking University Third Hospital, Research Center of Clinical Epidemiology Zhang, Hua; Peking University Third Hospital, Research Center of Clinical Epidemiology Wang, Xiaoxiao; Peking University Third Hospital, Research Center of Clinical Epidemiology Li, Fengwei; Beijing Aerospace General Hospital zhao, yiming; Peking University Third Hospital, Research Center of Clinical Epidemiology
Primary Subject Heading:	Qualitative research
Secondary Subject Heading:	Research methods, Qualitative research
Keywords:	QUALITATIVE RESEARCH, clinical study, data collection, research design

How Novice, Skilled and Advanced Clinical Researchers Include Variables in a Case Report Form for Clinical Research: A Qualitative Study

Chu Hongling, Zeng Lin, Michael D Fetters, Li Nan, Tao Liyuan, Shi Yanyan, Zhang Hua, Wang Xiaoxiao, Li Fengwei, Zhao Yiming

Corresponding author: Zhao Yiming, Peking University Third Hospital, Research Center of Clinical Epidemiology, 49#, Huayuanbei Road, Haidian District Beijing, P.R.China 100191. yimingzhao115@163.com, +86-010-82266577, +86-010-82265732.

Chu Hongling, Zeng Lin, Li Nan, Tao Liyuan, Shi Yanyan, Zhang Hua, Wang Xiaoxiao, Zhao Yiming: Peking University Third Hospital, Research Center of Clinical Epidemiology, Beijing, China.

Michael D Fetters: University of Michigan Department of Family Medicine, Ann Arbor, MI, US.

Li Fengwei: Beijing Aerospace General Hospital, Beijing, China.

Abstract

Objectives: Despite varying degrees in research training, most academic clinicians are expected to conduct clinical research. The objective of this research was to understand how clinical researchers of different skill levels include variables in a case report form for their clinical research.

Setting: The setting for this research was a major academic institution in Beijing, China.

Participants: The target population was clinical researchers with three levels of experience, namely, limited clinical research experience, clinicians with rich clinical

research experience and clinical research experts.

Methods: Using a qualitative approach, we conducted 13 individual interviews (face–to-face) and one group interview (n=4) with clinical researchers from June to September 2016. Based on maximum variation sampling to identify researchers with three levels of research experience: 8 clinicians with limited clinical research experience, 5 clinicians with rich clinical research experience, and 4 clinical research experts. These 17 researchers had diverse hospital-based medical specialties and or specialization in clinical research.

Results: Our analysis yields a typology of three processes developing a case report form that varies according to research experience level. Novice clinician researchers often have an incomplete protocol or none at all, and conduct data collection and publication based on a general framework. Experienced clinician researchers include variables in the case report form based on previous experience with attention to including domains or items at risk for omission, and by eliminating unnecessary variables. Expert researchers consider comprehensively in advance data collection and implementation needs and plan accordingly.

Conclusion: These results illustrate increasing levels of sophistication in research planning that increase sophistication in selection for variables in the case report form. These findings suggest that novice and intermediate-level researchers could benefit by emulating the comprehensive planning procedures such as those utilized by expert clinical researchers.

Keywords: qualitative research; clinical study; data collection, research design

Article summary

Strengths and limitations of this study

- The study used qualitative interviews to explore the novel topic of how researchers create a case report form.
- The study involved the development of three visual models to depict the approach of novice, excellent, and expert clinical researchers
- The expert clinical researchers shared guiding principles they use to conduct

research.

- A limitation of our study concerns the potential selection bias of our sample as clinician participants are from a premier academic hospital in Beijing, and these participants might have greater knowledge, cognitive skills, and opportunities for research than clinicians in many other hospitals in China.

INTRODUCTION

Conducting clinical research is a multistage process. It can be divided into three stages: the first stage is top-level design, the second stage is protocol design and implementation, the third stage is conducting data analysis, interpretation and writing of the paper[1]. During the first two stages, researchers should consider what variables are important, and how to collect the data.

A case report form (CRF) is an instrument to structure and facilitate collection of data for clinical research[2]. Most CRFs are customized to collect data specific to a particular clinical study protocol. Case report form development represents a significant part of the clinical trial process and can impact study success. A well-designed CRF is required for database construction, data accuracy, data query/cleaning, CRF completion and statistical analysis.

In our institution, to facilitate greater clinical research efficiency, our lab became interested in how to develop an approach to build a CRF. Our interest extended to both the structure and variables of the CRF. The structure has been the topic of concern among some researchers[3-11]. In one published paper, Li et al. emphasize that the design of CRF needs the cooperation and efforts of each member of the study group[7]. In another related paper, Wan et al note that the CRF should keep privacy for participants, include a tracked page or modules, and some other fields on forms[10]. However, the question about what variables should be included in CRF has received less consideration despite its crucial role in the quality of clinical research.

The goal of this qualitative research study was to explore how clinical researchers select variables for a CRF, and was part of a larger mixed-methods project

to develop an approach for clinical researchers to build systematic variables for clinical research. These findings of this research could provide an approach for choosing variables in a CRF by clinical researchers. This can serve as a reference to other researchers, and lay a foundation for further inquiry into what variables to include in a CRF.

METHOD

Qualitative inquiry is an approach particularly useful when little is known about the phenomenon under study[12]. As little is known about how clinicians determine variables to include in a CRF we deemed a qualitative approach as useful.

Setting:

The setting for this research was a heavily research-focused major academic institution in Beijing, China.

Study Population:

These hospitals are host to over 1000 faculty, and to some degree there is an expectation, or hope that all will engage in clinical research at some level. Purposive sampling was used in this study. The target population was clinical researchers with three levels of experience, namely, limited clinical research experience, clinicians with rich clinical research experience and clinical research experts. Seventeen clinical researchers chose to participate in this study.

Data collection instrument:

A semi-structured interview guide was first developed based on a group discussion and a preliminary pilot study with 2 clinical researchers. The primary interview questions developed to generate data for the study was, “What process do you use to design a case report form for clinical research?” Key probes were, “What are the difficulties and challenges encountered when you designed case report form for your project?” and “What is your previous experience joining a clinical research project?” The overall interview guide was designed for a study looking comprehensively at clinical researchers’ approaches to data collection and thus had additional questions in the interview guide as well.

Recruitment:

Individuals targeted for enrollment were contacted by email by a research assistant. We used maximum variation sampling to identify individuals with different degrees of research experience and different medical specialties in the project's host academic institution, a teaching hospital and clinical research institute in Beijing. The inclusion criteria were:

a. Meet one of the following conditions: 1) Clinicians with limited clinical research experience who were defined as rarely participating in clinical research, a criterion operationalized as no experience to one experience designing a CRF for clinical research; 2) clinicians with rich clinical research experience defined as researchers with experience participating in several clinical research studies and experience designing three to nine data collection reforms or case report forms for clinical research. 3) Clinical research experts defined as researchers with experience in clinical research for five years or more, and participation directly in 10 or more clinical research projects.
b. Being willing to participate in the study.

Data collection

Two research assistants (RAs), both females with a PhD, conducted interviews from June to September in 2016. The RAs first conducted one focus group interview with clinicians who had limited experience in clinical research. As gathering busy clinicians for focus group interviews proved difficult, the RAs changed to face–to-face semi-structured in-depth interviews with an additional nine clinicians and four clinical research experts. The same questions were posed in the same order to all the participants, whether the interviews were conducted in the focus group or individual interviews. One question was added for the four clinical research experts, namely, "What have you encountered when you directed other clinical researchers' project". The interviews lasted between (25-40) minutes and were conducted in a location that was quiet without interruption, either the participant's office or the interviewee's conference room. All interviews were audio-recorded and transcribed verbatim into

Chinese. Data collection was complete after 17 interviews, the point when saturation of themes was reached.

