The Necessity of Randomized Clinical Trials

Janus Christian Jakobsen[1,2*] and Christian Gluud[1]

[1]*The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark.*
[2]*Emergency Department, Holbæk Hospital, Denmark.*

Authors' contributions

This work was carried out in collaboration between both authors. Both authors conceived the study, contributed to the writing of the paper, and are guarantors. Both authors read and approved the final manuscript.

Policy Paper

ABSTRACT

Aims: The hierarchy of evidence-based medicine determines the inferential powers of different clinical research designs. We want to address the difficult question if observational evidence under some circumstances can validate intervention effects.

Methodology: Assessment of previous argumentation aiming at a clear conclusion for future decision-making.

Results: We present five arguments demonstrating the fundamental need of randomized clinical trials to sufficiently validate intervention effects. Furthermore, we argue that hindrances to the conduct of randomized clinical trials can be lessened through education, collaboration, infrastructure, and other measures. Our arguments validate why the randomized clinical trial should and must be the study design evaluating interventions. By choosing the randomized clinical trial as the primary study design, effective preventive, prognostic, diagnostic, and therapeutic interventions will reach more patients earlier.

Conclusion: Clinical experience or observational studies should never be used as the sole basis for assessment of intervention effects — randomized clinical trials are always needed. Therefore, always randomize the first patient as Thomas C Chalmers suggested in 1977. Observational studies should primarily be used for quality control after treatments are included in clinical practice.

*Corresponding author: Email: jcj@ctu.dk;

1. INTRODUCTION

Observational studies, such as non-randomized cohort studies or patient series, are usually viewed as producing results with less evidential weight compared to the results from randomized clinical trials [1,2]. However, quite often clinicians argue that their clinical experience sufficiently can assess the effects of some interventions [3], and some publications state that observational studies can adequately validate intervention effects [4-8]. Conducting observational studies require much less work and resources than conducting randomized clinical trials, and randomized clinical trials are often perceived as bureaucratic and difficult to conduct. Therefore, it is no surprise that many investigators choose observational studies to try to assess intervention effects.

We will in the following paragraphs consider if randomized clinical trials always are necessary and the best clinical study design to assess any kind of health-care intervention, including drugs, medical devices, surgery, nutrition, psychotherapy, in vitro diagnostic medical devices, etc. [9-13]. We are convinced that Thomas C. Chalmers was correct when he stated that we should always randomize the first patient [14]. However, we also acknowledge the difficulties that randomized clinical trials may cause and that they too may show erroneous results. We will, therefore, in the second part of the manuscript provide a list of the typical issues that represents a perceived or real hindrance for the conduct of randomized clinical trials and we will suggest some remedies to reduce these hindrances.

Randomized clinical trials cannot only assess the effects of many different forms of experimental interventions, but also many different forms of control interventions, e.g., no intervention, placebo, 'impure' placebo, nocebo, or an 'active' control intervention (the latter being a treatment backed by convincing evidence from randomized clinical trials with low risks of systematic errors due to bias; with low risks of systematic errors due to design flaws; and with low risks of random errors due to play of chance). The latter trials compare the effects of two interventions (so-called 'head-to-head' trials or 'comparative intervention research'). It is clear that the inferences of the results from the different forms of trials differ according to their design. We will in the following paragraphs use the term 'randomized clinical trials' as a collective term for all kinds of trials, as we believe that the fundamental principles are similar regardless of type of experimental intervention and control intervention. The fundamental construct of the randomized clinical trial allows that any intervention using quantitative or qualitative outcomes can be assessed using the same basic principles [15,16].

2. METHODS AND RESULTS

2.1 Five Arguments Demonstrating the Fundamental Need of Randomized Clinical Trials to Assess and Validate Intervention Effects

2.1.1 Development of interventions is a prospective process

It is important to make the correct choice of study design before the initial assessment of a new intervention. The optimal indication, effect size, and balance between harmful and beneficial effects (see the paragraphs below) will remain unknown if randomized clinical

trials are not conducted before an intervention is implemented into clinical practice. We fully agree with Thomas C. Charmers when he in 1977 wrote that we should always randomize the first patient [14]. Accordingly, when an investigator wants to assess if an intervention is effective or not, an observational design should never be used for the initial assessment of the intervention. We will in the paragraphs below consider if there are exemptions to this rule.

Large well-conducted observational studies can sometimes provide useful information about rare adverse events and intervention effects [17]. We acknowledge a few historical instances where observational evidence validly have demonstrated benefits of new interventions (e.g., insulin for diabetic coma and ether for anaesthesia) [5]. However, we cannot a priory identify such rare instances. It is only in retrospect it may be concluded that interventions have been validly assessed by observational studies [5], and evidence based on observational evidence will in most circumstances be uncertain [18-20]. Observational studies will often either grossly overestimate or underestimate intervention effects and adjustment with statistical analyses (logistic regression or propensity score) only seem to increase the problem [20]. If an intervention is implemented into clinical practice based on observational evidence and seems to work, it can be difficult to justify and to conduct randomized clinical trials assessing the correct balance between benefits and harms. In this situation, we may never know the 'true' balance between benefits and harms. If an intervention does not look rewarding in an observational study we will likely stop further assessment of the intervention and therefore risk 'throwing the baby out with the bath water'. Intervention research during the development of drugs, devices, and other interventions are in essence a prospective process and the correct research design has to be selected prospectively [21]. The correct design ought to be the randomized clinical trial [14,16].

