Statistics Corner

How to design a randomized clinical trial: tips and tricks for conduct a successful study in thoracic disease domain

Francesco Guerrera[1], Stéphane Renaud[2], Fabrizio Tabbò[3], Pier Luigi Filosso[1]

[1]Department of Thoracic Surgery, University of Torino, Torino, Italy; [2]Department of Thoracic Surgery, Nancy University Hospital, Nancy, France; [3]Department of Oncology, University of Torino, Torino, Italy

Correspondence to: Prof. Pier Luigi Filosso, MD. Department of Thoracic Surgery, Azienda Ospedaliera Universitaria Città della Salute e della Scienza di Torino, University of Torino, Corso Dogliotti 14, Torino 10126, Italy. Email: pfilosso@unito.it.

Abstract: Randomized controlled trials (RCTs) are considered one of the highest level of evidence in clinical practice, due to their strong confidence and robustness in producing data. The "randomization" (e.g., allocating patients randomly in each group of the study) allows eliminating many pre-analytical differences that might bias the entire study. Nevertheless, RCTs aren't free of internal pitfalls that might make them not easy to be developed or utilized. Our objective is to explain RCT management difficulties and suggest certain tips useful for the design of a RCT in thoracic disease domain. In particular have a realistic timeline, define a clear objective and precise endpoints, balance the study with a correct randomization and focus on the right equilibrium between strict selection criteria and more heterogeneous parameters are key elements that help researchers assuring a strong scientific validity.

Keywords: Randomized controlled trial (RCT); clinical research; study design; thoracic disease

Submitted Apr 16, 2017. Accepted for publication Jun 04, 2017.
doi: 10.21037/jtd.2017.06.147
View this article at: http://dx.doi.org/10.21037/jtd.2017.06.147

Introduction

Nowadays, medical decisions such as which type of surgical approach, whether to treat or not a patient and with which pharmacological intervention are evaluated considering the evidence-based medicine (1). In medicine, levels of evidence, as described by the National Cancer Institute, are arranged in "a ranking system used to describe the strength of the results measured in a clinical trial or research study. The design of the study and the endpoints measured affect the strength of the evidence" (2). So far different classifications have been proposed to classify levels of evidence. Among them, the United States Preventive Service Task Force (USPSTF) classified levels of evidence from level I (evidence obtained from at least one properly designed randomized trial) to level III (opinions of respected authorities based on clinical experience, descriptive studies, or reports of expert committees) (3). Another example is the Oxford CEBM levels of evidence, which ranges from level 1a (systematic reviews of randomized controlled trials) to level 5 (expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles") (4).

Regardless the chosen classification, randomized clinical trials (RCTs) are considered as one of the highest level of evidence in clinical practice, due to their strong confidence and robustness in producing data. Allocating patients randomly in each considered groups of the study, the "randomization" allows eliminating many pre-analytical differences that might bias the entire study. This is one of the major aspect that contribute to the robustness of this type of clinical trial and one of the reasons why RCTs have been extensively adopted by clinician-trialists.

Therefore, the ability, not only to properly interpret, but also to design correctly a RCT is mandatory if we aim to obtain a high-level clinical trial, which may impact on the scientific community.

Definition, strengths and limitations

RCTs are usually utilized to assess treatments' efficacy or effectiveness, in term of "superiority" (to determine if

a novel treatment is better than a placebo or a standard intervention), "non-inferiority" (the novel treatment is no worse than a placebo or a standard intervention) or "equivalence" (analogous to non-inferiority, but permit of the novel intervention is no better than a placebo or a standard intervention) (5). These studies may analyze a comparison between treatment and not treatment or placebo, different treatment strategies or different dosage/intensity of the same treatment. The term "randomized" defines the random assignment to each study group. This implies a balance in baseline characteristics (known and unknown) amongst patients' groups, reducing confounding factors and improving internal validity of the results (6).

On the other hand, the "ideal condition" to correctly perform RCTs determines its proper limitations: the population analyzed in RCT is for definition "selected" and obviously differs from the usual care population (7,8). Finally, RCTs are performed in a clinical trial setting, which is obviously dissimilar from the actual clinical practice; thus, specialization, rate of control and visit results are more intense and often not generalizable or comparable with a general practice setting.

Road map and documents

Tips: prepare in advance a detailed study protocol, a realistic timeline and proper data collection forms. A procedure manual must be necessary according to the complexity of the RCT and of the demanded tasks.

