

The Necessity of Randomized Clinical Trials

Janus Christian Jakobsen^{1,2*} and Christian Gluud¹

¹The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen, Denmark

²Emergency Department, Holbæk Hospital, Denmark

ABSTRACT

Aims: ~~To assess. 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: ~~A Systematic 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~~ ~~issues that can hinder the~~ ~~to the~~ conduct of randomized clinical trials can be lessened through education, collaboration, infrastructure, and other measures. These arguments validate why the randomized clinical trial should and must be the study design evaluating new 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 ~~cannot usually and should never~~
~~sufficiently~~ validate intervention effects — randomized clinical trials are always needed.
Therefore, randomize the first patient as Thomas C Chalmers suggested in 1977.

Keywords: evidence-based medicine; randomized clinical trials; observational studies; clinical research; clinical experience; intervention research

Formatted: None, Space Before: 0 pt, Line spacing: Double, Don't keep with next, Don't keep lines together

Formatted: Left, None, Space Before: 0 pt, After: 0 pt, Line spacing: Double, Don't keep with next, Don't keep lines together

Formatted: Font: Not Bold

Formatted: None, Space Before: 0 pt, Line spacing: Double, Don't keep with next, Don't keep lines together

* Corresponding author: Janus Christian Jakobsen Tel.: +45 26186242.
E-mail address: 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 a number of publications state that observational studies in some circumstances can adequately validate intervention effects [4-7]; and clinicians often argue that their clinical experience sufficiently can assess the effects of some interventions [7]. Conducting observational studies require much less work and resources than conducting randomized clinical trials, and randomized clinical trials are often perceived as ~~complex and~~ 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, psychotherapy, etc. effect [8-12]s. We are convinced that Thomas C. Chalmers was correct when he stated that we should always randomize the first patient [13]. 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 We also offer 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 provide 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 (i.e., a

treatment backed by sufficient evidence) or placebo. The latter trials compare the effects of two active 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 accordingly. 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 which experimental intervention and is being assessed against which control or interventions being assessed 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^[14].

2. METHODS AND RESULTS

2.1 Five arguments demonstrating the fundamental need of randomized clinical trials to validate intervention effects

2.1.1 Development of interventions is a prospective process subheading

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. ~~When a researcher wants to assess if an intervention is effective or not, an observational design should therefore never be used.~~ We fully agree with Thomas C. Chalmers when he in 1977 wrote that we should always randomize the first patient "Randomize the first patient"^[13].

Formatted: Font: Not Bold

Formatted: Font: Not Bold

~~suggesting that we should always randomise the first patient. Accordingly, w~~
~~When an~~
~~investigator-researcher wants to assess if an intervention is effective or not, an observational~~
~~design should therefore never be used for the initial assessment of the intervention. We will~~
~~in the paragraphs below consider~~~~address~~ if there are exemptions to this rule.

~~Large well-conducted observational studies can sometimes provide useful information about~~
~~rare adverse events and intervention effects.~~ [15]. ~~and~~~~W~~~~W~~~~e~~ 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 priori identify such rare instances. It is only in retrospect it may be concluded that
interventions ~~would~~ have been validly assessed by observational studies [5], and evidence
based on observational evidence will in most circumstances be uncertain [16-18].

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~~ [18]. ~~and both circumstances will pose problems after an initial~~
~~assessment.~~ ~~If~~ When an intervention is ~~already~~ implemented ~~into~~ clinical ~~practically~~ ~~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~~ When 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 and~~ development of drugs, devices, and other interventions
are in essence a prospective process and the correct research design has to be selected
prospectively [19]. ~~The correct design decision ought to be the randomized clinical controlled~~
~~trial~~ [13].

2.1.2 Implementation of scientific results into clinical practice

Formatted: Font: Not Bold

Formatted: Font: Not Bold

If an intervention offers more benefit than harm ~~compared with and is superior to~~ 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 it can reach the global market fast if 'proved' effective [20]. Many medical devices have, for example, been implemented into clinical practice on the basis of observational evidence alone [21]. 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 [16-18]. 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 by~~ The Cochrane Collaboration Handbook [1]. Even if an intervention has an almost parachute-like beneficial intervention effect [22], a fast way to the global market might be blocked if the intervention is only assessed in observational studies. ~~Although more complex, t~~he results of properly conducted large randomized clinical trials will be more readily accepted by more clinicians than results from observational studies and will therefore probably offer a faster access to a larger market compared to market ~~penetration~~ implementation via an observational design.