2.1.2 Implementation of scientific results into clinical practice

If an intervention offers more benefit than harm compared with previous treatment options, it is an ethical obligation and hence necessary to get that intervention offered to as many patients as possible, as fast as possible. In the discussion about choice of design for assessing new interventions, investigators often claim that it is important to conduct a quick observational study so the potential treatment can speedily reach the global market if 'proved' effective [22]. Many medical devices have, for example, been implemented into clinical practice on the basis of observational evidence alone [23]. However, if only observational evidence backs the intervention it may be difficult to reach clinical consensus about a given intervention effect because clinicians might rightly question the validity of such results [18-20]. It is much more easy to reach clinical consensus based on results from randomized clinical trials preferably assessed in systematic reviews ad modum those conducted according to The Cochrane Collaboration Handbook [1]. Even if an intervention has a parachute-like beneficial intervention effect [24], a fast way to the global market might be blocked if the intervention is only assessed in observational studies. The results of properly conducted randomized clinical trials will be more readily accepted by more clinicians than results from observational studies and the randomized clinical trial will therefore probably offer a faster access to a larger market compared to market penetration via an observational design.

2.1.3 Balance between beneficial and harmful effects

It is theoretically possible to quantify a beneficial intervention effect size via observational evidence if the disease is stable and without any fluctuation in symptoms and if the

intervention effects are large enough to be recognised by 'observation'. However, very few diseases show such stability and interventions with large easily observable effects are extremely rare [15]. Most interventions have no beneficial effects or relatively small beneficial effects. It is among the latter we shall find the interventions of tomorrow. Moreover, large 'surprising' beneficial effects shown in observational studies may be due to random errors, systematic errors, or confounding. Randomized clinical trials are, therefore, needed to assess when potential beneficial effects outweigh the potential harmful effects. Randomization is able to construct the optimal control group, which, at baseline, becomes fully comparable with the experimental group regarding all known and all unknown prognostic factors — provided that the randomized groups become large enough. Without randomization and without an appropriate control group it is often unclear if a change in symptoms is caused solely by an intervention effect — or if some, or all, of the change is a natural fluctuation of the symptoms (often a combination of 'regression towards the mean' and the natural fluctuation of the symptoms). Observational studies including some kind of matched control group do not provide valid information about effect sizes, because the participants in the control group will almost never be fully comparable to the participants in the experimental group [20]. It is therefore impossible to quantify and have an overview of the relative effect sizes via observational evidence only (Box 1).

BOX 1

It can be 'observed' that an operation for heartburn can normalize pH in the oesophagus [25], but the surgical procedure also carry some risks [26,27]. Observational evidence cannot assess when the degree of heartburn justifies an operation with possible harmful effects [27]. Furthermore, without randomization it is unclear whether a change in symptoms is caused by the operation or by other factors.

Long-acting beta$_2$-agonists can improve lung function in asthma patients [28], but after a large number of participants have been assessed evidence has indicated that long-acting beta$_2$-agonists also cause a small increase in mortality [28]. Such rare harmful effects would be impossible to detect without randomized clinical trials. It would be unclear whether the relatively few deaths were caused by the long-acting beta$_2$-agonists or by other factors.

Without an assessment of the balance between benefits and harms it is impossible to assess the clinical significance of a preventive, prognostic, diagnostic, or therapeutic intervention. It is important to use the appropriate control group of a randomized clinical trial in order to make valid inferences. If a trial comparing the effects of two active interventions shows no difference in effect it is not on the face of it clear whether the two interventions are equally effective or equally ineffective. The interpretability of results from randomized trials using placebo as control intervention will on the face of it in a similar way be unclear because the effects of a placebo may be unknown. E.g., if trial results show no difference in effect between a placebo intervention and an experimental intervention and the placebo intervention does have significant effects, then the placebo effects can mask effects from the experimental trial intervention. It is always of great importance to consider if a placebo intervention (traditional placebo, nocebo, or 'active' placebo) might have a clinical effect. The optimal 'placebo' is a substance which on the face of it is identical to the experimental intervention but without any 'active' effects. Nevertheless, robust evidence has shown that most placebo interventions have very small effects or no effects at all compared with no intervention [29]. Therefore, placebo-controlled clinical trials will most likely demonstrate the effects of the experimental intervention. Randomized clinical trials assessing the effects of

experimental interventions versus placebo are therefore in general the optimal method to accurately assess the effects of an intervention (Table 1). If effective treatments exist, then such treatments may either be used as the control intervention or as basis treatment for all participants in all of the trial intervention groups, i.e., an experimental intervention may then be assessed as an add-on intervention versus placebo or another intervention while all groups receive the already known effective treatment. Here The Declaration of Helsinki and medical regulatory agencies have been too kind to the product and ignored the patient [30-32] – and even the 2013 suggested amendments to The Declaration seem to have missed this point [33].

