Common Pitfalls in Research Design and Its Reporting

3

Fikri M. Abu-Zidan

3.1 Introduction

Being involved in reviewing articles for high impact surgical journals for more than 25 years, I have repeatedly encountered certain errors in research methodology and statistical analysis regardless of the origin of the manuscripts, whether stemming from developed or developing countries. These errors can be easily avoided by asking for advice and proper planning. Some of these errors, although seem trivial, can be fatal because they cannot be saved retrospectively. Occasionally researchers may concentrate so much on the details, technicality, and complexity while missing the overall picture. That is similar to visualizing a sky tower or reading a chest X-ray. Details can be missed either because you are so far from it, or alternatively so close to it. Taking care of the overall aim and structure of a research project is as important as looking into the small details. This chapter aims to highlight some common research design and reporting errors, hoping that they will be avoided when performing a research project.

Learning Objectives

- Highlight the importance of properly defining a focused research question.
- Stress the importance of involving a research methodologist in the research project.

F. M. Abu-Zidan (✉)
The Research Office, College of Medicine and Health Sciences, United Arab Emirates University, Al-Ain, United Arab Emirates

M. Ceresoli et al. (eds.), *Statistics and Research Methods for Acute Care and General Surgeons*, Hot Topics in Acute Care Surgery and Trauma,
https://doi.org/10.1007/978-3-031-13818-8_3

- Recognize that fatal errors in research design include using invalid measurement tools and testing the wrong population.
- Recognize the difference between correlation and prediction.
- Understand the difference between clinical and statistical significance.
- Report the data properly and describe how to deal with missing data.

3.2 Unclear Research Question

The main research question is "What do you want to find in your study?" If this question is not focused, it will be difficult to have a proper plan (map) to reach that aim. I think that a proper research question is the most important component for a research project. Let us give a practical example. We know that road traffic collisions cause death. We may ask ourselves what causes this death. This may be caused by speed, slipping of a car in a rainy weather, distraction of the driver when using a cell phone, or not using a seatbelt. Real life situations are complex, and we will not be able to answer all these questions at the same time in a single study. You have to define exactly what you want to study. Selecting a wrong research question makes the whole study flawed. This is similar to horse racing in which the eyes of the horse are covered by eye blinkers so the horse can go only in one direction (forward) to win the race (Fig. 3.1). If the blinkers are removed, the horse will look around and slow down. Accordingly, the researcher should spend significant time to define the aim of the study and concentrate on answering it. After reaching the first aim, then the researcher can remove the eye blinkers, look around, and think of his/her next target, and so on. Each question will generate multiple new questions. It is then the duty of a good researcher to define the next important, relevant, and feasible question to answer. I personally aim at answering one question in each study. I discourage my research students to have a multifactorial design in which they try to answer more than one question because this approach has the risk of not being able to

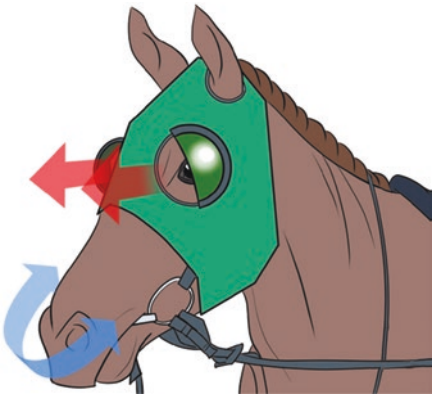

Fig. 3.1 Researchers should be like racing horses in their research in which the eyes of horses are covered by eye blinkers so that they can follow the racing track in one direction (forward) to win the race without being distracted. (*Illustrated by Mohammad F. Abu-Zidan*)

answer any of the questions. That is logical because having multiple questions to answer at the same time needs extreme care in the methods to be able to answer all questions. Making the aim more focused makes the methods simpler, direct, and more precise.