Data analysis

The transcribed data were analyzed in Chinese using thematic analysis with an inductive approach. Nvivo 11 was used to assist in the analysis of the data. Two researchers (CH and Zl) independently coded and analyzed the transcripts by: selecting the units of analysis, making sense of the transcribed data, developing codes, categorizing the data and abstracting[13]. The analysis focused on text from the four primary questions noted above but also used related information from other interview guide questions and context-specific language probes. The analysts discussed and reached an agreement on the coding and categorization after reviewing one interview. The two researchers independently coded the remaining transcripts. Differences were minimal. All research team members agreed with the final results. After constructing the models, the team confirmed the findings by checking the results with the interviews transcript. Member checking was used to share the results with all the interviewees by email--they raised no objections or new considerations, and agreed with the findings.

Below, the findings of the study are presented by illustrating the three identified models and illustrative quotes from the interviews. The researchers who conducted the data collection and analysis translated the quotes selected for this article into English. Original quotes are available upon request.

Ethical considerations

Each participant was informed about the study procedures and was free to withdraw from the research. All participants provided consent to participate in the study including audiotaping and transcription of the interviews by providing written informed consent. Potentially identifying information was removed from each transcript and each interviewee was assigned a unique identification number to protect his/her anonymity. The institutional review board (IRB) at Peking University Third

Hospital approved the study.

RESULTS

The 17 clinical researchers had diverse backgrounds, including otolaryngology, pharmacy, endocrinology, orthopedics, anesthesia, radiology, neurology, cardiology, nephrology, Hematology, thoracic surgery, epidemiology and biostatistics, clinical epidemiology, clinical research data management, and clinical research methodology. (Table1)

Table 1 Demographics of the participants

Participants level	Clinicians with limited clinical research experience	Clinicians with rich clinical research experience	Clinical research experts
Number	8	5	4
Gender			
Male (N)	2	2	3
Female (N)	6	3	1
Average age (mean± SD)	27±4	35±2	44±12
Department	Otolaryngology, Pharmacy, Endocrinology, Orthopedics, Anesthesia, Radiology, Neurologist, Cardiology.	Nephrology, Otolaryngology, Thoracic surgery, Hematology	Epidemiology Or Clinical research methodology
Experience in clinical research (projects)	Participating in clinical research work for 1-3 years	Being involved in 5-12 clinical research projects	Working in clinical research for 11-20 years
Data collection	4 semi-structured in-depth interviews, and 1 focus group interview for 4 clinicians	Semi-structured in-depth interviews	Semi-structured in-depth interviews

A typology of how clinical researchers of different skill levels include variables in CRF for clinical research

Based on our analysis, we developed a typology of how clinical researchers of different skill levels select variables in a CRF for clinical research. These three

models are illustrated as Figures 1, 2, and 3. The findings are supported by quoted comments from the research participants. The models are illustrated using a flow chart showing the different process for each of three levels of the participants, e.g., novice clinicians with limited clinical research experience, intermediate-level clinicians with rich clinical research experience, and clinical research experts.

a. The novice clinical researcher's approach to selecting variables in a CRF for clinical research.

Novice clinicians with limited clinical research experience described a multi-faceted approach to selecting variables for inclusion in their CRFs. When they planned the variables to include in their CRFs, most had no clearly defined comprehensive approach.

Finding the template from similar research and imitating it. Most novice clinical researchers mentioned that they would find a similar data collection form for reference and modify it according to the needs of their own study. As one endocrinologist noted, "First, I search on the Internet to find a template from similar research, then I modify the template based on my research". One otolaryngology clinician opined, "I ask for a data collection template from other experienced researchers in my department. I then imitate the template to make my own data collection template, and modify it in places according my research." An Orthopedics clinician reflected his opinion that most novice clinical researchers are doing like this, "I will imitate another data collection template, then copy what is applicable to my research, and modify those variables that are not applicable. I think most of us are doing like this, because it is easier for us."

Discussing with clinicians with rich clinical research experience. Some novice clinical researchers reported discussing the CRF with experienced clinical researchers. One participant recounted, "I discussed the CRF with clinicians with rich experience in clinical research. They deleted useless items directly."

Modifying again if any problems are found during use. Some novice clinical researchers directly used the data collection form without a pilot study, and modified it only when they found problems during use. An anesthesia clinician said, "When the

CRF is used to collect data in a clinical research project, some variables are found not to be applicable or absent. In this case, I always think about it and then modify/add the variables." Another cardiology clinician also remarked, "When I use the CRF to collect data from the first few participants, I find some variables can't get data at all. So I realize that these variables also can't get data from other participants, and these variables can contribute little to my research. Then I delete them."

The following flow chart was refined from the above in-depth interviews with four clinicians and a group interview with another four clinicians.

Figure 1. Novice clinical researchers' approach to developing a CRF for clinical research

b. The experienced clinical researcher's approach to selecting variables in a CRF for clinical research.

Experienced clinical researchers also described a multi-faceted approach for how they include variables in a CRF. Most of them have a clear purpose and protocol firstly, and think about the importance, feasibility and statistics of these variables according to their experience in clinical research.

Confirming the purpose and protocol. These clinical researchers first concern themselves with confirming the purpose and developing a protocol before designing the CRF. For example, a nephrologist said, "First, you should confirm the purpose of your study, disease features or prognosis, follow-up study or cross-sectional study."

Conducting a literature search for references and related materials. Most of the experienced clinical researchers would review the literature and related materials to consider the variables for inclusion in the CRF. "I think searching the related literature and other references is important. You can find many variables mentioned in other's research. You also can list all the related items for reference."

Finding a template from similar research and intimating it. Experienced clinical researchers also find a template to imitate for their own study. However, they consider the variables according to the research protocol, literature and their previous experience. One hematologist explained, "(I) search materials first, then find a

template from the literature or similar research previously conducted in my department. Then I imitate the template to make CRF for my project, and optimize and perfect the CRF according to the research protocol."

Consider the importance, feasibility and statistics of fields. The experienced clinical researchers note that the importance, feasibility and statistical analysis of collected variables should be considered before deciding whether to collect the data in the first place. One hematologist said, "Even if we do the optimizing and perfecting, we can also find many problems when we use it to collect data in clinical research. We can just modify it again and again. In the final version, many variables are deleted that are not easy to collect, or not very clear, or have little relationship with the primary outcome."

Conduct a pilot study and modify the CRF. These experienced researchers also note that piloting the CRF before using it to collect data is very important. If any problems are found in the pilot study, the problem can be resolved in a timely way. A thoracic surgeon shared, "A pilot study was conducted after the draft CRF was completed. We then recruited some patients to complete it. We found that some items or scales could not be completed. For example, the tinnitus handicap inventory is too complex for patients. And it is not very important for the final evaluation about the treatment, so we deleted it in follow-up visits."

The flow chart depicted in Figure 2 illustrates the overall steps employed by these experienced clinicians based on their interviews.

Figure 2. Experienced clinical researcher's approach to selecting variables in a CRF for clinical research.

c. The clinical research experts' approach to selecting variables in a CRF for clinical research.

Clinical research experts described their multi-faceted approach to identifying variables for inclusion in a CRF or how they help other researchers design a CRF. They would consider the study comprehensively, and they raised some different and crucial views.

Confirming the clinical and scientific issues, hypotheses, and research protocol. A clear research question and hypotheses are essential before you consider which variables should be included in the CRF. One expert working in the epidemiology research center stated, "First, you complete the top-level design of the clinical research to confirm the clinical issues, scientific issues and hypotheses, and then confirm the research purpose and protocol."