2.1.3 Balance between beneficial and harmful effects

~~Large well-conducted observational studies can provide useful information about rare adverse events and intervention effects [9], and i~~it is theoretically possible to quantify a beneficial intervention effect size via observational evidence if ~~thea~~ disease is stable and without any fluctuation in symptoms and if the intervention effects are large enough to be recognized by 'observation'. However, very few diseasesconditions show such stability and

128 interventions with large easily clearly observable effects occur extremely rarely^[14]are almost
129 never occurring. Most interventions have no beneficial effects or relatively small effects. It is
130 among the latter we shall find the interventions of tomorrow. Moreover, large 'surprising'
131 beneficial effects shown in observational studies may be due to random errors, systematic
132 errors, or confounding, and rRandomized clinical trials are, therefore, needed to assess
133 when potential beneficial effects outweigh the potential harmful effects. Randomization is
134 able to construct the perfect control, which, at baseline, becomes fully comparable to the
135 experimental group regarding all known and all unknown prognostic factors——provided that
136 the randomized groups become large enough. Without randomization and without an
137 appropriate control group it is often unclear if a change in symptoms is caused solely by an
138 intervention effect — or if some, or all, of the change is a natural fluctuation of the symptoms
139 (often a combination of 'regression towards the mean' and the natural fluctuation of the
140 symptoms). Observational studies including some kind of matched control group do not
141 provide valid information about effect sizes, because the participants in the control group will
142 almost never be fully comparablecomparablecompletely similar to the participants in the
143 experimental group^[18]. It is therefore impossible to quantify and have an overview of the
144 relative effect sizes via observational evidence only (Box 1).

Formatted: Font: Bold

153 BOX 1

It can be 'observed' that an operation for heartburn can normalize pH in the oesophagus[23]. but the surgical procedure also carry some risks[24,25]. Observational evidence cannot assess when the degree of heartburn justifies an operation with possible harmful effects[25]. Furthermore, without randomization it is unclear whether a change in symptoms is caused by the operation or by other factors.

Long-acting beta₂-agonists can improve lung function in asthma patients[26]. but after a large number of participants have been assessed evidence has indicated that long-acting beta₂-agonists also cause a small increase in mortality[26]. 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 beta₂-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 an appropriate control group of in a randomized clinical trial in order to make valid inferences if effect sizes are to be assessed. 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 placebo effects may be unknown will often be unknown. However, placebo has often very small effects or no effects compared with no intervention[27] and placebo-controlled trials will therefore often demonstrate the effect sizes of the experimental interventions. Ethical issues can be a hindrance, but R randomized clinical trials assessing the effects of experimental interventions versus placebo no intervention are therefore in general the optimal method to accurately assess effect sizes (Table 1). If effective treatments exist, then such treatments may either be used

Formatted: Left, Space After: 0 pt, Line spacing: Double, Don't keep with next

Formatted: Line spacing: Double

Formatted: Left, Space After: 0 pt, Line spacing: Double, Don't keep with next

Formatted: Line spacing: Double

Formatted: Left, Space After: 0 pt, Line spacing: Double, Don't keep with next

Formatted: Line spacing: Double

Formatted: Font: Bold

as the control intervention or as basis treatment for participants in all of the trial intervention groups, i.e., an experimental intervention may then be assessed as an add-on intervention to one of the intervention groups. Here The Declaration of Helsinki and medical regulatory agencies have been too kind to the product and ignored the patient[28-30].

We have in **Table 2** presented an overview of the different types of randomized trials and summarized the corresponding methodological strengths and limitations. It is impossible to quantify and have an overview of the relative effect sizes via observational evidence only (Box 1). If one is not able to assess the balance between benefits and harms it is impossible to assess the clinical significance of a preventive, prognostic, diagnostic, or therapeutic intervention.