3.3 Lack of Planning *(Failing to Plan Is Planning to Fail)*

Genuine time should be spent in designing and planning a study before it starts. Involving a methodologist at this early stage will avoid errors and improve the chance of accepting a scientific paper. Submitted papers without involvement of a methodologist are more likely to be directly rejected without sending them to the reviewers. Even if they were sent for review, they are more likely to be rejected [1]. Inappropriate statistics and overinterpretation of the results are the most common causes for paper rejections [2]. Methodologists can be involved in the whole process of research including formulating the research question, research design, research audit, analyzing the data, participating in writing the manuscript, critically reading it, and finally approving it [3]. I will give a personal practical example highlighting this important point. Twenty years ago, I developed an experimental animal model for training Focused Assessment Sonography of Trauma (FAST) [4]. Designing and planning this study took 2 months, while performing the animal experiments and collecting the data took only 2 days. The paper was reviewed and accepted in less than 3 weeks. The ratio between the design/planning: performing the study in this example was 30:1. Although this may be an extreme example, it highlights the importance of thinking deeply and discussing the study with a methodologist to finalize the research design and plan for executing it.

3.4 Using the Wrong Research Tool

When measuring outcome variables in a research study, you need the proper tools to accurately measure these variables. Let us assume that you want to measure the mean arterial pressure in a critically ill septic patient in the intensive care unit. There are important characteristics in the measurement tool that have to be fulfilled. These are: (1) The tool should be *valid*, which means that it can measure the mean arterial pressure. Using a thermostat to measure the mean arterial pressure is not valid; (2) It should be *accurate*, this means that it will measure the real value; (3) It should be reliable, this means that it will give the same result if the measurements are repeated. Kindly note that accuracy is different from reliability. You may get the same result when the tool is reliable but this may cause a systematic error if it is not accurate. The most serious error in the study is using an invalid tool. This error cannot be corrected after finishing the study or experiment and will spoil the whole experiment.

It is common that acute care surgeons use surveys in their research. Although surveys look easy to perform and collect information about needs assessment, they are very tricky. They are not simply sending few questions and collecting the answers. It is important that these surveys should be valid and reliable. A lot of attention should be taken to have simple, clear, useful, well understood, and precise questions in these questionnaires [5].

3.5 Selecting the Wrong Population

This is a very fatal mistake that should be avoided. Any experiment or intervention should be tested in the population that are expected to benefit from it. I have repeatedly encountered clinical studies that aim to investigate a diagnostic test for a certain disease and then studied it in a population that has the final diagnosis of that disease (prior probability of almost 100%). An example of that is studying the role of ultrasound in diagnosing acute appendicitis. The authors studied ultrasound only in those operated (prior probability of 90%). That is the wrong population to be studied because ultrasound should be tested in those suspected to have appendicitis and not those already decided to be operated on. What is the value of ultrasound if you have already decided for surgery?

Let us have another example of an interventional procedure. Assume that we are going to study the role of Resuscitative Endovascular Balloon Occlusion of the Aorta (REBOA) in trauma abdominal bleeding patients. Figure 3.2 shows what is called the therapeutic window of an intervention. If REBOA was used in those having very mild disease, then it will be harmful (line is horizontal, no benefit). Similarly, if severity is above a certain limit (the line is also horizontal with severe injury), then it may not be useful. It is then very important to carefully select the population that may benefit from an intervention to be properly tested.

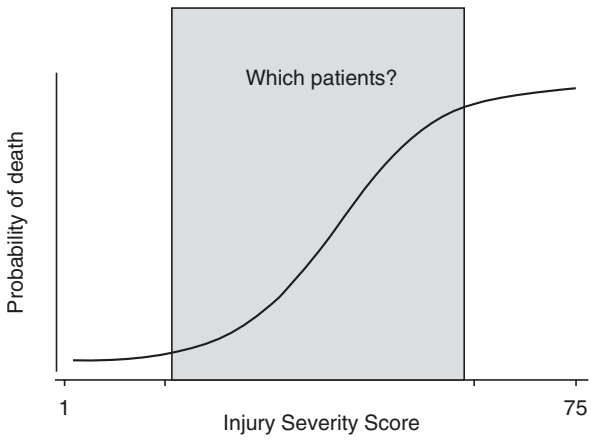

Fig. 3.2 An illustration showing the principle of the therapeutic window when using Resuscitative Endovascular Balloon Occlusion of the Aorta (REBOA). REBOA will be harmful if used in mild injured patients (lower horizonal line). It will not be useful if used above a certain limit (upper horizontal line)

3.6 Addressing the Missing Data

The authors have to be transparent regarding handling their missing data [3]. Prospective studies should generally have missing data of less than 10%. Retrospective studies usually have more missing data (up to 30%). Missing data are usually not random in high risk situations (like death) and may affect the analysis. Patients who die especially in the Emergency Department tend to have more missing data.