PICO model is considered to build the causality model. The PICO Model is a format used to help define the clinical research question. P stands for Patient /Population/problem; I stands for Intervention/Prognostic factor /Exposure; C stands for Comparison; O stands for Outcome sought to measure or achieve[14]. The PICO model as commonly used in evidence-based medicine, can be employed to build a causality model and clearly confirm the related variables in the CRF. Another clinical research expert working in the epidemiology research center explained, "In clinical research, you need to confirm the PICO, and build the causality model based on the PICO model. Then confirm the primary outcome and secondary outcome based on O (outcome) in PICO model." "You need to document the domains based on the PICO model. These domains include but are not limited to population characteristics, intervention or exposure and related outcome, and related confounding factors affecting the causality between the intervention and outcome."

What's more, "few to many" and "many to few" are used to list and screen domains and variables in the research. Many clinical researchers worry about omitting key variables in the CRF, so they often list as many variables as possible. However, many variables are not necessary, and the added burden of collecting these data, may even lower the quality of key variables. One expert in the clinical research institute emphasized that, "'From less to more', you need to list variables from all domains as much as possible through multiple ways such as clinical practice experience and literature. 'Many to few', we should organize related experts including clinicians, statisticians, data manager and experts on clinical research design to discuss the necessity and feasibility of all the domains and variables. It should be considered in terms of accuracy, sensitivity and difficulty of detection and collection,

statistics, ethical problems, cost, quality control and so on." The flow chart depicted in Figure 3 illustrates the overall steps used by these 4 clinical research experts.

Figure 3. The expert clinical researcher's approach to selecting variables in a CRF for clinical research

DISCUSSION

This qualitative study reveals a different process of selecting variables for inclusion in a CRF for clinical research among three different levels of clinical researchers. Novice, experienced and expert clinical researchers have progressively sophisticated approaches for selecting variables in a CRF for clinical research. The novice clinician researcher approach of finding a template, modifying it according to their own study, reviewing it with clinicians with rich clinical research experience, and then completing the CRF when collecting data can be problematic. This process can truncate thinking about the whole protocol, study purpose, statistics, etc., and risks compromising the study quality. This point is consistent with previous discussion of Ionnides et al[15] that research in a previously understudied domain might supply too little information to be useful. The resulting risk is, small uninformative studies that remain common in several specialties.

Relative to the existing literature about the development of CRFs, this research has expanded understanding about what variables should be included in a CRF, procedures used by expert researchers for first exploding, and then limiting the range of variables to be studied. The notion of "few to many" and "many to few" to first comprehensively generate, then severely truncate to limit the domains is similar to previous advice. Li et al suggest collecting data so as to be able to just answer the scientific issues or hypothesis, nothing more and nothing less[7].

Based on views from the novice to the expert clinician, the take away message from this research is that increasing levels of sophistication in research planning reflect increasing levels of sophisticated selection of variables on the case report form.

As alluded to by Chalmers et al, there are several principles that can help guide clinical researchers to conduct efficient and high quality studies[16]:

(a) **Build the CRF to align closely with the research design, and develop and modify the protocol and template according to the feasibility of collecting data in clinical practice**. The risk of designing a CRF after completing the research protocol, is that some variables might not be feasible to collect as part of the clinical enterprise. The consequence of deviating from the protocol is the risk of compromised research quality. These findings further remind researchers to be certain to modify the CRF when the research protocol is modified.

(b) **Align the CRF with the research protocol.** The variables in the CRF should be consistent with protocol. That is, all domains and variables mentioned in the protocol should be included in CRF, to avoid missing any important domains or variables.

(c) **Develop the CRF to be concise and to the point.** In general, variables only related to the study purpose could be included in the CRF. Unnecessary or redundant variables can require unnecessary time, effort collecting and monitoring. This can distract from limited time and energy of researchers to ensure the quality of key variables in clinical research. Berge et al also identify that data collection forms can be too complex and burden centers with a requirement to collect data items that are never analyzed or reported[17].

(d) **Refine the template to be readily understood and operational.** In the final template of CRF, every variable should be clear. The structure and order of the CRF should be consistent with research flow and clinical practice.

These findings speak to fundamentals of conducting high quality research, and while based on research in a single institution, can reasonably be expected to hold true in a broad range of settings.

A potential limitation of our study concerns the risk of selection bias of our sample. Clinician participants are from a premier academic hospital in Beijing, and these participants might have greater knowledge, cognitive skills, and opportunities for research than clinicians in many other hospitals. We acknowledge the potential for

variations on the findings in other settings, while also acknowledging the results are genuine and representative of the participants' experiences. Despite variations that exist in practice, we believe the lessons learned based on this study are transferable and robust ideas. Another potential limitation is that the interviewers were aware of the interviewed researchers' experience and the study hypotheses, and this could raise concern about confirmation bias, however the models were built and reviewed by the researchers themselves and this suggests that the processes delineated represent their views.

Further studies are needed to assess whether the approaches and variations described are robust across a wider variety of settings. The proposed expert model could be further validated by applying it to design of CRF for clinical research in different levels of hospitals and clinicians from other settings. Regardless, we believe that while there may be additional considerations, these findings are a robust and meaningful reference for any clinicians engaged in clinical research, and particularly, novice and moderately experienced researchers who wish to learn lessons from experts.

In conclusion, this research illustrates that increasing levels of sophistication in research planning reflect increasing levels of sophistication in the selection of variables for inclusion in as case report form. Thus, novice and intermediate-level clinical researchers alike could benefit by emulating the comprehensive planning procedures utilized by expert clinical researchers.

Acknowledgements: We sincerely appreciate all the participants who made this research possible.

Funding: This study was funded by grant 2013BAI09B14 from the Ministry of Science and Technology of the People's Republic of China. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Contributors: Chu Hongling and Zeng Lin conducted the interviews and all analyses, wrote the first draft of the manuscript and rewrote new drafts based on input from co-authors. Michael Fetters provided extensive input on the structuring of the

manuscript, and revisions of the manuscript drafts. Zhao Yiming designed the research project, planned the analyses and provided input and revision of manuscript drafts. All authors read and approved the final manuscript.

Competing interests: None to declare.

Participant consent. Obtained written consent.

Ethics approval: The Peking University Third Hospital Ethics Review Board approved the study.

Provenance and peer review: Not commissioned; externally peer reviewed.

Data sharing statement: Original quotes and audio data of interviews are available on request by emailing the corresponding author. The presented data were anonymized and the risk of identification is low.

References:

1. Zhao YM, Ding J, Zeng L. Top level design on clinical research. National Medical Journal of China. 2010;90(14):942-946. Chinese.

2. Nahm M, Shepherd J, Buzenberg A, Rostami R, Corcoran A, McCall J, et al. Design and implementation of an institutional case report form library. *Clin Trials* 2011;8(1):94-102.

3. Cao AP, Su BH, Ding DY, et al. The case report form in clinical research. Chinese Journal of Health Statistics. 2000(03):61-64. Chinese.

4. Fu, H.J. and X.X. Luo, Key factors in design of case report form. Yao Xue Xue Bao, 2015. 50(11): p. 1452-1455. Chinese.

5. Li QN, Lu F, An FH, Gao R. The design of case report form and analysis of common related problem. Chin J Clin Pharmacol Ther 2013(08):901-906. Chinese.

6. Li ST. The design of case report form and its importance in drug clinical trials. Chines Journal of New Drugs 2004(01):53-55. Chinese.

7. Li XY, Wen ZH, Tang XC, Liang WX. The principle and process of case report form design in clinical trials. Traditional Chinese Drug Research &Clinical Pharmacology 2013(02):206-209. Chinese.

8. Lou H. The design and application of case report form in evidence-based traditional Chinese medicine research. Shandong Traditional Chinese Medicine University, 2015. Chinese.

