CASE STUDY

Selection of the Questions: Finding the Research Question

Andre Brunoni & Felipe Fregni[1]

Dr. L. Heart is a scientist working on cardiovascular diseases in a large, busy emergency room of a tertiary hospital specialized in acute coronarian syndromes. While searching PUBMED, she found an interesting article on a new drug - which animal studies have demonstrated to be a powerful anti-thrombotic agent - showing its safety in healthy volunteers. She then feels that it would be the right time to perform a phase II trial testing this new drug in patients presenting myocardial infarction. She sees this as her big career breakthrough. However, when Dr. Heart starts writing a study proposal for the internal review board (ethics committee), she asks herself: What is my research question?

Introduction

Defining the research question is, perhaps, the most important part of the planning of a research study. That is because the wrong question will eventually lead to a poor study design and therefore all the results will be useless – on the other hand, choosing an elegant, simple question will probably lead to a good study that will be meaningful to the scientific community even if the results are negative. In fact, the best research question is the one that, regardless of the results (negative or positive), produces interesting findings. In addition, a study should be designed with only one main question in mind.

---

[1] Fregni, F. & Illigens, B.M.W. Selection the Research Question. In: Critical Thinking in Clinical Research: Applied theory and Practice Using Case Studies. Oxford Pres, 2017, p38. Dr. Andre Brunoni and Professor Felipe Fregni prepared this case. Course cases are developed solely as the basis for class discussion. The situation in this case is fictional. Cases are not intended to serve as endorsements or sources of primary data.

However, choosing the most appropriate question is not always easy as such a question might not be feasible to be answered. For instance, when researching acute myocardial infarction (MI), the most important question would be whether or not a new drug decreases mortality.

However, for economic and ethical reasons, such an approach can only be considered when previous studies have already *suggested* the new drug is a potential candidate. Therefore, the investigator needs to deal with the important issue of feasibility vs. clinical relevance. Dr. Heart soon realized that her task would not be an easy one and also that this task may take some time; she kept thinking back about one of the citations in an article she recently read *"One third of a trial's time between the germ of your idea and its publication in the New England Journal of Medicine should be spent fighting about the research question."*[2]

"So what?" - Test For the Research Question

Dr. Heart knows that an important test for the research question is to ask: "so what?" In other words, is the researching question addressing an important issue? She knows for example that the main agency funding in the USA, NIH (National Institutes of Health), considers significance and innovation as important factors to fund grant applications. Dr. Heart also remembers something that her mentor used to tell her at the beginning of her career: "A house built on a weak foundation will not stand". She knows that even if she has the most refined design and uses the optimal statistical tests, her research would be of very little interest or utility if it does not advance the field. But regarding this point, she is confident that her research will have a significant impact in the field.

---

[2] Riva JJ, Malik KM, Burnie SJ, Endicott AR, Busse JW. What is your research question? An introduction to the PICOT format for clinicians. J Can Chiropr Assoc. 2012 Sep;56(3):167-71.

Next Step for the Research Question: How to Measure the Efficacy of the Intervention?

Dr. Lonely Heart is in a privileged position. She works in a busy hospital that receives a significant amount of acute cardiovascular patients. She has also just received huge departmental support for her research, meaning that she can run a wide range of blood exams to measure specific biological markers related to death in myocardial infarction. Finally, she has a PhD student who is a psychologist working with quality of life post-MI. Therefore, she asks herself whether she should rely on biological markers, on the assessment of quality of life, or if she should go to a more robust outcome to prove the efficacy of the new drug. She knows that this is one of the most critical decisions she has to make. It was a Friday afternoon. She had just packed up her laptop and the articles she was reading knowing that she will have to make a decision by the end of the weekend.

Dr. Heart is facing a common problem: what outcome should be used in a research study? This needs to be defined for the research question. She knows that there are several options. For instance, the outcome might be mortality, new myocardial infarction, days admitted to the emergency room, quality of life, a specific effect of disease such as angina, a laboratory measure (cholesterol levels), or the cost of the intervention. Also, she might use continuous or categorical outcomes. For instance, if she is measuring angina, she might measure the number of days with angina (continuous outcome) or dichotomize the number of angina days in two categories (less than 100 days with angina vs. more or equal to a 100 days with angina). She then lays out her options:

- − Use of clinical outcomes (such as mortality or new myocardial infarction): she knows that by using this outcome, her results would be easily accepted by her colleagues; however, using these outcomes will increase the trial duration and costs.
- − Use of surrogates (for instance, laboratorial measurements): one attractive alternative for her is to use some biomarkers or radiological exams (such as a catheterism). She knows a colleague in the infectious disease field that only uses CD4 for HIV trials as the main outcome. This would increase the trial feasibility. However, she is concerned that her biomarkers might not really represent disease progression.