Table 1. Some hindrances of randomized clinical trials and possible solutions

Typical issues perceived or realized as hindrances for the conduct of randomized clinical trials	Potential solutions and counter arguments
Practical issue: It is time consuming to conduct randomized clinical trials.	**Potential solutions:** Investigators must be taught the most effective way of conducting randomized clinical trials — how to use the resources in the most efficient way. Counselling from competent trialists or trial units is essential.
Practical issue: Difficulties recruiting enough trial participants.	**Potential solutions:** Realistic sample size estimation must be calculated based upon the primary outcome early on in trial planning. More participants will be recruited in multicentre trials compared to single centre trials and through the use of broad inclusion criteria and appropriately selected exclusion criteria [34,35].
Methodological issue: Lack of methodological know-how and lack of practical experience conducting randomized clinical trials.	**Potential solutions:** Establishment of academic industry independent trial units and infrastructures of such units with know-how about evidence-based medicine [36] and trial design can lessen and solve some of the many problems conducting randomized clinical trials.
Ethical issue: It can be difficult to ethically justify the conduct of a randomized clinical trial especially if the control group is receiving no intervention or placebo.	**Potential solutions:** It may be unethical to treat patients with interventions that are not based on evidence. Furthermore, if an evidence-based treatment exists, then all intervention groups should ideally receive this treatment (see text). A new experimental intervention can then be assessed as an add-on intervention in the experimental intervention group versus placebo or another add-on intervention in the control group. All participants will receive the treatment that previous evidence has shown offers more benefits than harms and the trial is ethically justified.
Typical misconception: Trial participants differ from patients in common clinical settings [4,37,38]. Strict inclusion and exclusion criteria are believed to put together trial populations not representative of patients in the clinic questioning the clinical relevance of	**Counter argument:** It is not necessary to use narrow criteria for selecting trial participants [1,35,39]. Using fewer inclusion and exclusion criteria will also make trial populations more similar to patients in the clinic. Moreover, patients that receive similar interventions within and outside

results from randomized clinical trials [4,37,38].

randomized clinical trials seem to have similar prognosis [39,40].

Typical misconception:
Intervention effects in a trial setting are not representative of intervention effects in the clinic. Trial participants are often subjected to strict thorough treatment protocols and repetitive follow-up assessments of different kinds. It has been postulated that this might specifically benefit trial participants (and hence the trial results) compared to patients in the clinic [4,41,42].

Counter argument: Allocation to an experimental intervention in a trial setting compared to a similar treatment outside a trial setting has been shown to have similar effects [39,40,43]. Moreover, it is not necessary to use strict treatment protocols in a randomized clinical trial [1]. It is possible to randomize participants to, e.g., a non-standardized care versus 'no intervention'.

Typical misconception:
Interventions cannot be standardized without compromising efficacy. It is believed that randomized trials cannot assess the effects of individualized patient treatment, where clinicians effectively treat each patient according to clinical expertise and experience [22,44].

Counter argument: Standardized interventions based on evidence-based practice are most often superior to non-standardized interventions [45-48]. Furthermore, it is possible in a randomized clinical trial to compare the effects of treating patients according to clinical experience with a standardized intervention or another comparator. Any intervention can be assessed in a randomized clinical trial using a given outcome.

Typical misconception:
It is costly to conduct randomized clinical trials.

Counter argument: If you think clinical research is costly, consider clinical practice. It has been calculated that investment in randomized clinical trials usually gives a reasonable or high return on investment [49]. Politicians and other decision makers must be taught the key position of the randomized clinical trial regarding knowledge about intervention effects. The more effective the healthcare system becomes, the cheaper it will be.

We have in Table 2 presented an overview of the different types of randomized clinical trials and summarized the corresponding methodological strengths and limitations.

Studies have shown that observational studies compared to randomized clinical trials often overestimate benefits and underestimate harms, i.e., produce biased results [18-20]. To accurately and objectively assess the balance between benefits and harms, we need randomized clinical trials with blinded outcome assessment. Blinded randomized clinical trials compared to unblinded randomized clinical trials show significantly less biased results [50,51]. A valid and unbiased assessment of benefits and harms are impossible to achieve in an observational design where blinding usually is impossible.

Table 2. Different comparisons in randomized clinical trials and associated methodological strengths and limitations.

Different types of control groups					
Experimental intervention versus no intervention		**Experimental intervention versus placebo, impure placebo*, 'active' placebo (nocebo)**, or a sham intervention**		**Experimental intervention versus 'treatment as usual'*****	
Strengths	**Limitations**	**Strengths**	**Limitations**	**Strengths**	**Limitations**
The beneficial and harmful effects of the experimental intervention can be assessed by the results.	Results of the trial may be biased due to lack of blinding of the participants. It may be ethically wrong to conduct the trial if an effective treatment exists.	Allows blinding of trial participants; investigators; treatment providers; outcome assessors; data managers; statisticians; and conclusion drawers. Allows assessment of experimental intervention effect sizes controlling for non-specific treatment factors****.	The 'effect' of placebo may be unclear in certain conditions. Participants can often because of beneficial effects or adverse effects figure out if they are treated with the active intervention or the control intervention.	The trial results demonstrate what a given average patient gains by an experimental intervention compared with the treatment the patient usually receive.	Treatment as usual most often contains some non-specific treatment elements with unknown effects. Results may be biased as no blinding is involved, unless one uses double placebo ('double dummy').