If imputations are used to replace the missing data (usually for retrospectively collected data), the authors have to justify this approach and demonstrate that missing data were random. This can be addressed by demonstrating that: (1) the groups have the same percentage of missing data before imputation for each studied variable and (2) they were statistically similar before the imputation. Our Trauma Group follows a school that does not replace missing data because this depends on assumptions which may increase the uncertainty in our statistical findings. We found that the best approach in establishing our trauma registry is to collect data prospectively by trained researchers and regularly audit the data which increased the trust in our data [6, 7].

3.7 Correlation and Prediction

There is great difference between correlation and prediction which should be clear. Correlation addresses the relationship between two variables regardless of whether one of them depends on the other. The correlation (association) does not imply a cause–effect relationship or the sequence in which they happen [8]. In comparison, prediction tries to define the outcome of one variable (dependent factor) depending on one or more factors (independent factors). The size of the p value does not reflect the strength of the correlation. Statistical significance having a small p value can occur when the sample size is large despite a weak correlation [9, 10]. Although there may be a statistically significant correlation, this may not be a strong correlation and the variable cannot be used as a predictor (Fig. 3.3). Predictors for important clinical outcomes, which can affect serious decisions, should be strong and simple to be useful in clinical practice.

The test for defining the correlation depends on whether the data have a normal distribution or not. When the data have a normal distribution, then Pearson's correlation test can be done. If the data are ordinal or do not have a normal distribution then Spearman's rank correlation test should be performed [8]. Figure 3.4 demonstrates this point. Spearman's rank correlation was used because the Likert type scale has ordinal data of 1–7. This analysis correlates the ranks and not the actual numbers. The scatterplot clearly shows that Pearson's correlation cannot be used in this scenario.

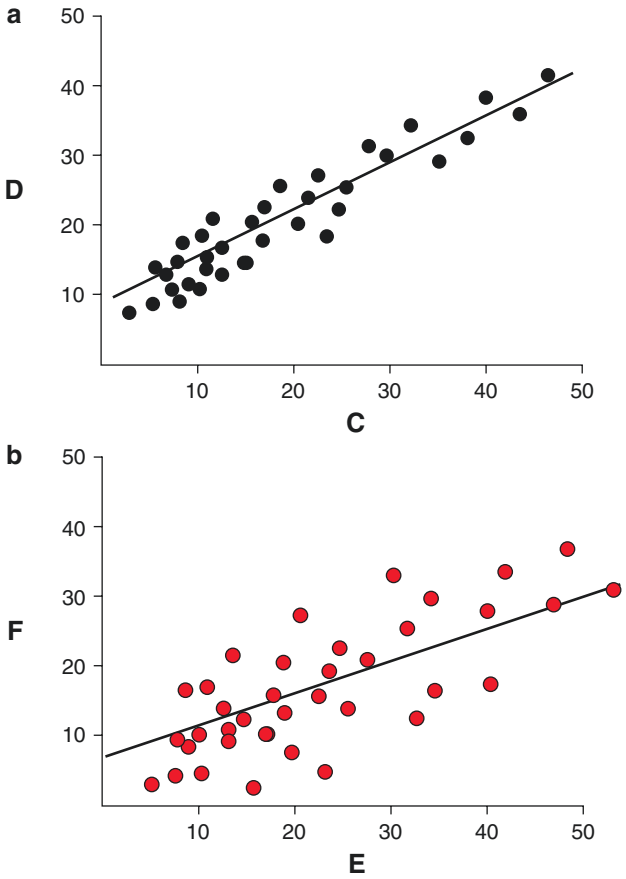

Fig. 3.3 This figure illustrates two situations of significant correlations; one of them has a strong correlation (**a**) that may be used for prediction, while the other has a weak correlation (**b**). Kindly notice the distribution of the data points around the correlation line and the slope of the line in each situation. Statistical significance with a small *p* value can occur despite a weak correlation