− *Use of quality of life scales:* this might be an intermediate solution for her. However, she is still concerned with the interpretation of the results if she decides to use quality of life scales.

More on the response variable: *categorical or continuous?*

Even before making the final decision, Dr. Heart needs to define if she is using a continuous or categorical variable. She wishes now that she knew the basic concepts of statistics. However, she calls a colleague that explains to her the main issue of categorical vs. continuous outcomes – in summary, the issue is the trade-off of power vs. clinical significance.

A categorical outcome usually has two categories: for instance, a yes/no answer, while a continuous outcome can express any value. A categorical approach might be more robust than a continuous one, and it also has more clinical significance, but it also decreases the power of study due to the use of less information[3]. She is now at the crossroad of feasibility vs. clinical significance.

Choosing the study population

Now that Dr. Heart went through the difficult decision of finding the best outcome measure, she needs to define the target population – i.e., in which patients is she going to test the new drug? Her first idea is to select only patients that have a high probability of dying – for instance, males who smoke, older than 75 years, with insulin-dependent diabetes and hypercholesterolemia. "Then", she thinks, "it will be easier to prove that the new drug is useful regardless of the population she studies. But does that really sound like a good idea?"

The next step is to define the target population. Dr. Heart is inclined to restrict the study population, as she knows that this drug might be effective to a particular population of patients and therefore increasing her chances of getting a good result. In addition, she did

[3] These concepts will be discussed in details in the statistical module.

remember from her statistical courses that this would imply a smaller variability and therefore she would gain power (power is an important currency in research as it makes the study more efficient, decreasing costs and time to complete the study) – which makes her thrilled. On the other hand, she is concerned that she might put all her efforts in one basket – this is a risky approach as this specific population might not respond and she knows that broadening the population also has some advantages, for instance, the results would be more generalizable and it would be easier to recruit patients. But this would also increase the costs of the study.

But How About Other Ideas?

After the weekend of reflection, Dr. Heart called the staff for a team meeting and proudly explained the scenario and stated her initial thoughts. The staff was very eager to start a new study, and they made several suggestions: "We should also use echocardiography to assess the outcome!" "Why don't we perform a genotypic analysis on these patients?" "We need to follow them until one year after discharge". She started to get anxious again. What should she do with these additional ideas? As they all sound like good ideas.

When designing a clinical trial, researchers expose a number of subjects to a new intervention. Therefore, they want to extract as much data as possible from studies. On the other hand, it might not be possible to ask all of the questions, since this will increase the study duration, money and personnel. Also, researchers should be aware that all the other outcomes assessed will be exploratory, i.e., their usefulness remains in suggesting possible associations and future studies. That is because studies are designed to answer a primary question only – and, as a principle of statistics, there is a 5% probability of observing a positive result just by chance – if you perform 20 tests, for instance, one of them will be positive just by chance! But Dr. Heart knows that she can test additional hypotheses as secondary questions. She knows that there is another issue to go through: the issue of primary vs. secondary questions.

Defining Her Hypothesis

After going through this long process, Dr. Heart is getting close to her research question. But now she needs to define the study hypotheses. In other words, what is her educated guess regarding the study outcome?

An important step when formulating a research question is to define the hypothesis of the study. This is important as to design the analysis plan and also to estimate the study sample size. Usually researchers come up with study hypotheses after reviewing the literature and preliminary data. Dr. Heart can choose between a simple and a complex hypothesis. In the first case, her hypothesis would only have one dependent variable (i.e., the response variable) and one independent variable (i.e., the intervention, for instance). Complex hypotheses have more than one independent and/or dependent variable and might not be easy to use to plan the data analysis.

By the end of the day, Dr. Heart was overwhelmed with the first steps to put this study together. Although she is confident that this study might be her breakthrough and she is needing to get her tenure track position at the institution she works at, she also knows she has only one chance and needs to be very careful at this stage. After wrestling with her thoughts, she finished her espresso and walked back to her office confident that she knows what to do.