Co-interventions

All three types of trials can include different kinds of co-interventions delivered similarly to all intervention groups. If there is no interaction between these co-interventions and the experimental and control interventions, the effects of the co-interventions will even out between the two comparison groups

* A substances with pharmacological effects but not considered to have an effect on the condition being treated (e.g., antibiotics in viral infections or vitamins for prevention of death).
** A placebo preparation that mimics the adverse effects (nocebo) of the experimental intervention.
*** An intervention where participants are treated, as they would have been if they had not been included in the trial. Terms like treatment as usual, standard care, or usual care (synonyms) are often collective terms used for different non-specific interventions.
**** Trial participants might benefit from, e.g., believing that an intervention is effective or just from being in contact with a treatment provider. Placebo-controlled blinded trials can assess the specific effects of an intervention because the outcome of the control group will ideally show the effects of the non-specific treatment factors.

2.1.4 Patient-relevant and clinically relevant outcomes

Intervention effects on patient-relevant and clinically relevant outcomes such as psychological distress, quality of life, patient satisfaction, and pain are impossible to assess accurately by 'observation' (Box 2). Such outcomes should be reported and assessed by the

patient and not by a clinician and are by nature subjective, fluctuating, and a placebo effect can be significant [29]. Therefore, randomized clinical trials enabling blinding of all parties (participants; investigators; health-care providers; outcome assessors; data managers; statisticians; conclusion drawers) are mandatory to validly assess patient relevant and clinically relevant outcomes [1].

BOX 2

A clinician can observe that laser intervention can reduce redness of a 'port-wine stain' on the skin of a patient [52]; or that chemotherapy seems to prolong survival in incurable cancer patients [53]. However, the most clinically relevant outcomes in these two examples would likely be long-term patient satisfaction after the cosmetic laser treatment in patients with port-wine stains [52] and 'quality of life' and QUALY (quality adjusted life years) of the cancer patients [54]. These outcomes are impossible or difficult to assess only by clinical 'observation'.

2.1.5 Indications for an intervention

Most diseases have varying degrees of severity. When a disease is on the borderline between severe and 'not severe', only randomized clinical trials can determine if we should intervene or not. Randomized clinical trials are necessary to determine the most optimal indication for an intervention — when to treat or when not to treat. We have illustrated this in the two examples in Box 3. Randomized clinical trials, with low risk of bias, low risk of design errors, and low risk of random errors can via prospectively planned subgroup analyses suggest such indications [1,55]. However, because of concerns of multiplicity and of the small sample sizes often involved, subgroup analyses should be viewed only as hypothesis generating exercises [56,57]. If subgroup analyses show effect in only one or more of the subgroups, then new confirmatory randomized clinical trials on these subgroups ought to be conducted [58].

BOX 3

Tracheostomy can be lifesaving for patients with risk of obstructed airways, but tracheostomy can also cause serious complications such as fatal bleeding and airway stenosis [59]. Without randomized clinical trials it is not apparent how severe the hypoxia should be before performing tracheostomy [59].

It can be observed that defibrillation can convert ventricular fibrillation to normal sinus rhythm in patients with cardiac arrest. However, randomized clinical trials are needed to determine when defibrillation for long-term cardiac arrest will lead to a meaningful life of the patient — and when it will not [60].

2.2 Typical Hindrances for the Conduct of Randomized Clinical Trials and some Remedies to Reduce These

Conducting randomized clinical trials generally require more resources than conducting observational studies. Researchers can be reluctant to conduct randomized clinical trials because they are costly and time consuming. Lack of methodological and statistical know-how can hinder the making of randomized clinical trials; it can be difficult to recruit enough trial participants, etc. Typical misconceptions about the usefulness of results from

randomized clinical trials can also hinder that such trials are conducted. It is, e.g., often stated that trial populations are not representative of patients in the clinic [4,37,38]. Strict inclusion and exclusion criteria (e.g., the need of informed consent) are believed to put together trial populations not representative of patients in the clinic. The ethically need of informed consent can theoretically affect trial populations so they are different from the everyday patients, but such fears are often overestimated [39,40]. Besides the need of informed consent it is generally not necessary to use narrow criteria for selecting trial participants, as this may impair the external validity of a trial [35]. We acknowledge all of these difficulties regarding randomized clinical trials. Nevertheless, the establishment of academic industry independent trial units with know-how about evidence-based medicine [36] can lessen and solve some of the many problems conducting randomized clinical trials [61-66]. Furthermore, regional, national, international, and global research collaboration between trial units and clinical sites (e.g., The European Clinical Research Infrastructures (ECRIN), The UK Clinical Research Collaboration (UKCRC) Clinical Trials Units Network [67], and The Nordic Trial Alliance (NTA) [68]) may reduce problems with recruitment of a sufficient number of trial participants, etc. [69,70]. Well-conducted multicentre clinical trials also offer better external validity than well-conducted single centre trials. It must be recognized how much health-care costs can be reduced if patient treatment becomes more effective through evidence-based research. It has been calculated that investment in randomized clinical trials usually gives a reasonable or high return on investment [49]. Politicians and decision makers must be taught the key positions of the randomized clinical trial and of systematic reviews of such trials in clinical intervention research.