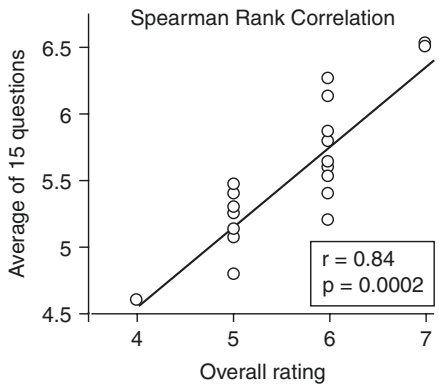

Fig. 3.4 The shown figure of the data of a Likert type scale has ordinal data of 1–7. The scatterplot clearly shows that Pearson's correlation cannot be used in this scenario. Spearman's rank correlation test should be used because the data do not have a normal distribution. This test correlates the ranks and not the actual numbers

3.7.1 Statistical and Clinical Significance

It is very important to be aware of the difference between statistical and clinical significance. We should not look through the pinhole of the *p* value but concentrate on the clinical implications of the statistical findings (Fig. 3.5). The "*p*" value estimates the probability that the reported result occurred by chance. It does not show the difference in the mean nor its direction. In contrast, confidence intervals can show the effect size, the direction of the change, the precision of the findings besides the statistical significance [11, 12].

Although statisticians can perform advanced analysis, clinicians may have more in-depth understanding of what do the findings mean because they are aware of their clinical importance and implications. Clinical significance depends on its effect on the existing clinical practice. When the sample size is large, there may be highly statistically significant findings but these may not translate to an effect size that can change clinical practice [13]. Occasionally when statisticians lead the clinical research, they may not appreciate the difference between dependent and independent factors if they do not have a clinical background or have close interactions with clinicians. I have personally reviewed articles in which the analysis tried to predict a clinically independent factor from a dependent factor which should be the opposite. Statisticians and clinicians should work together as one team before starting the clinical studies in designing the research protocol, during the study, and after completion of the study up to its publication. Team work is very important for acute care surgery including its research.

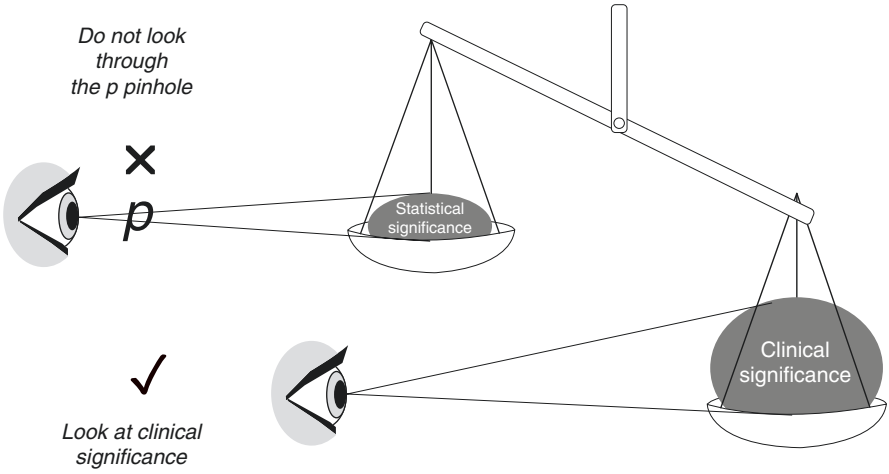

Fig. 3.5 We should not look through the pinhole of the "*p*" value but concentrate on the clinical implications of the statistical findings. The "*p*" value only estimates the probability that the reported result occurred by chance. It does not show the effect size nor its direction. (*Illustrated by Mohammad F. Abu-Zidan*)