Formatted: Font: Bold

Formatted: Font: Not Bold

BOX 1

It can be 'observed' that an operation for heartburn can normalize pH in the oesophagus.[19] but the procedure also carry some risks.[20,21] Observational evidence cannot assess when the degree of heartburn justifies an operation with possible harmful effects.[21] Furthermore, without a control group it is unclear whether a change in symptoms is caused by the operation or by other factors.

Long-acting beta₂-agonists can improve lung function in asthma patients,[22] but after a large number of participants have been assessed evidence has indicated that long-acting beta₂-agonists also cause a small overall increase in mortality.[22] Such rare harmful effects would be impossible to detect without randomized trials. It would be unclear whether the relatively few deaths were caused by the long-acting beta₂-agonists or by other factors.

Studies have shown that observational studies compared to randomized clinical trials often overestimate benefits and underestimate harms, i.e., produce biased results [16-18]. 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 on the same interventions show significantly less biased results [31,32]. A valid and unbiased assessment of benefits and

Formatted: Line spacing: Double

Formatted: Font: Not Bold

Formatted: Font: Not Bold

harms are impossible to achieve in an observational design where blinding usually is impossible.

2.1.4 Patient relevant and clinically relevant and patient relevant outcomes

Intervention effects on ~~clinically relevant and~~ 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 [27]. 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 ~~It can be observed by a clinician that a~~ laser intervention can reduce redness of ~~at the otherwise non-disappearing~~ 'port-wine stain' on the skin of a patient;[33]. or that chemotherapy seems to prolong survival in incurable cancer patients.[34]. 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[33] and 'quality of life' and QUALY (quality adjusted life years) of the cancer patients.[35]. These outcomes are impossible or difficult to assess only by clinical 'observation'.

Formatted: Font: Bold

Formatted: Font: Not Bold

Formatted: Line spacing: Double, Don't keep with next

Formatted: None, Space Before: 0 pt, Line spacing: Double, Don't keep with next, Don't keep lines together

Formatted: Line spacing: Double

2.1.5 Indications for an intervention

Most diseases have varying degrees of severity. When diseases are 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 [1,36]. [However, because of concerns of multiplicity and of small sample sizes often involved, subgroup analyses should be viewed](#) [are best viewed only as a hypothesis generating exercises](#) [37,38]. If [such](#) 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 [39].

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.[40]. Without randomized clinical trials it is, [e.g.](#), not apparent how severe the hypoxia should be before performing tracheostomy.[40].

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.[41].

Formatted: Left, None, Space Before: 0 pt, After: 0 pt, Line spacing: Double, Don't keep with next, Don't keep lines together

Formatted: Left, Space After: 0 pt, Line spacing: Double, Don't keep with next

Formatted: None, Space Before: 0 pt, Line spacing: Double, Don't keep with next, Don't keep lines together

2.2 Typical hindrances for the conduct of randomized clinical trials and some remedies to reduce these

~~It is clear that~~ 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 randomized trials are conducted. It is, e.g., often stated that trial populations are not representative of patients in the clinic [4,42,43]. 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 [44,45]. 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 [46]. 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 [47] can lessen and solve some of the many problems conducting randomized clinical trials [48-53]. 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 [54], and The Nordic Trial Alliance (NTA) [55]) may reduce problems with recruitment of a sufficient number of trial participants etc. [56,57]. 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

Formatted: Line spacing: Double

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Formatted: Font: 10 pt

255 [randomized clinical trials usually gives a reasonable or high return on investment](#) [58].
256 [Politicians and decision makers must be taught the key positions of the randomized clinical](#)
257 [trial and of systematic reviews of such trials in clinical intervention research.](#)
258
259 We have in **Table 1** listed typical issues [and misconceptions](#) that are perceived or realized
260 as obstacles for the conduct of randomized clinical trials and pointed out how the problems
261 may be minimized.