We have in Table 1 listed typical issues and misconceptions that are perceived or realized as obstacles for the conduct of randomized clinical trials and pointed out how the problems may be minimized.

3. DISCUSSION

We have pointed out the dangers with observational evidence and concurred with others, that the randomized clinical trial is the optimal design to use when new interventions are to be assessed and when questions arise about the advantages of treatments already in use in clinical practice. Our recommendations should not be surprising, as they represent the opinion of drug regulatory agencies all over the globe. We just stress that these recommendations should be expanded to all interventions. We acknowledge that conducting randomized clinical trials is more difficult than conducting observational studies. However, typical issues hindering the conduct of trials can be overcome (Table 1).

We believe that clinical experience and observational studies cannot and should not validate the effects of interventions. Observational studies can sufficiently assess associations between certain interventions and outcomes, but the randomized clinical trials are always needed to avoid falsely negating (type I error) or falsely confirming (type II error) the null hypothesis. Randomized clinical trials are needed to sufficiently validate intervention effects and to assess causality between interventions and outcomes. Observational evidence should be used primarily to detect very rare adverse events, very late adverse events, or to monitor the quality of medical treatments once they have been introduced in clinical practice [71].

A report from the Patient-Centered Outcomes Research Institute was recently published for public comment [72]. This report claims that the use of observational studies to make causal inference is potentially much stronger than it has been in the past [72], and similar

arguments are often published in highly esteemed journals [3-7,73]. We believe that the fundamental construct of the observational studies limits the reliability of the results from observational studies [18,20]. To assess if an intervention causes more benefit than harm randomized clinical trials are, in practical terms, always needed. Deeks and colleagues have in a comprehensive report compared results from randomized trials and observational studies [20]. They showed that results from observational studies can be seriously misleading and that adjusted results in observational studies may even appear more misleading than unadjusted results [20]. Compared to small randomized clinical trials, small observational studies often showed effects that were far from the 'true' intervention effect [20]. Ioannidis and colleagues also showed that significant discrepancies do occur between the results of randomized clinical trials and observational studies [18] — and that results from observational studies are more often contradicted than results from randomized clinical trials [74]. Observational studies can be the only possible option regarding assessment of very rare adverse events, very late occurring effects, or of very long-term interventions. Observational studies can also have their place when it is difficult to include large enough sample sizes assessing extremely rare diseases or when lack of funds hinders the conduct of randomized clinical trials. Observational studies also have an important role in monitoring the quality of evidence-based medicine through use of patient registers and databases [71]. Observational studies have their place under such circumstances but their inferential power should always be considered threatened by random errors, confounding by indication, unmeasured confounding, and other systematic errors. Therefore, the randomized clinical trial would still in such circumstances be the optimal design regardless of hindrances making them infeasible. It may, as mentioned, be possible to present a few historical examples where intervention effects have been sufficiently validated by observational evidence [5]. However, these exceptions do not justify that observational evidence generally should be used prospectively to validate intervention effects. As it has been clearly expressed by Heiberg already in 1897 and reiterated by others both before and since [75-77] — regarding the vast majority of interventions randomized clinical trials are necessary to assess their effects.

We acknowledge that randomized clinical trials may also get intervention effects wrong. However, the likelihood of this occurring decreases with improved quality of the trial methodology (reducing the risks of systematic errors), with increasing sample sizes of the trials (reducing the risks of random errors), and with limiting the number of outcomes (reducing the risks of random errors) [1,50,51,55,78,79]. Moreover, the conduct of systematic reviews assessing all randomized clinical trials on an intervention as conducted by The Cochrane Collaboration also reduces these risks [1,55,78,79]. We therefore need to invest more in education in clinical research as well as in infrastructures for clinical research and for systematic reviewing of randomized clinical trials.

Another group of arguments also exposes the weaknesses of observational studies. For observational studies we do not yet have the requirements of making public peer-reviewed protocols before the epidemiologic work is started; we do not yet publish all data on individual participants in observational studies on a repository; and we do not yet have practices of systematically reviewing all observational studies on a topic. Regarding randomised clinical trials all of these issues have been solved or are in the making to be solved [1,80].

It may be frustrating for clinicians to realize that clinical experience and observational studies do not provide valid knowledge about intervention effects — especially because many interventions in clinical use have not been assessed in randomized clinical trials [72].

Randomized clinical trials or systematic reviews with low methodological quality (high risks of systematic errors due to bias and design errors) and insufficient sample sizes (high risks of random error) [81-85] should not be used to guide decision makers and clinicians about which intervention to choose. We aim to support the development and use of truly effective health-care interventions to the benefit of patients as well as health-care systems. This can be obtained by much wider use of randomized clinical trials for the proper assessment of benefits and harms. In times of austerity, the need of randomized clinical trials seems increasingly urgent. We must as clinicians realize the uncertainty of our knowledge if randomized clinical trials have not been conducted and remember the validity of the evidence hierarchy [86]. Systematic reviews of randomized clinical trials is and should be considered the highest level of evidence followed by single randomized trials [86]. We should not, necessarily, stop using all interventions not based on results from randomized clinical trials. However, we believe that patients most often should be treated with interventions that have been proved effective in randomized clinical trials. Regarding many conditions it might be best not to intervene unless randomized clinical trials with low risks of systematic errors ('bias'), low risks of design errors ('bias'), and low risks of random error ('play of chance') have shown more benefit than harm [1,55].