9. Lv XY, Zhang ZL, Ai YK, et al. Talking about the design of case report form from the point of data management. Modernization of Traditional Chinese Medicine and Materia-World Science and Technology 2014(03):614-617. Chinese.
10. Wan X, Yang H, Liu JP. The design of case report form in clinical research. Journal of Traditional Chinese Medicine 2007(10):885-887. Chinese.
11. Wang YH, Jiao YZ. The design of case report form and analysis of related problem in traditional Chinese medicine research. China Journal of Traditional Chinese Medicine and Pharmacy 2005(10):620-623. Chinese..
12. Meyer J. Qualitative research in health care. Using qualitative methods in health related action research. *BMJ* 2000;320(7228):178-81.
13. Elo S, Kyngas H. The qualitative content analysis process. *J Adv Nurs* 2008;62(1):107-15.
14. Kleijnen J, Chalmers I. How to practice and teach evidence-based medicine: role of the Cochrane Collaboration. *Acta Anaesthesiol Scand Suppl* 1997;111:231-3.
15. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383(9912):166-75.
16. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;383(9912):156-65.
17. Berge E, Salman RA, van der Worp HB, Stapf C, Sandercock P, Sprigg N, et al. Increasing value and reducing waste in stroke research. *Lancet Neurol* 2017;16(5):399-408.

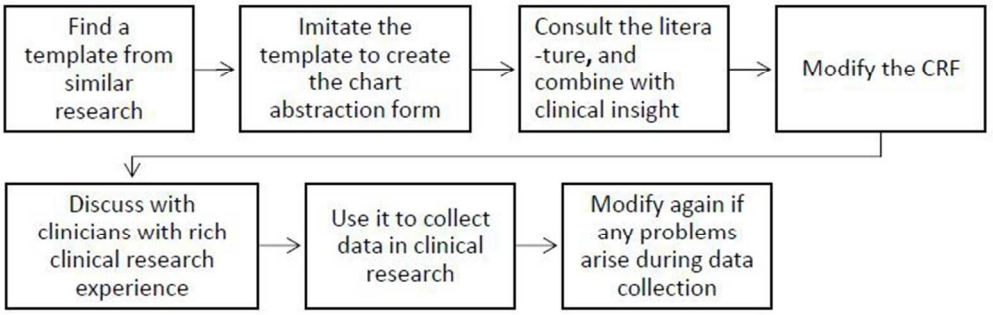


Figure 1. Novice clinical researchers' approach to developing a CRF for clinical research

67x21mm (300 x 300 DPI)

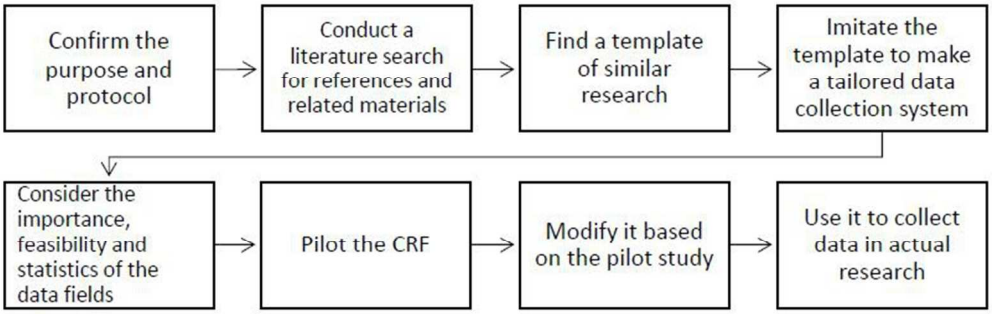


Figure 2. Experienced clinical researcher's approach to selecting variables in a CRF for clinical research.

68x22mm (300 x 300 DPI)

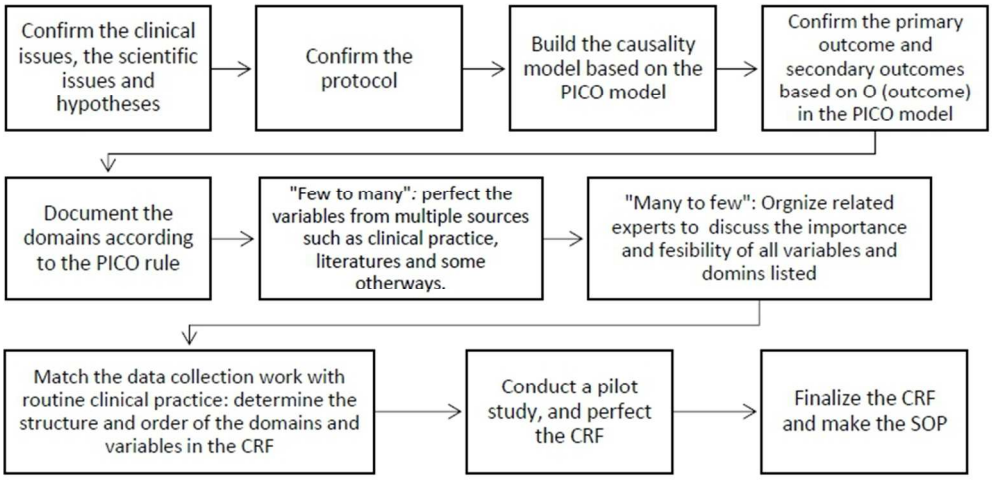


Figure 3. The expert clinical researcher's approach to selecting variables in a CRF for clinical research

70x34mm (300 x 300 DPI)

Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist		
No Item	**Guide questions/description**	**Page number**
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer/facilitator	Which author/s conducted the interview or focus group?	5
2. Credentials	What were the researcher's credentials? E.g. PhD, MD	5
3. Occupation	What was their occupation at the time of the study?	5
4. Gender	Was the researcher male or female?	5
5.Experience and training Relationship with participants	What experience or training did the researcher have?	5
6.Relationship established	Was a relationship established prior to study commencement?	5
7.Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	N
8. Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	N
Domain 2: study design		
Theoretical framework		
9. Methodological orientation and Theory	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	6
Participant selection		
10. Sampling	How were participants selected? e.g. purposive, convenience, consecutive, snowball	4
11. Method of approach	How were participants approached? e.g. face-to-face, telephone, mail, email	6
12. Sample size	How many participants were in the study?	4
13. Non-participation	How many people refused to participate or dropped out? Reasons?	N
Setting		
14. Setting of data collection	Where was the data collected? e.g. home, clinic, workplace	6
15. Presence of non-participants	Was anyone else present besides the participants and researchers?	6
16. Description of sample	What are the important characteristics of the sample? e.g. demographic data, date	7
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	5

18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	6
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	6
20. Field notes	Were field notes made during and/or after the interview or focus group?	N
21. Duration	What was the duration of the interviews or focus group?	6
22. Data saturation	Was data saturation discussed?	6
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	N
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	6
25. Description of the coding tree	Did authors provide a description of the coding tree?	N
26. Derivation of themes	Were themes identified in advance or derived from the data?	8
27. Software	What software, if applicable, was used to manage the data?	6
28. Participant checking	Did participants provide feedback on the findings?	6
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes / findings? Was each quotation identified? e.g. participant number	8--12
30. Data and findings consistent	Was there consistency between the data presented and the findings?	8--13
31. Clarity of major themes	Were major themes clearly presented in the findings?	8--13
32. Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	8--13