A detailed study protocol and manual of operations have central roles in RCT design and progress. The study protocol should contain in details the background, the objective, the rationale and the importance, as well as the design, the methodology, the Institutional Review Board approval, the informed consent and the statistical considerations of the RCT. In addition, the selection criteria for patient eligibility (i.e., inclusion and exclusion) and the concrete organization of the RCT (e.g., recruitment, baseline data collection, randomization, treatment administration, control visits and follow-up) should be described in the document (9).

Even if small or simple clinical trial require only a study protocol document, a procedures manual should be needed in more complex or large RCT. The procedures manual should contain study definitions, descriptions and instructions of each procedure and each task/item of data collection process. The manual should be written in details and could contain figure or diagrams to elucidate and forecast possible problems in procedures finalization (10).

Data collection forms should include all crucial items for evaluate baseline characteristics and outcomes. In order to easily collect, to avoid interferences and to limit related cost in data collection, these forms should be consistent and organized in logical order, according to the timing of procedures amongst RCTs. Data collection forms should be easy to be completed properly and unnecessary secondary variables should be avoided, as well as possible non-response and write-in responses. Baseline data collection should include items needed to confirm eligibility, to permit randomization and to collect predictors for possible stratification. Follow-up data collection encompasses information on primary and secondary outcomes, as well as treatment toxicity and morbidity.

Finally, another crucial and sometimes underestimated element to be prepared in advance is a realistic timeline document. Timeline document should report all the crucial steps of the starting RCT, with realistic and achievable time-objective.

A comprehensive study protocol with a detailed procedures manual, a realistic timeline and proper data collection forms define the Road Map needed to perform a correct RCT.

Hypothesis and outcome

Tips: formulate a single, simple and clear main hypothesis, accompanied by limited number of secondary ones. Select an intervention or a treatment that is clinically relevant and could be correctly assessed. Choose a significant endpoint that could be simply and practically verified.

RCTs typically assess a single intervention or a treatment in a limited and controlled setting. This because of strengths and restrictions associated to the nature of this kind of study. For these reasons, they have certain limitations to explore composite interventions in complex populations (e.g., elder patients, multiple pharmacological interactions, numerous comorbidity), which is the common scenario in actual clinical practice (11,12). Mostly, RCTs demonstrate to be an excellent setting for phase II (evaluation of treatment efficacy and safety) and phase III (evaluation of treatment efficacy and effectiveness in comparison with 'gold standard' or placebo) studies. Nevertheless, they have limitations in investigate rare outcomes or delayed effects.

Thus, results imperative to choose a clear hypothesis, verifiable by a limited number of strong and clinically significant endpoints. The general objective of RCTs is

to obtain results of easy and concrete applicability for clinicians and that can be easy implemented in ordinary clinical settings. Is about transfer scientific knowledge in medical decision-making strategies.

For example, Wolfe *et al.* recently reported how the surgical intervention of thymectomy improves clinical outcomes in non-thymomatous myasthenic patients in a period over 3 years when compared with only prednisone treatment. The robust underlying hypothesis, the clear endpoints, and the relevance of the intervention, such as thymectomy, highlight how this RCT has a strong validity and helps changing clinical practice (13).

On the other hand, Kleinberg *et al.* demonstrated the value of choosing an appropriate endpoint; using as parameter the pathological complete response at the primary site, which may not predict correctly the overall survival outcome, might jeopardize study's results. Indeed, the proposed alternative neoadjuvant treatments, with considerable toxic effects, didn't impact on long-term survival when compared to standard therapeutic protocols (14).

Selection criteria and sample size

Tips: find an equilibrium between very strict and selective criteria (standardized patient group) and more heterogeneous conditions (external validity of the results). Always taking account of possible under recruitment and loss to follow up.

A precise statistical preparation of the RCT must taking account of the selection criteria and the power needed to obtain valuable results. The patient selection criteria (inclusion and exclusion, both) must be chosen to avoid possible confounding factors, to exclude patients in whom the intervention is useless or dangerous. Moreover, the selection criteria should be not too severe; the risk is to conduct the RCT in an overly selected population and to obtain results not generalizable at the actual clinical practice.

A sufficient sample size is fundamental to detect a reliable statistical difference among the study groups. The sample size needed to reach an adequate power in a study is inversely proportional to the intervention effect squared (15). Consequently, considering that frequently the effect of the studied intervention is relatively small, the number of patients needed is relatively large. Nevertheless, an insufficient sample size is a frequent problematic issue in several published RCTs.

For example, Portier *et al.* compared adjuvant fluorouracil and folinic acid administration with surgery alone after colorectal liver metastases resection (16). The results indicated an effect of adjuvant chemotherapy with an improved progression-free survival but not a statistically significant effect on overall survival. However, a trend towards an improvement on overall survival was observed and, probably, significance was not reached because a lacking statistical power due to a small sample size. Indeed, diverse RCTs on adjuvant chemotherapy regimens after liver metastasectomy are usually underpowered to find significant conclusion (17).