4. CONCLUSION

Clinical experience or observational studies cannot sufficiently assess and validate intervention effects — randomized clinical trials are always needed. We therefore disagree with authors claiming that observational designs can be employed for assessing interventions. Observational evidence should be restricted to assess rare adverse events; late adverse events; and to monitor the quality of evidence-based medicine through use of patient registers and databases.

CONSENT

Not applicable.

ETHICAL APPROVAL

Not applicable.

ACKNOWLEDGEMENTS

FUNDING

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Higgins JPT, Green S (2011) The Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. The Cochrane Collaboration; 2011. Available: http://www.cochrane-handbook.org.
2. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol. 2013;66(2):151-157.
3. Ulvenes LV, Aasland O, Nylenna M, Kristiansen IS. Norwegian physicians' knowledge of and opinions about evidence-based medicine: cross-sectional study. PLoS ONE. 2009;4(11):e7828.
4. Vincent JL We should abandon randomized controlled trials in the intensive care unit. Crit Care Med. 2010;38(10 Suppl):S534-538.
5. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials unnecessary? Picking signal from noise. BMJ. 2007;334(7589):349-351.
6. Stolberg HO, Norman G, Trop I. Randomized controlled trials. Am J Roentgenol. 2004;183(6):1539-1544.
7. Hawkins NG, Sanson-Fisher RW, Shakeshaft A, D'Este C, Green LW. The multiple baseline design for evaluating population-based research. Am J Prev Med. 2007;33(2):162-168.
8. Manchikanti L, Pampati V. Research designs in interventional pain management: is randomization superior, desirable or essential? Pain Physician. 2002;5(3):275-284.
9. Gluud C, Kubiak C, Whitfield K, Byrne J, Huemer KH, Thirstrup S, et al. Typical investigational medicinal products follow relatively uniform regulations in 10 European Clinical Research Infrastructures Network (ECRIN) countries. Trials. 2012;13:27.
10. Kramer DB, Xu S, Kesselheim AS. How does medical device regulation perform in the United States and the European union? A systematic review. PLoS Med. 2012;9(7):e1001276.
11. Feltham C. Which Psychotherapy?: Leading Exponents Explain Their Differences (page 1-224): SAGE Publications Ltd; 1997.
12. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, et al. No surgical innovation without evaluation: the IDEAL recommendations. Lancet. 2009;374(9695):1105-1112.
13. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. Lancet. 2011;377(9778):1693-1702.
14. Chalmers TC. Randomize the first patient. N Eng J Med. 1977;296(2):107.
15. Chalmers I, Milne I, Trohler U, Vandenbroucke J, Morabia A, Tait G, et al. The James Lind Library: explaining and illustrating the evolution of fair tests of medical treatments. J R Coll Physicians Edinb. 2008;38(3):259-264.
16. Evans I, Thornton H, Chalmers I, Glasziou P. Testing Treatments: Better Research for Better Healthcare. 2. Edition; 2011 Available: http://wwwtestingtreatmentsorg/wp-content/uploads/2012/09/TT_2ndEd_English_17oct2011pdf: 1-226.
17. Vandenbroucke JP. What is the best evidence for determining harms of medical treatment? CMAJ. 2006;174(5):645-646.

18. Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. JAMA. 2001;286(7):821-830.

19. Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. CMAJ. 2006;174(5):635-641.

20. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, et al. Evaluating non-randomised intervention studies. Health Technol Assess. 2003;7(27): iii-x, 1-173.

21. Chalmers I. Well informed uncertainties about the effects of treatments. BMJ. 2004;328(7438):475-476.

22. Cooper J. Randomized clinical trials for new surgical operations: square peg in a round hole? J Thorac Cardiovasc Surg. 2010;140(4):743-746.

23. Huot L, Decullier E, Maes-Beny K, Chapuis FR. Medical device assessment: scientific evidence examined by the French national agency for health - a descriptive study. BMC Public Health. 2012;12:585.

24. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ. 2003;327(7429):1459-61.

25. Cuschieri A, Hunter J, Wolfe B, Swanstrom LL, Hutson W. Multicenter prospective evaluation of laparoscopic antireflux surgery. Preliminary report. Surg Endosc. 1993;7(6):505-510.

26. Kahrilas PJ, Lin S, Spiess AE, Brasseur JG, Joehl RJ, Manka M. Impact of fundoplication on bolus transit across esophagogastric junction. Am J Physiol. 1998;275(6 Pt 1):G1386-1393.

27. Wileman SM, McCann S, Grant AM, Krukowski ZH, Bruce J. Medical versus surgical management for gastro-oesophageal reflux disease (GORD) in adults. Cochrane Database Syst Rev. 2010;(3):CD003243.