3.8 Reporting of the Data

Accuracy and completion of the published statistical data will have long term implications in the future. Michalczyk and Lewis found that nearly half of the studies published in the Journal of Medical Education did not report enough statistical data [14]. Other researchers may need to compare the data with their own or pool the results in future systematic reviews. For example, it is not enough to report the mean alone without its variation (the standard deviation). In 2011, I had a disappointing personal experience trying to perform a systematic review on internal fixation of flail chest. After performing a lengthy detailed search, we could locate only two randomized controlled trials which were ready for the analysis [15, 16]. The paper of Granetzny et al. [15] reported only the mean without the standard deviation. We tried to contact the authors to get this data but we failed. Missing this simple data aborted the systematic review. When the mean and standard deviation of an independent variable of a group is given with the sample size, then it is possible to compare it or pool it with other studies [11]. It is becoming now a requirement in some highly ranked journals to publish the set of data that generated the results.

Do and Don't

- Define a focused, relevant, important, and feasible research question to answer.
- Plan your study properly with the help of a methodologist before you start data collection.
- Use valid tools to measure the outcome variables in the proper population.
- Report your data accurately.
- Concentrate on the clinical significance of your findings.
- Don't interpret a correlation relationship as a predictor.
- Don't ignore the missing data of your study.

Take Home Messages

- The research question is the most important pillar of a study.
- Failing to plan is planning to fail.
- Be transparent and accurate in using your research tools and reporting your results.
- Use your clinical sense.

Conflict of Interest None declared by the author.

References

1. Altman DG, Goodman SN, Schroter S. How statistical expertise is used in medical research. JAMA. 2002;287(21):2817–20.
2. Bordage G. Reasons reviewers reject and accept manuscripts: the strengths and weaknesses in medical education reports. Acad Med. 2001;76(9):889–96.

3. Clark GT, Mulligan R. Fifteen common mistakes encountered in clinical research. J Prosthodont Res. 2011;55(1):1–6.
4. Abu-Zidan FM, Siösteen AK, Wang J, Al-Ayoubi F, Lennquist S. Establishment of a teaching animal model for sonographic diagnosis of trauma. J Trauma. 2004;56:99–104.
5. Cummings SR, Hulley SB. Designing questionnaires and interviews. In: Hulley SB, Cummings SR, Browner WS, Grady DC, Newman TB, editors. Designing clinical research. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2007. p. 241–55.
6. Shaban S, Eid HO, Barka E, Abu-Zidan FM. Towards a national trauma registry for the United Arab Emirates. BMC Res Notes. 2010;3:187.
7. Shaban S, Ashour M, Bashir M, El-Ashaal Y, Branicki F, Abu-Zidan FM. The long term effects of early analysis of a trauma registry. World J Emerg Surg. 2009;4:42.
8. Chan YH. Biostatistics 104: correlational analysis. Singapore Med J. 2003;44:614–9.
9. Guyatt G, Walter S, Shannon H, Cook D, Jaeschke R, Heddle N. Basic statistics for clinicians: 4. Correlation and regression. CMAJ. 1995;152:497–504.
10. Aggarwal R, Ranganathan P. Common pitfalls in statistical analysis: the use of correlation techniques. Perspect Clin Res. 2016;7:187–90.
11. Bailar JC 3rd, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. Ann Intern Med. 1988;108:266–73.
12. Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: “P” values, statistical significance and confidence intervals. Perspect Clin Res. 2015;6:116–7.
13. Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: clinical versus statistical significance. Perspect Clin Res. 2015;6:169–70.
14. Michalczyk AE, Lewis LA. Significance alone is not enough. J Med Educ. 1980;55:834–8.
15. Granetzny A, Abd El-Aal M, Emam E, Shalaby A, Boseila A. Surgical versus conservative treatment of flail chest. Evaluation of the pulmonary status. Interact Cardiovasc Thorac Surg. 2005;4:583–7.
16. Tanaka H, Yukioka T, Yamaguti Y, Shimizu S, Goto H, Matsuda H, Shimazaki S. Surgical stabilization of internal pneumatic stabilization? A prospective randomized study of management of severe flail chest patients. J Trauma. 2002;52:727–32.