BMJ Open

How Novice, Skilled and Advanced Clinical Researchers Include Variables in a Case Report Form for Clinical Research:A Qualitative Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016760.R2
Article Type:	Research
Date Submitted by the Author:	18-Jul-2017
Complete List of Authors:	Chu, Hongling; Peking University Third Hospital, Research Center of Clinical Epidemiology Zeng, Lin; Peking University Third Hospital, Research Center of Clinical Epidemiology Fetters, Micheal; University of Michigan Department of Family Medicine Li, Nan; Peking University Third Hospital, Research Center of Clinical Epidemiology Tao, Liyuan; Peking University Third Hospital, Research Center of Clinical Epidemiology Shi, Yanyan; Peking University Third Hospital, Research Center of Clinical Epidemiology Zhang, Hua; Peking University Third Hospital, Research Center of Clinical Epidemiology Wang, Xiaoxiao; Peking University Third Hospital, Research Center of Clinical Epidemiology Li, Fengwei; Beijing Aerospace General Hospital zhao, yiming; Peking University Third Hospital, Research Center of Clinical Epidemiology
Primary Subject Heading:	Qualitative research
Secondary Subject Heading:	Research methods, Qualitative research
Keywords:	QUALITATIVE RESEARCH, clinical study, data collection, research design

How Novice, Skilled and Advanced Clinical Researchers Include Variables in a Case Report Form for Clinical Research: A Qualitative Study

Chu Hongling, Zeng Lin, Michael D Fetters, Li Nan, Tao Liyuan, Shi Yanyan, Zhang Hua, Wang Xiaoxiao, Li Fengwei, Zhao Yiming

Corresponding author: Zhao Yiming, Peking University Third Hospital, Research Center of Clinical Epidemiology, 49#, Huayuanbei Road, Haidian District Beijing, P.R.China 100191. yimingzhao115@163.com, +86-010-82266577, +86-010-82265732.

Chu Hongling, Zeng Lin, Li Nan, Tao Liyuan, Shi Yanyan, Zhang Hua, Wang Xiaoxiao, Zhao Yiming: Peking University Third Hospital, Research Center of Clinical Epidemiology, Beijing, China.

Michael D Fetters: University of Michigan Department of Family Medicine, Ann Arbor, MI, US.

Li Fengwei: Beijing Aerospace General Hospital, Beijing, China.

Abstract

Objectives: Despite varying degrees in research training, most academic clinicians are expected to conduct clinical research. The objective of this research was to understand how clinical researchers of different skill levels include variables in a case report form for their clinical research.

Setting: The setting for this research was a major academic institution in Beijing, China.

Participants: The target population was clinical researchers with three levels of experience, namely, limited clinical research experience, clinicians with rich clinical

research experience and clinical research experts.

Methods: Using a qualitative approach, we conducted 13 individual interviews (face–to-face) and one group interview (n=4) with clinical researchers from June to September 2016. Based on maximum variation sampling to identify researchers with three levels of research experience: 8 clinicians with limited clinical research experience, 5 clinicians with rich clinical research experience, and 4 clinical research experts. These 17 researchers had diverse hospital-based medical specialties and or specialization in clinical research.

Results: Our analysis yields a typology of three processes developing a case report form that varies according to research experience level. Novice clinician researchers often have an incomplete protocol or none at all, and conduct data collection and publication based on a general framework. Experienced clinician researchers include variables in the case report form based on previous experience with attention to including domains or items at risk for omission, and by eliminating unnecessary variables. Expert researchers consider comprehensively in advance data collection and implementation needs and plan accordingly.

Conclusion: These results illustrate increasing levels of sophistication in research planning that increase sophistication in selection for variables in the case report form. These findings suggest that novice and intermediate-level researchers could benefit by emulating the comprehensive planning procedures such as those utilized by expert clinical researchers.

Keywords: qualitative research; clinical study; data collection, research design

Article summary

Strengths and limitations of this study

- The study used qualitative interviews to explore the novel topic of how researchers create a case report form.
- The study involved the development of three visual models to depict the approach of novice, excellent, and expert clinical researchers
- The expert clinical researchers shared guiding principles they use to conduct

research.

- A limitation of our study concerns the potential selection bias of our sample as clinician participants are from a premier academic hospital in Beijing, and these participants might have greater knowledge, cognitive skills, and opportunities for research than clinicians in many other hospitals in China.

INTRODUCTION

Conducting clinical research is a multistage process. It can be divided into three stages: the first stage is top-level design, the second stage is protocol design and implementation, the third stage is conducting data analysis, interpretation and writing of the paper[1]. During the first two stages, researchers should consider what variables are important, and how to collect the data.

A case report form (CRF) is an instrument to structure and facilitate collection of data for clinical research[2]. Most CRFs are customized to collect data specific to a particular clinical study protocol. Case report form development represents a significant part of the clinical trial process and can impact study success. A well-designed CRF is required for database construction, data accuracy, data query/cleaning, CRF completion and statistical analysis.

In the Research Center of Clinical Epidemiology, to facilitate greater clinical research efficiency, the question that how to develop an approach to build a CRF became meaningful. This interest extended to both the structure and variables of the CRF. The structure has been the topic of concern among some researchers[3-11]. In one published paper, Li et al. emphasize that the design of CRF needs the cooperation and efforts of each member of the study group[7]. In another related paper, Wan et al note that the CRF should maintain privacy for participants, include a tracked page or modules, and some other fields on forms[10]. However, the question about what variables should be included in CRF has received less consideration despite its crucial role in the quality of clinical research.

The goal of this qualitative research study was to explore how clinical

researchers select variables for a CRF, and was part of a larger mixed-methods project to develop an approach for clinical researchers to build systematic variables for clinical research. These findings of this research could provide an approach for choosing variables in a CRF by clinical researchers. This can serve as a reference to other researchers, and lay a foundation for further inquiry into what variables to include in a CRF.

METHOD

Qualitative inquiry is an approach particularly useful when little is known about the phenomenon under study[12]. As little is known about how clinicians determine variables to include in a CRF it was deemed a qualitative approach as useful.

Setting:

The setting for this research was a heavily research-focused major academic institution and also is a University hospital in Beijing, China.

Study Population:

These hospitals are host to over 1000 faculty, and to some degree there is an expectation, or hope that all will engage in clinical research at some level. Maximum variation sampling[13] was used in this study. The target population was clinical researchers with three levels of experience, namely, limited clinical research experience, clinicians with rich clinical research experience and clinical research experts. Seventeen clinical researchers chose to participate in this study.

Data collection instrument:

A semi-structured interview guide was first developed based on a group discussion and a preliminary pilot study with 2 clinical researchers. The primary interview question developed to generate data for the study was, “What process do you use to design a case report form for clinical research?” Key probes were, “What are the difficulties and challenges encountered when you designed case report form for your project?” and “What is your previous experience joining a clinical research project?” The overall interview guide was designed for a study looking comprehensively at clinical researchers’ approaches to data collection and thus had

additional questions in the interview guide as well.

Recruitment:

Individuals targeted for enrollment were contacted by email by a research assistant. We used maximum variation sampling to identify individuals with different degrees of research experience and different medical specialties in the project's host academic institution, a teaching hospital and clinical research institute in Beijing. The inclusion criteria were:

a. Meet one of the following conditions: 1) Clinicians with limited clinical research experience who were defined as rarely participating in clinical research, a criterion operationalized as no experience to one experience designing a CRF for clinical research; 2) clinicians with rich clinical research experience defined as researchers with experience participating in several clinical research studies and experience designing three to nine data collection reforms or case report forms for clinical research. 3) Clinical research experts defined as researchers with experience in clinical research for five years or more, and participation directly in 10 or more clinical research projects.

Data collection

Two research assistants (RAs), both females with a PhD, conducted interviews from June to September in 2016. The RAs first conducted one focus group interview with clinicians who had limited experience in clinical research. As gathering busy clinicians for focus group interviews proved difficult, the RAs changed to face–to-face semi-structured in-depth interviews with an additional nine clinicians and four clinical research experts. The same questions were posed in the same order to all the participants, whether the interviews were conducted in the focus group or individual interviews. One question was added for the four clinical research experts, namely, "What have you encountered when you directed other clinical researchers' project". The interviews lasted between (25-40) minutes and were conducted in a location that was quiet without interruption, either the participant's office or the interviewee's conference room. All interviews were audio-recorded and transcribed verbatim into

Chinese. Data collection was complete after 17 interviews, the point when saturation of themes was reached.