28. Cates C, Cates M. Regular treatment with salmeterol for chronic asthma: serious adverse events. Cochrane Database Syst Rev. 2008;16(3):CD006363.

29. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. Cochrane Database Syst Rev. 2010;(1):CD003974.

30. World Medical Association. Declaration of Helsinki. Ethical principles for medical research involving human subjects. J Indian Med Assoc. 2009;107(6):403-405.

31. Bertele V, Banzi R, Gluud C, Garattini S. EMA's reflection on placebo does not reflect patients' interests. Eur J Clin Pharmacol. 2012;68(5):877-879.

32. Garattini S, Bertele V, Banzi R. Placebo? no thanks, it might be bad for me! Eur J Clin Pharmacol. 2013;69(3):711-714.

33. World Medical Association. Declaration of Helsinki Working Group. Draft revised text for public consultation, 15 April – 15 June; 2013. Available: http://wwwwmanet/en/20activities/10ethics/10helsinki/15publicconsult/indexhtml.

34. Nichol AD, Bailey M, Cooper DJ, POLAR, EPO Investigators. Challenging issues in randomised controlled trials. Injury. 2010;41(Suppl 1):S20-23.

35. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. JAMA. 2007;297(11):1233-1240.

36. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-72.

37. Pincus T. Limitations of randomized clinical trials to recognize possible advantages of combination therapies in rheumatic diseases. Semin Arthritis Rheum. 1993;23(2 Suppl1):2-10.

38. Cartwright N, Munro E. The limitations of randomized controlled trials in predicting effectiveness. J Eval Clin Pract. 2010;16(2):260-266.

39. Vist G, Bryant D, Somerville L, Birminghem T, Oxman A. Outcomes of patients who participate in randomized controlled trials compared to similar patients receiving similar interventions who do not participate. Cochrane Database Syst Rev. 2008;(3):MR000009.

40. West J, Wright J, Tuffnell D, Jankowicz D, West R. Do clinical trials improve quality of care? A comparison of clinical processes and outcomes in patients in a clinical trial and similar patients outside a trial where both groups are managed according to a strict protocol. Qual Saf Health Care. 2005;14(3):175-178.

41. Joffe S, Weeks JC. Views of American oncologists about the purposes of clinical trials. J Natl Cancer Inst. 2002;94(24):1847-1853.

42. Davis S, Wright PW, Schulman SF, Hill LD, Pinkham RD, Johnson LP, et al. Participants in prospective, randomized clinical trials for resected non-small cell lung cancer have improved survival compared with nonparticipants in such trials. Cancer. 1985;56(7):1710-1718.

43. Gross CP, Krumholz HM, Van Wye G, Emanuel EJ, Wendler D. Does Random Treatment Assignment Cause Harm to Research Participants? PLoS Med. 2006;3(6):e188.

44. Appelbaum P, Roth L, Lidz C, Benson P, Winslade W (1987) False hopes and best data: consent to research and the therapeutic misconception. The Hastings Center Report. 1987;17(2):20-4.

45. Middleton S, McElduff P, Ward J, Grimshaw JM, Dale S, D'Este C, et al. Implementation of evidence-based treatment protocols to manage fever, hyperglycaemia, and swallowing dysfunction in acute stroke (QASC): a cluster randomised controlled trial. Lancet. 2011;378(9804):1699-1706.

46. Becker A, Held H, Redaelli M, Chenot JF, Leonhardt C, Keller S, et al. Implementation of a Guideline for Low Back Pain Management in Primary Care - A Cost-Effectiveness Analysis. Spine (Phila Pa 1976). 2011;37(8):701-710.

47. Shorr AF, Micek ST, Jackson WL, Jr., Kollef MH. Economic implications of an evidence-based sepsis protocol: can we improve outcomes and lower costs? Crit Care Med. 2007;35(5):1257-1262.

48. Karjalainen S, Palva I. Do treatment protocols improve end results? A study of survival of patients with multiple myeloma in Finland. BMJ. 1989;299(6707):1069-1072.

49. Health Economics Research Group, Office of Health Economics, RAND Europe Medical Research: What's it worth? Estimating the economic benefits from medical research in the UK; 2008. Avaiable: http://wwwwellcomeacuk/stellent/groups/corporatesite/@sitestudioobjects/documents/web_document/wtx052110pdf.

50. Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med. 2012;157(6):429-438.

51. Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ. 2008;336(7644):601-605.

52. Faurschou A, Olesen AB, Leonardi-Bee J, Haedersdal M. Lasers or light sources for treating port-wine stains. Cochrane Database Syst Rev. 2011;11:CD007152.

53. Boulaamane L, Boutayeb S, Errihani H. Bevacizumab based chemotherapy in first line treatment of HER2 negative metastatic breast cancer: results of a Moroccan observational institutional study. BMC Res Notes. 2012;5(1):162.