TABLE 1

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 randomised clinical trials.	Potential solutions: Investigators Trialists 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: A-r 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 rather than s single centre trials and through the use of broad inclusion criteria and appropriately selected exclusion criteria trials . [46,59].
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 [47] 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, no intervention	Potential solutions: It may be is unethical to treat patients with interventions that are not evidence-based. 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 one of the experimental intervention groups versus placebo as an 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 can easily be ethically justified.
Typical misconception: Trial participants differ from patients in common clinical settings. [4,42,43]. Strict inclusion and exclusion criteria are believed to put together trial populations not representative of patients in the clinic questioning the clinical relevance of results from randomized clinical trials. [4,42,43].	Counter argument: It is not necessary to use narrow criteria for selecting trial participants. [1,45,46]. Using fewer inclusion and exclusion criteria will also make trial populations more similar to patients in the clinic. Moreover, patients that receive similar treatments and i interventions within and outside randomized clinical trials seem to have similar prognosis. [44,45].
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-	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. [44,45,62]. Moreover, it is not necessary to use

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:[4,60,61].	strict treatment protocols in a randomized 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:[20,63].	Counter argument: Standardized interventions based on evidence-based practice are most often superior to non-standardized interventions:[64-67]. 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: It has been calculated that investment in randomized clinical trials usually gives a reasonable or high return on investment:[58]. 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.

264

265

266
267

TABLE 2

Formatted: Font: 11 pt, Bold

Different types of control groups in randomized clinical trials					
<u>Experimental intervention versus no intervention</u>		<u>Experimental intervention versus placebo, impure placebo*, 'active' placebo (nocebo)**, or a sham intervention</u>		<u>Experimental intervention versus 'treatment as usual'***</u>	
<u>Methodological strengths (+) and limitations (-)</u>		<u>Methodological strengths (+) and limitations (-)</u>		<u>Methodological strengths (+) and limitations (-)</u>	
+	-	+	-	+	-
<u>The beneficial and harmful effects of the experimental intervention can be shown by the results.</u>	<u>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.</u>	<u>Allows blinding of trial participants; investigators; treatment providers; outcome assessors; data managers; statisticians; and conclusion drawers.</u> <u>Allows assessment of experimental intervention effect sizes controlling for non-specific treatment factors****.</u>	<u>The 'effect' of placebo may be unclear in certain conditions.</u> <u>Participants can often because of beneficial effects or adverse effects figure out if they are treated with the active intervention or the control intervention.</u>	<u>The trial results demonstrate what a given average patient gains by an experimental intervention compared with the treatment the patient usually receive.</u>	<u>Treatment as usual most often contains some non-specific treatment elements with unknown effects.</u> <u>Results may be biased as no blinding is involved, unless one uses double placebo ('double dummy').</u>
<p align="center"><u>Co-interventions</u></p> <p><u>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</u></p>					

268
269
270
271
272
273

* 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).

** 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 of different non-specific interventions

274 | [**** Trial participants might benefit from, e.g., believing that an intervention is effective or just from being in contact with a treatment provider. Placebo-](#)
275 | [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-](#)
276 | [specific treatment factors.](#)

277

278

279

280

281
282 Researchers can be reluctant to conduct randomized clinical trials because they are costly
283 and time consuming. Lack of methodological and statistical know-how can hinder the making
284 of randomized clinical trials; it can be difficult to recruit enough trial participants, etc. We
285 acknowledge all of these difficulties. Nevertheless, the establishment of industry
286 independent trial units with know-how about evidence-based medicine [32] can lessen and
287 solve some of the many problems conducting randomized trials [46-51]. Furthermore,
288 regional, national, international, and global research collaboration between trial units and
289 clinical sites (e.g., The European Clinical Research Infrastructures (ECRIN), The UK Clinical
290 Research Collaboration (UKCRC) Clinical Trials Units Network [52], and The Nordic Trial
291 Alliance (NTA) [53]) may reduce problems with recruitment of a sufficient number of trial
292 participants and other problems [54,55]. Well-conducted multicentre trials also offer better
293 external validity than well-conducted single-centre trials. It must be recognized how much
294 health care costs can be reduced if patient treatment becomes more effective through
295 evidence-based research. It has been calculated that investment in randomized clinical trials
296 usually gives a reasonable or high return on investment [45]. Politicians and decision makers
297 must be taught the key position of the randomized clinical trial in clinical intervention
298 research.
299

300 3. DISCUSSION

301
302 We believe that clinical experience and observational studies cannot and should not validate
303 the effects of new interventions effects. Observational studies can sufficiently assess
304 associations between certain interventions and outcomes, but the randomized clinical
305 trials are always needed to avoid falsely negating (type I error) or falsely confirming (type II
306 error) the null hypothesis and to assess causality between interventions and outcomes, i.e.,
307 randomized clinical trials are needed to sufficiently validate intervention effects sufficiently
308 validate intervention effects. Typical issues hindering the conduct of trials can be overcome
309 (Table 1).