Data analysis

The transcribed data were analyzed in Chinese using thematic analysis with an inductive approach. Nvivo 11 was used to assist in the analysis of the data. Two researchers (CH and ZL) independently coded and analyzed the transcripts by: selecting the units of analysis, making sense of the transcribed data, developing codes, categorizing the data and abstracting[14]. The analysis focused on text from the four primary questions noted above but also used related information from other interview guide questions and context-specific language probes. The analysts discussed and reached an agreement on the coding and categorization after reviewing one interview. The two researchers independently coded the remaining transcripts. Differences were minimal. All research team members agreed with the final results in a final discussion meeting. After constructing the models, the team confirmed the findings by checking the results with the interviews transcript. Member checking was used to share the results with all the interviewees by email--they raised no objections or new considerations, and agreed with the findings.

Below, the findings of the study are presented by illustrating the three identified models and illustrative quotes from the interviews. The researchers who conducted the data collection and analysis translated the quotes selected for this article into English. Original quotes are available upon request.

Ethical considerations

The institutional review board (IRB) at Peking University Third Hospital approved the study. Each participant was informed about the study procedures and was free to withdraw from the research. All participants provided consent to participate in the study including audiotaping and transcription of the interviews by providing written informed consent. Potentially identifying information was removed from each transcript and each interviewee was assigned a unique identification

number to protect his/her anonymity.

RESULTS

The 17 clinical researchers had diverse backgrounds, including otolaryngology, pharmacy, endocrinology, orthopedics, anesthesia, radiology, neurology, cardiology, nephrology, Hematology, thoracic surgery, epidemiology and biostatistics, clinical epidemiology, clinical research data management, and clinical research methodology. (Table1)

Table 1 Demographics of the participants

Participants level	Clinicians with limited clinical research experience	Clinicians with rich clinical research experience	Clinical research experts
Number	8	5	4
Gender			
Male (N)	2	2	3
Female (N)	6	3	1
Average age (mean± SD)	27±4	35±2	44±12
Department	Otolaryngology, Pharmacy, Endocrinology, Orthopedics, Anesthesia, Radiology, Neurologist, Cardiology.	Nephrology, Otolaryngology, Thoracic surgery, Hematology	Epidemiology Or Clinical research methodology
Experience in clinical research (projects)	Participating in clinical research work for 1-3 years	Being involved in 5-12 clinical research projects	Working in clinical research for 11-20 years
Data collection	4 semi-structured in-depth interviews, and 1 focus group interview for 4 clinicians	Semi-structured in-depth interviews	Semi-structured in-depth interviews

A typology of how clinical researchers of different skill levels include variables in CRF for clinical research

Based on our analysis, we developed a typology of how clinical researchers of different skill levels select variables in a CRF for clinical research. These three

models are illustrated as Figures 1, 2, and 3. The findings are supported by quoted comments from the research participants. The models are illustrated using a flow chart showing the different process for each of three levels of the participants, e.g., novice clinicians with limited clinical research experience, intermediate-level clinicians with rich clinical research experience, and clinical research experts.

a. The novice clinical researcher's approach to selecting variables in a CRF for clinical research.

Novice clinicians with limited clinical research experience described a multi-faceted approach to selecting variables for inclusion in their CRFs. When they planned the variables to include in their CRFs, most had no clearly defined comprehensive approach.

Finding the template from similar research and imitating it. Most novice clinical researchers mentioned that they would find a similar data collection form for reference and modify it according to the needs of their own study. As one endocrinologist noted, “First, I search on the Internet to find a template from similar research, then I modify the template based on my research”. One otolaryngology clinician said, “I ask for a data collection template from other experienced researchers in my department. I then imitate the template to make my own data collection template, and modify it in places according my research.” An Orthopedics clinician reflected his opinion that most novice clinical researchers are doing like this, “I will imitate another data collection template, then copy what is applicable to my research, and modify those variables that are not applicable. I think most of us follow, because it is easier for us.”

Discussing with clinicians with rich clinical research experience. Some novice clinical researchers reported discussing the CRF with experienced clinical researchers. One participant recounted, “I discussed the CRF with clinicians with rich experience in clinical research. They deleted useless items directly.”

Modifying again if any problems are found during use. Some novice clinical researchers directly used the data collection form without a pilot study, and modified it only when they found problems during use. An anesthesia clinician said, “When the

CRF is used to collect data in a clinical research project, some variables are found not to be applicable or absent. In this case, I always think about it and then modify/add the variables." Another cardiology clinician also remarked, "When I use the CRF to collect data from the first few participants, I find some variables can't get data at all. So I realize that these variables also can't get data from other participants, and these variables can contribute little to my research. Then I delete them."

The following flow chart was refined from the above in-depth interviews with four clinicians and a group interview with another four clinicians.

Figure 1. Novice clinical researchers' approach to developing a CRF for clinical research

b. The experienced clinical researcher's approach to selecting variables in a CRF for clinical research.

Experienced clinical researchers also described a multi-faceted approach for how they include variables in a CRF. Most of them have a clear purpose and protocol firstly, and think about the importance, feasibility and statistics of these variables according to their experience in clinical research.

Confirming the purpose and protocol. These clinical researchers first concern themselves with confirming the purpose and developing a protocol before designing the CRF. For example, a nephrologist said, "First, you should confirm the purpose of your study, disease features or prognosis, follow-up study or cross-sectional study."

Conducting a literature search for references and related materials. Most of the experienced clinical researchers would review the literature and related materials to consider the variables for inclusion in the CRF. "I think searching the related literature and other references is important. You can find many variables mentioned in other's research. You also can list all the related items for reference."

Finding a template from similar research and intimating it. Experienced clinical researchers also find a template to imitate for their own study. However, they consider the variables according to the research protocol, literature and their previous experience. One hematologist explained, "(I) search materials first, then find a

template from the literature or similar research previously conducted in my department. Then I imitate the template to make CRF for my project, and optimize and perfect the CRF according to the research protocol."

Consider the importance, feasibility and statistics of fields. The experienced clinical researchers note that the importance, feasibility and statistical analysis of collected variables should be considered before deciding whether to collect the data in the first place. One hematologist said, "Even if we do the optimizing and perfecting, we can also find many problems when we use it to collect data in clinical research. We can just modify it again and again. In the final version, many variables are deleted that are not easy to collect, or not very clear, or have little relationship with the primary outcome."

Conduct a pilot study and modify the CRF. These experienced researchers also note that piloting the CRF before using it to collect data is very important. If any problems are found in the pilot study, the problem can be resolved in a timely way. A thoracic surgeon shared, "A pilot study was conducted after the draft CRF was completed. We then recruited some patients to complete it. We found that some items or scales could not be completed. For example, the tinnitus handicap inventory is too complex for patients. And it is not very important for the final evaluation about the treatment, so we deleted it in follow-up visits."

The flow chart depicted in Figure 2 illustrates the overall steps employed by these experienced clinicians based on their interviews.

Figure 2. Experienced clinical researcher's approach to selecting variables in a CRF for clinical research.

c. The clinical research experts' approach to selecting variables in a CRF for clinical research.

Clinical research experts described their multi-faceted approach to identifying variables for inclusion in a CRF or how they help other researchers design a CRF. They would consider the study comprehensively, and they raised some different and important views.

Confirming the clinical and scientific issues, hypotheses, and research protocol. A clear research question and hypotheses are essential before you consider which variables should be included in the CRF. One expert working in the epidemiology research center stated, “First, you complete the top-level design of the clinical research to confirm the clinical issues, scientific issues and hypotheses, and then confirm the research purpose and protocol.”