54. Montazeri A. Health-related quality of life in breast cancer patients: a bibliographic review of the literature from 1974 to 2007. J Exp Clin Cancer Res. 2008;27:32.
55. Keus F, Wetterslev J, Gluud C, van Laarhoven CJ. Evidence at a glance: error matrix approach for overviewing available evidence. BMC Med Res Methodol. 2010;10:90.
56. Magnusson BP, Turnbull BW. Group sequential enrichment design incorporating subgroup selection. Stat Med. 2013;doi: 10.1002/sim.5738.
57. Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. Health Technol Assess. 2001;5(33):1-56.
58. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. JAMA. 1994;272(17):1367-1371.
59. Jackson LS, Davis JW, Kaups KL, Sue LP, Wolfe MM, Bilello JF, et al. Percutaneous tracheostomy: to bronch or not to bronch - that is the question. J Trauma. 2011;71(6):1553-1556.
60. Stiell IG, Nichol G, Leroux BG, Rea TD, Ornato JP, Powell J, et al. Early versus later rhythm analysis in patients with out-of-hospital cardiac arrest. N Eng J Med. 2011;365(9):787-797.
61. Fleischmann R. Primer: establishing a clinical trial unit--obtaining studies and patients. Nat Clin Pract Rheumatol. 2007;3(8):459-463.
62. Fleischmann R. Primer: establishing a clinical trial unit - regulations and infrastructure. Nat Clin Pract Rheumatol. 2007;3(4):234-239.
63. Seiler CM, Wente MN, Diener MK, Frohlich BE, Buchler MW, Knaebel HP. Center for clinical studies in a surgical department - an approach for more evidence-based medicine. Contemp Clin Trials. 2006;27(3):211-214.
64. Ohmann C, Bruns I, Wolff S. The Coordinating Centers for Clinical Trials (KKS) and the KKS Network: competence for clinical research. Onkologie. 2010;33(Suppl7):11-15.
65. Gluud C, Sorensen TI. New developments in the conduct and management of multi-center trials: an international review of clinical trial units. Fundam Clin Pharmacol. 1995;9(3):284-289.
66. Gluud C, Sorensen TI. Organisation of a clinical trial unit--a proposal. Fundam Clin Pharmacol. 1998;12(3):298-305.
67. The UK Clinical Research Collaboration (UKCRC) (http://www.ukcrc-ctu.org.uk/).
68. Nordic Trail Alliance (NTA). Available: http://www.nordforsk.org/no/programs/nordic-trial-alliance-nta).
69. Demotes-Mainard J, Ohmann C. European Clinical Research Infrastructures Network: promoting harmonisation and quality in European clinical research. Lancet. 2005;365(9454):107-108.
70. Demotes-Mainard J, Ohmann C, Gluud C, Chëne G, Fabris N, Garattini S. European clinical research infrastructures network meeting report: towards an integration of clinical research infrastructures in Europe. Int J Pharm Med. 2005;19(1):43-45.
71. Winkel P, Zhang NF. Statistical Development of Quality in Medicine (page 1-224). Wiley; 2007.
72. Helfand M, Berg A, Flum D, Gabriel S, Normand S (2012) Draft Methodology Report: "Our Questions, Our Decisions: Standards for Patient-Centered Outcomes Research" (Page 1-186). Patient-Centeret Outcomes Research Institute Published for Public Comment July 23; 2012. Available: http://www.pcori.org/assets/Preliminary-Draft-Methodology-Report.pdf).

73. Higgins J. Convincing evidence from controlled and uncontrolled studies on the lipid-lowering effect of a statin [editorial]. Cochrane Database of Syst Rev. 2012;12(ED000049).
74. Ioannidis J. Contradicted and initially stronger effects in highly cited clinical research. JAMA. 2005;294(2):218-228.
75. Heiberg P. Studier over den statistiske undersøgelsesmetode som hjælpemiddel ved terapeutiske undersøgelser [Studies on the statistical study design as an aid in therapeutic trials]. Bibl Laeger. 1897;89:1-40.
76. Korn EL, McShane LM, Freidlin B. Statistical challenges in the evaluation of treatments for small patient populations. Sci Transl Med. 2013;5(178):178sr173.
77. James Lind Library. The Library and Information Services Department, The Royal College of Physicians of Edinburgh; 2003. Available: http://wwwjameslindlibraryorg/.
78. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev. 2012;12:MR000033.
79. Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, et al. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis - a simulation study. PLoS ONE. 2011;6: e25491.
80. +AllTrials All Trials Registered, All Results Reported. Avaiable: http://www.alltrials.net.
81. Shao S-CC, Hansheng W, Jun. Sample Size Calculations in Clinical Research, Second Edition (Chapman & Hall/CRC Biostatistics Series): Chapman and Hall/CRC. Pages; 1-480; 2007.
82. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Int Med. 2010;152(11):726-732.
83. Scales DC, Rubenfeld GD. Estimating sample size in critical care clinical trials. J Crit Care. 2005;20(1):6-11.
84. Turner RM, Bird SM, Higgins JP. The Impact of Study Size on Meta-analyses: Examination of Underpowered Studies in Cochrane Reviews. PLoS ONE. 2013;8(3):e59202.
85. Pereira TV, Horwitz RI, Ioannidis JP. Empirical evaluation of very large treatment effects of medical interventions. JAMA. 2012;308:1676-1684.
86. Gugiu PC, Gugiu MR. A critical appraisal of standard guidelines for grading levels of evidence. Eval Health Prof. 2010;33(3):233-255.