310
311 A report from the Patient-Centered Outcomes Research Institute was recently published for
312 public comment [68]. This report claims that the use of observational studies to make causal
313 inference is potentially much stronger than it has been in the past [68], and similar
314 arguments are often published in highly esteemed journals [3-7,69]. We believe that the
315 fundamental construct of the observational studies limits the reliability of the results from
316 observational studies [16,18]. To assess if an intervention causes more benefit than harm

Formatted: Line spacing: Double

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 [18]. [Theyhis report](#) 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 [18]. Compared to small randomized [clinical](#) trials, small observational studies often showed effects that were far from the 'true' intervention effect [18]. Ioannidis and colleagues [have](#) also observed that significant discrepancies do occur between the results of randomized clinical trials and observational studies[16][\[44,43\]](#)— and that results from observational studies are more often contradicted than results from randomized [clinical](#) trials [70]. [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 hinderthe conduct of randomized clinical trials. Observational studies can of course have their place in such circumstances other research settings 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. As mentioned, it may, as mentioned, be possible to present a few historical examples where intervention effects have been sufficiently validated by observational evidence\[5\]. However, the \[ese exceptions is\]\(#\) does not justify that observational evidence \[generally\]\(#\) should be used prospectively to validate intervention effects. \[As it has been clearly expressed by Heiberg already in 1897 and reiterated by others both before and since\]\(#\)\[71-73\]\[—r\]\(#\)Regarding the vast majority of \[present-day\]\(#\) 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 increasing sample sizes of the trials,](#)
[number of outcomes \(reducing the risks of random errors\), as well as with improved quality](#)
[of the methodology \(reducing the risks of systematic errors\)\[1,31,32,36,74,75\].Moreover, the](#)
[conduct of systematic reviews assessing all randomized clinical trials on an intervention as](#)
[conducted by The Cochrane Collaboration reduces these risks \[1,36,74,75\]. 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 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; 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\[1,76\].](#)

It may be frustrating for clinicians to realize that clinical experience [and observational studies](#)
[does](#) not provide valid knowledge about intervention effects — especially because many
interventions [in clinical use](#) have not been assessed in randomized [clinical](#) trials[68]. [We aim](#)
[to support the development and use of effectivehealth-careinterventionsto 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 aBut-as](#)
rational clinicians ~~we-must-consequently~~ realize the uncertainty of our knowledge if
randomized [clinical](#) trials have not been conducted [and remember the validity of the](#)
[evidential hierarchy\[77\].- Systematic reviews of randomized clinical trials is and should be](#)
[considered the highest level of evidence followed by single randomized trials\[77\]. This does](#)

~~not necessarily mean~~ 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 [1,36].

4. CONCLUSION

Clinical experience or observational studies cannot sufficiently validate intervention effects — randomized clinical trials are always needed. We therefore ~~strongly~~ disagree with authors claiming that observational designs can be employed for assessing ~~new~~ interventions.

ACKNOWLEDGEMENTS

We thank Jørn Wetterslev for helpful discussions. ~~We have received no funding.~~

COMPETING INTERESTS

Authors have declared that no competing interests exist. We have received no specific funding.

AUTHORS' CONTRIBUTIONS

Both authors conceived the study, ~~and~~ contributed to the writing of the paper, and are ~~JCJ~~ is guarantors.