PICO model is considered to build the causality model. The PICO Model is a format used to help define the clinical research question. P stands for Patient /Population/problem; I stands for Intervention/Prognostic factor /Exposure; C stands for Comparison; O stands for Outcome sought to measure or achieve[15]. The PICO model as commonly used in evidence-based medicine, can be employed to build a causality model and clearly confirm the related variables in the CRF. Another clinical research expert working in the epidemiology research center explained, “In clinical research, you need to confirm the PICO, and build the causality model based on the PICO model. Then confirm the primary outcome and secondary outcome based on O (outcome) in PICO model.” “You need to document the domains based on the PICO model. These domains include but are not limited to population characteristics, intervention or exposure and related outcome, and related confounding factors affecting the causality between the intervention and outcome.”

What's more, “few to many” and “many to few” are used to list and screen domains and variables in the research. Many clinical researchers worry about omitting key variables in the CRF, so they often list as many variables as possible. However, many variables are not necessary, and the added burden of collecting these data, may even lower the quality of key variables. One expert in the clinical research institute emphasized that, “'From less to more', you need to list variables from all domains as much as possible through multiple ways such as clinical practice experience and literature. ‘Many to few’, we should organize related experts including clinicians, statisticians, data manager and experts on clinical research design to discuss the necessity and feasibility of all the domains and variables. It should be considered in terms of accuracy, sensitivity and difficulty of detection and collection,

statistics, ethical problems, cost, quality control and so on." The flow chart depicted in Figure 3 illustrates the overall steps used by these 4 clinical research experts.

Figure 3. The expert clinical researcher's approach to selecting variables in a CRF for clinical research

DISCUSSION

This qualitative study reveals a different process of selecting variables for inclusion in a CRF for clinical research among three different levels of clinical researchers. Novice, experienced and expert clinical researchers have progressively sophisticated approaches for selecting variables in a CRF for clinical research. The novice clinician researcher approach of finding a template, modifying it according to their own study, reviewing it with clinicians with rich clinical research experience, and then completing the CRF when collecting data can be problematic. This process can truncate thinking about the whole protocol, study purpose, statistics, etc., and risks compromising the study quality. This point is consistent with previous discussion of Ioannidis et al[16] that research in a previously understudied domain might supply too little information to be useful. The resulting risk is, small uninformative studies that remain common in several specialties.

Relative to the existing literature about the development of CRFs, this research has expanded understanding about what variables should be included in a CRF, procedures used by expert researchers for first greatly increasing, and then limiting the range of variables to be studied. The notion of "few to many" and "many to few" to first comprehensively generate, then severely truncate to limit the domains is similar to previous advice. Li et al suggest collecting data so as to be able to just answer the scientific issues or hypothesis, nothing more and nothing less[7].

Based on views from the novice to the expert clinician, the take away message from this research is that increasing levels of sophistication in research planning reflect increasing levels of sophisticated selection of variables on the case report form. As alluded to by Chalmers et al, there are several principles that can help guide

clinical researchers to conduct efficient and high quality studies[17]:

(a) **Build the CRF to align closely with the research design, and develop and modify the protocol and template according to the feasibility of collecting data in clinical practice**. The risk of designing a CRF after completing the research protocol, is that some variables might not be feasible to collect as part of the clinical enterprise. The consequence of deviating from the protocol is the risk of compromised research quality. These findings further remind researchers to be certain to modify the CRF when the research protocol is modified.

(b) **Align the CRF with the research protocol.** The variables in the CRF should be consistent with protocol. That is, all domains and variables mentioned in the protocol should be included in CRF, to avoid missing any important domains or variables.

(c) **Develop the CRF to be concise and to the point.** In general, variables only related to the study purpose could be included in the CRF. Unnecessary or redundant variables can require unnecessary time, effort collecting and monitoring. This can distract from limited time and energy of researchers to ensure the quality of key variables in clinical research. Berge et al also identify that data collection forms can be too complex and burden centers with a requirement to collect data items that are never analyzed or reported[18].

(d) **Refine the template to be readily understood and operational.** In the final template of CRF, every variable should be clear. The structure and order of the CRF should be consistent with research flow and clinical practice.

These findings speak to fundamentals of conducting high quality research, and while based on research in a single institution, can reasonably be expected to hold true in a broad range of settings. This is consistent with our research expectations.

A potential limitation of our study concerns the risk of selection bias of our sample. Clinician participants are from a premier academic hospital in Beijing, and these participants might have greater knowledge, cognitive skills, and opportunities for research than clinicians in many other hospitals. We acknowledge the potential for variations on the findings in other settings, while also acknowledging the results are

genuine and representative of the participants' experiences. Despite variations that exist in practice, we believe the lessons learned based on this study are transferable and robust ideas. Another potential limitation is that the interviewers were aware of the interviewed researchers' experience and the study hypotheses, and this could raise concern about confirmation bias, however the models were built and reviewed by the researchers themselves and this suggests that the processes delineated represent their views.

Further studies are needed to assess whether the approaches and variations described are robust across a wider variety of settings. The proposed expert model could be further validated by applying it to design of CRF for clinical research in different levels of hospitals and clinicians from other settings. Regardless, we believe that while there may be additional considerations, these findings are a robust and meaningful reference for any clinicians engaged in clinical research, and particularly, novice and moderately experienced researchers who wish to learn lessons from experts.

In conclusion, this research illustrates that increasing levels of sophistication in research planning reflect increasing levels of sophistication in the selection of variables for inclusion in a case report form. Thus, novice and intermediate-level clinical researchers alike could benefit by emulating the comprehensive planning procedures utilized by expert clinical researchers.

Acknowledgements: We sincerely appreciate all the participants who made this research possible.

Funding: This study was funded by grant 2013BAI09B14 from the Ministry of Science and Technology of the People's Republic of China. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Contributors: Chu Hongling and Zeng Lin conducted the interviews and all analyses, wrote the first draft of the manuscript and rewrote new drafts based on input from co-authors. Michael Fetters provided extensive input on the structuring of the manuscript, and revisions of the manuscript drafts. Zhao Yiming designed the

research project, planned the analyses and provided input and revision of manuscript drafts. All authors read and approved the final manuscript.

Competing interests: None to declare.

Participant consent. Obtained written consent.

Ethics approval: The Peking University Third Hospital Ethics Review Board approved the study.

Provenance and peer review: Not commissioned; externally peer reviewed.

Data sharing statement: Original quotes and audio data of interviews are available on request by emailing the corresponding author. The presented data were anonymized and the risk of identification is low.

References:

1. Zhao YM, Ding J, Zeng L. Top level design on clinical research. National Medical Journal of China. 2010;90(14):942-946. Chinese.

2. Nahm M, Shepherd J, Buzenberg A, Rostami R, Corcoran A, McCall J, et al. Design and implementation of an institutional case report form library. *Clin Trials* 2011;8(1):94-102.

3. Cao AP, Su BH, Ding DY, et al. The case report form in clinical research. Chinese Journal of Health Statistics. 2000(03):61-64. Chinese.

4. Fu, H.J. and X.X. Luo, Key factors in design of case report form. Yao Xue Xue Bao, 2015. 50(11): p. 1452-1455. Chinese.

5. Li QN, Lu F, An FH, Gao R. The design of case report form and analysis of common related problem. Chin J Clin Pharmacol Ther 2013(08):901-906. Chinese.

6. Li ST. The design of case report form and its importance in drug clinical trials. Chines Journal of New Drugs 2004(01):53-55. Chinese.

7. Li XY, Wen ZH, Tang XC, Liang WX. The principle and process of case report form design in clinical trials. Traditional Chinese Drug Research &Clinical Pharmacology 2013(02):206-209. Chinese.