Contrariwise, Pompili *et al.* performed a power analysis to determinate a correct sample size to detect a difference in duration of chest tube placement after segmentectomy or lobectomy of at least 1 day (18). The proper number of recruited patients permit them to demonstrate a significant improvement from digital chest drain utilization.

Another recurrent problem in most of RCTs is a low recruitment rate. This is due to recruitment difficulties, inadequate selection criteria or patient willing. Moreover, some patients leave the intervention group due to patient or physician choices or treatment complications. Finally, other patients will be lost at the follow-up, and define the outcome for this patient will be not possible. Therefore, is mandatory to forecast that the RCT will be completed only in a relatively small percentage of all potentially eligible population.

For example, the ENG trial on adjuvant systemic treatment after liver resection for colorectal metastasis, closed prematurely due to poor recruitment. The final analysis was carried out on 107 patients and no effect on overall survival was observed. However, a successive pooled analysis that merged this trial with FFCD ACHBTH AURC 9002 Trial, showed a significant effect of chemotherapy on overall survival and progression-free survival, both (19).

Randomization, stratification, blind and intention to treat analysis

Tips: choose and report the methods of Randomization correctly. Balance the study group using stratification technique. At least, outcomes evaluation should be blinded. Always apply intention to treat analysis.

A key aspect of RCTs is the method of randomization. Random allocation of patients in the study or in the control group assures that all participants' known and unknown

characteristics are similar and balanced between groups at the beginning of RCTs (6,11). Therefore, the study group will differ for treatments type assigned only, avoiding the selection bias (20). One of the most important aspects of randomization is the impossibility to determine *a priori* the allocation of each patient. Consequently, is mandatory to report all the aspects of the randomization process: the randomization method (e.g., a coin toss, a table of random numbers, computer based schedule), personnel involved (physician, nurse, technician), randomization timing, existence of a randomizations register.

For example, Pompili *et al.* performed a trial on advantage of digital chest drain and accurately describe randomization method (a randomization list, generated by a computer software, blinded in consecutively numbered sealed envelopes) and timing (end of each surgical procedure) (18).

Despite randomization, especially in small RCTs, the risk of unbalance in important prognostic factors amongst groups could remain relevant. The stratification method allows to balance groups over the predictors of interest and to increase analysis power.

Blind methodology is used to prevent the possible bias derivate from the knowledge of group allocation of a patient (e.g., subjective evaluation of results). A "double blinded" approach is when both physician and patients don't know which is the treatment received. Indeed, even if exist numerous methods to maintain a "double blinding" of a RCT (e.g., placebo treatment in drug therapies, sham procedure in surgical studies), sometimes is not possible to disguise treatment allocation. In these cases, at least the personnel dedicated to the evaluation of the response to the treatment should do not have information regarding group allocation. This assures to avoid subjectivity in outcome assessing and to permit the reliability and the objectivity of the results (21,22).

For example, in the recent trial, evaluating effect of surgical thymectomy for improve clinical outcomes in non-thymomatous myasthenia patients, double blinding is fairly difficult. Thus, to preserve rater blinding almost in outcome, patients were evaluated 4 months after surgical procedure by a neurologist who was not aware of the trial-group assignments.

The intention to treat analysis (ITT) analysis is a solid method to avoid analytical bias (23). The ITT considers each patient belonging to the group to which he/she was initially allocated, regardless if he/she finally will or will not be submitted to assigned treatment. The patients that did not performed the planned intervention will be not excluded from RCTs, and this prevent the possible bias of patient withdraw or crossover.

For example, Corris *et al.* used ITT analysis to evaluate the efficacy of azithromycin treatment *vs.* placebo in bronchiolitis obliterans post-transplantation. In this way, the authors avoid the bias effect due to patients' withdrawal that did not complete the 12-weeks of study drug (24).

Conclusions

RCTs are incredible and irreplaceable tools for clinical researchers; nevertheless, aren't free of internal pitfalls that might render them not easy to be developed or utilized. Our aim in this article is to suggest certain tips useful for the design of a RCT; have a realistic timeline, define a clear objective and precise endpoints, balance the study with a correct randomization are key elements that help us assuring a strong study's validity.

RCTs conducted in perfect and ideal conditions often are not easy to be applied in routine clinical contest. A major point is to focus on the right equilibrium between strict selection criteria and more heterogeneous parameters that can help in a "real life" contest.