8. Lou H. The design and application of case report form in evidence-based traditional Chinese medicine research. Shandong Traditional Chinese Medicine University, 2015. Chinese.

9. Lv XY, Zhang ZL, Ai YK, et al. Talking about the design of case report form from the point of data management. Modernization of Traditional Chinese Medicine and Materia-World Science and Technology 2014(03):614-617. Chinese.
10. Wan X, Yang H, Liu JP. The design of case report form in clinical research. Journal of Traditional Chinese Medicine 2007(10):885-887. Chinese.
11. Wang YH, Jiao YZ. The design of case report form and analysis of related problem in traditional Chinese medicine research. China Journal of Traditional Chinese Medicine and Pharmacy 2005(10):620-623. Chinese.
12. Meyer J. Qualitative research in health care. Using qualitative methods in health related action research. *BMJ* 2000;320(7228):178-81.
13. Devers KJ, Frankel RM. Study design in qualitative research--2: Sampling and data collection strategies. *Educ Health (Abingdon)* 2000;13(2):263-71.
14. Elo S, Kyngas H. The qualitative content analysis process. *J Adv Nurs* 2008;62(1):107-15.
15. Kleijnen J, Chalmers I. How to practice and teach evidence-based medicine: role of the Cochrane Collaboration. *Acta Anaesthesiol Scand Suppl* 1997;111:231-3.
16. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383(9912):166-75.
17. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;383(9912):156-65.
18. Berge E, Salman RA, van der Worp HB, Stapf C, Sandercock P, Sprigg N, et al. Increasing value and reducing waste in stroke research. *Lancet Neurol* 2017;16(5):399-408.

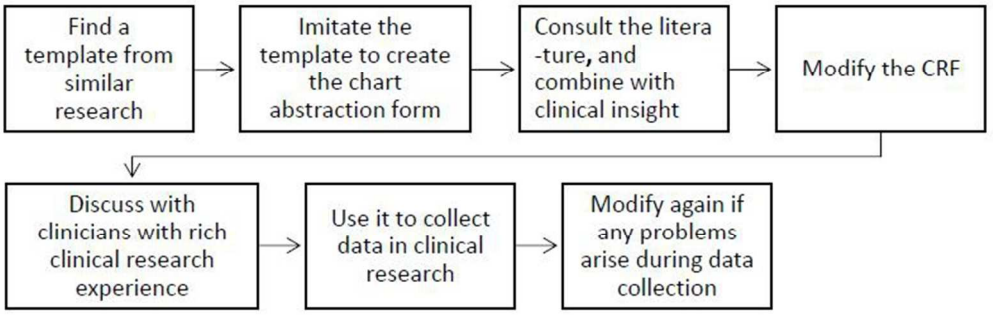


Figure 1. Novice clinical researchers' approach to developing a CRF for clinical research

67x21mm (300 x 300 DPI)

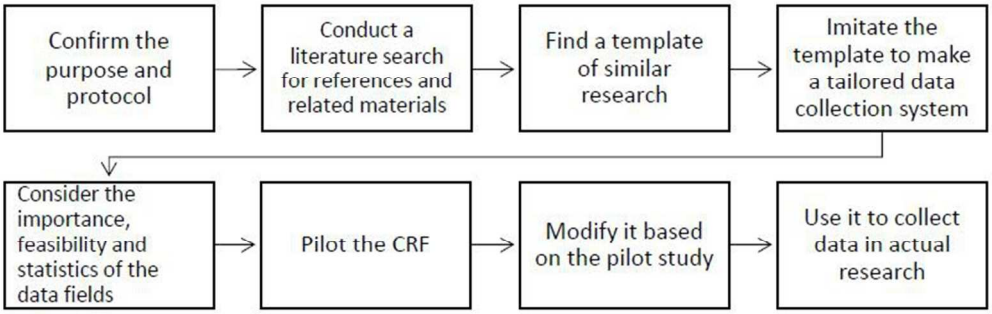


Figure 2. Experienced clinical researcher's approach to selecting variables in a CRF for clinical research.

68x22mm (300 x 300 DPI)

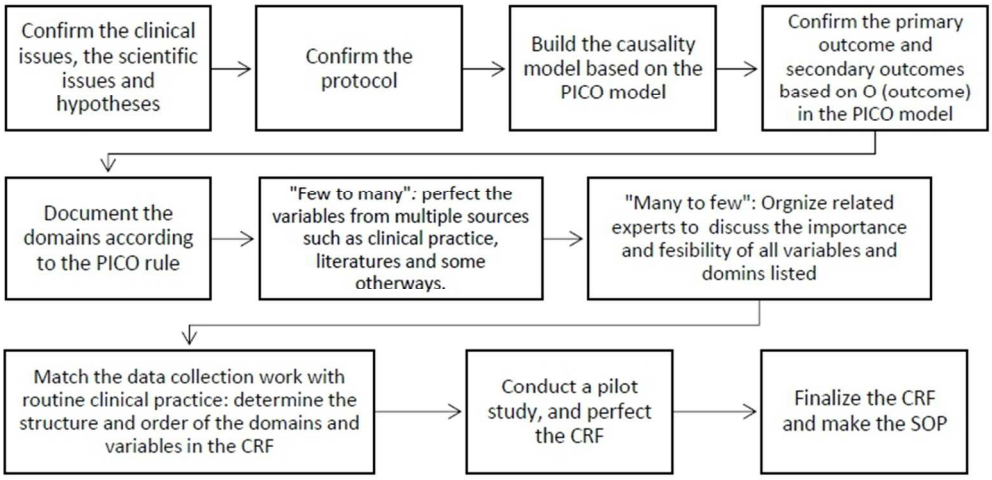


Figure 3. The expert clinical researcher's approach to selecting variables in a CRF for clinical research

70x34mm (300 x 300 DPI)

Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist		
No Item	**Guide questions/description**	**Page number**
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer/facilitator	Which author/s conducted the interview or focus group?	5
2. Credentials	What were the researcher's credentials? E.g. PhD, MD	5
3. Occupation	What was their occupation at the time of the study?	5
4. Gender	Was the researcher male or female?	5
5.Experience and training Relationship with participants	What experience or training did the researcher have?	5
6.Relationship established	Was a relationship established prior to study commencement?	5
7.Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	N
8. Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	N
Domain 2: study design		
Theoretical framework		
9. Methodological orientation and Theory	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	6
Participant selection		
10. Sampling	How were participants selected? e.g. purposive, convenience, consecutive, snowball	4
11. Method of approach	How were participants approached? e.g. face-to-face, telephone, mail, email	6
12. Sample size	How many participants were in the study?	4
13. Non-participation	How many people refused to participate or dropped out? Reasons?	N
Setting		
14. Setting of data collection	Where was the data collected? e.g. home, clinic, workplace	6
15. Presence of non-participants	Was anyone else present besides the participants and researchers?	6
16. Description of sample	What are the important characteristics of the sample? e.g. demographic data, date	7
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	5

18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	6
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	6
20. Field notes	Were field notes made during and/or after the interview or focus group?	N
21. Duration	What was the duration of the interviews or focus group?	6
22. Data saturation	Was data saturation discussed?	6
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	N
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	6
25. Description of the coding tree	Did authors provide a description of the coding tree?	N
26. Derivation of themes	Were themes identified in advance or derived from the data?	8
27. Software	What software, if applicable, was used to manage the data?	6
28. Participant checking	Did participants provide feedback on the findings?	6
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes / findings? Was each quotation identified? e.g. participant number	8--12
30. Data and findings consistent	Was there consistency between the data presented and the findings?	8--13
31. Clarity of major themes	Were major themes clearly presented in the findings?	8--13
32. Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	8--13