Based on these indications, is clear that each point from the design to the realization of a RCT is important for the study's results. If we aim to obtain RCT strengthening evidence for clinical practice, we have to build them on strong hinges that allow us to influence the scientific literature and change the clinical decision-making activity of physicians involved in thoracic disease management.

Acknowledgements

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Riegelman R. editor. Studying a study and testing a test: how to read the medical evidence. Philadelphia, PA: Lippincott Williams & Wilkins, 2000.
2. National Cancer Institute at the National Institutes of Health. NCI Dictionary of Cancer Terms: Levels of

evidence. US DHHS-National Institutes of Health. Available online: https://www.cancer.gov/publications/dictionaries/cancer-terms?CdrID=446533
3. United States. Prevention Services Task Force. Guide to clinical preventive services. DIANE Publishing, 1989.
4. Oxford Centre for Evidence-based Medicine – Levels of Evidence (March 2009). Centre for Evidence-Based Medicine. Retrieved 25 March 2015.
5. Jaillon P. Controlled randomized clinical trials. Bull Acad Natl Med 2007;191:739-56.
6. Nichol A, Bailey M, Cooper DJ, et al. Challenging issues in randomised controlled trials. Injury 2010;41:S20e3.
7. Albert RK. "Lies, damned lies ..." and observational studies in comparative effectiveness research. Am J Respir Crit Care Med 2013;187:1173-7.
8. Ware JH, Hamel MB. Pragmatic trials--guides to better patient care? N Engl J Med 2011;364:1685-7.
9. Meinert CL. editor. Clinical trials: design, conduct, and analysis. New York, NY: Oxford University Press, 1986.
10. Knatterud GL, Forman SA, Canner PL. The Coronary Drug Project. Design of data forms. Control Clin Trials 1983;4:429-40.
11. Nallamothu BK, Hayward RA, Bates ER. Beyond the randomized clinical trial: the role of effectiveness studies in evaluating cardiovascular therapies. Circulation 2008;118:1294-303.
12. Stanley K. Design of randomized controlled trials. Circulation 2007;115:1164-9.
13. Wolfe GI, Kaminski HJ, Aban IB, et al. Randomized trial of thymectomy in myasthenia gravis. N Engl J Med 2016;375:511-22.
14. Kleinberg LR, Catalano PJ, Forastiere AA, et al. Eastern Cooperative Oncology Group and American College of Radiology Imaging Network Randomized Phase 2 Trial of Neoadjuvant Preoperative Paclitaxel/Cisplatin/Radiation Therapy (RT) or Irinotecan/Cisplatin/RT in Esophageal Adenocarcinoma: Long-Term Outcome and Implications for Trial Design. Int J Radiat Oncol Biol Phys 2016;94:738-46.
15. Stolberg HO, Norman G, Trop I. Randomized controlled trials. AJR Am J Roentgenol 2004;183:1539-44.
16. Portier G, Elias D, Bouche O, et al. Multicenter randomized trial of adjuvant fluorouracil and folinic acid compared with surgery alone after resection of colorectal liver metastases: FFCD ACHBTH AURC 9002 trial. J Clin Oncol 2006;24:4976-82.
17. O'Connell MJ, Adson MA, Schutt AJ, et al. Clinical trial of adjuvant chemotherapy after surgical resection of colorectal cancer metastatic to the liver. Mayo Clin Proc 1985;60:517-20.
18. Pompili C, Detterbeck F, Papagiannopoulos K. Multicenter international randomized comparison of objective and subjective outcomes between electronic and traditional chest drainage systems. Ann Thorac Surg 2014;98:490-6; discussion 496-7.
19. Mitry E, Fields AL, Bleiberg H. Adjuvant chemotherapy after potentially curative resection of metastases from colorectal cancer: a pooled analysis of two randomized trials. J Clin Oncol 2008;26:4906-11.
20. McKee M, Britton A, Black N, et al. Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies. BMJ 1999;319:312-5.
21. Estellat C, Torgerson DJ, Ravaud P. How to perform a critical analysis of a randomised controlled trial. Best Pract Res Clin Rheumatol 2009;23:291-303.
22. Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008;336:601-5.
23. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999;319:670-4.
24. Corris PA, Ryan VA, Small T, et al. A randomised controlled trial of azithromycin therapy in bronchiolitis obliterans syndrome (BOS) post lung transplantation. Thorax 2015;70:442-50.

Cite this article as: Guerrera F, Renaud S, Tabbò F, Filosso PL. How to design a randomized clinical trial: tips and tricks for conduct a successful study in thoracic disease domain. J Thorac Dis 2017;9(8):2692-2696. doi: 10.21037/jtd.2017.06.147