397 REFERENCES

- 398 1. Higgins JPT, Green S. The Cochrane Handbook for Systematic Reviews of
399 Interventions, Version 5.1.0. The Cochrane Collaboration. 2011;Available
400 from <http://www.cochrane-handbook.org>.
401
- 402 2. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, et al.
403 GRADE guidelines 11-making an overall rating of confidence in effect
404 estimates for a single outcome and for all outcomes. *J Clin Epidemiol*.
405 2013;66(2):151-157.
- 406 3. Ulvenes LV, Aasland O, Nylenna M, Kristiansen IS. Norwegian physicians'
407 knowledge of and opinions about evidence-based medicine: cross-sectional
408 study. *PloS one*. 2009;4(11):e7828.
- 409 4. Vincent JL. We should abandon randomized controlled trials in the intensive care
410 unit. *Crit Care Med*. 2010;38(10 Suppl):S534-538.
- 411 5. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials
412 unnecessary? Picking signal from noise. *BMJ*. 2007;334(7589):349-351.
- 413 6. Stolberg HO, Norman G, Trop I. Randomized controlled trials. *Am J Roentgenol*.
414 2004;183(6):1539-1544.
- 415 7. Hawkins NG, Sanson-Fisher RW, Shakeshaft A, D'Este C, Green LW. The
416 multiple baseline design for evaluating population-based research. *Am J*
417 *Prev Med*. 2007;33(2):162-168.
- 418 8. Gluud C, Kubiak C, Whitfield K, Byrne J, Huemer KH, Thirstrup S, et al. Typical
419 investigational medicinal products follow relatively uniform regulations in 10
420 European Clinical Research Infrastructures Network (ECRIN) countries.
421 *Trials*. 2012;1327.
- 422 9. Kramer DB, Xu S, Kesselheim AS. How does medical device regulation perform
423 in the United States and the European union? A systematic review. *PLoS*
424 *Med*. 2012;9(7):e1001276.
- 425 10. Feltham C (1997) Book: Which Psychotherapy?: Leading Exponents Explain
426 Their Differences (page 1-224): SAGE Publications Ltd.
- 427 11. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, et
428 al. No surgical innovation without evaluation: the IDEAL recommendations.
429 *Lancet*. 2009;374(9695):1105-1112.
- 430 12. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet*.
431 2011;377(9778):1693-1702.
- 432 13. Chalmers TC. Randomize the first patient. *N Engl J Med*. 1977;296(2):107.
- 433 14. Chalmers I, Milne I, Trohler U, Vandenbroucke J, Morabia A, Tait G, et al. The
434 James Lind Library: explaining and illustrating the evolution of fair tests of
435 medical treatments. *The journal of the Royal College of Physicians of*
436 *Edinburgh*. 2008;38(3):259-264.
- 437 15. Vandenbroucke JP. What is the best evidence for determining harms of medical
438 treatment? *CMAJ*. 2006;174(5):645-646.
- 439 16. Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, et al.
440 Comparison of evidence of treatment effects in randomized and
441 nonrandomized studies. *JAMA*. 2001;286(7):821-830.
- 442 17. Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms
443 of medical interventions in randomized and nonrandomized studies. *Can*
444 *Med Assoc J*. 2006;174(5):635-641.

- 445 18. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, et al.
446 Evaluating non-randomised intervention studies. *Health Technol Assess.*
447 2003;7(27):iii-x, 1-173.
- 448 19. Chalmers I. Well informed uncertainties about the effects of treatments. *BMJ.*
449 2004;328(7438):475-476.
- 450 20. Cooper J. Randomized clinical trials for new surgical operations: square peg in a
451 round hole? *J Thorac Cardiovasc Surg.* 2010;140(4):743-746.
- 452 21. Huot L, Decullier E, Maes-Beny K, Chapuis FR. Medical device assessment:
453 scientific evidence examined by the French national agency for health - a
454 descriptive study. *BMC public health.* 2012;12585.
- 455 22. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to
456 gravitational challenge: systematic review of randomised controlled trials. *Int*
457 *J Prosthodont.* 2006;19(2):126-128.
- 458 23. Cuschieri A, Hunter J, Wolfe B, Swanstrom LL, Hutson W. Multicenter
459 prospective evaluation of laparoscopic antireflux surgery. Preliminary report.
460 *Surg Endosc.* 1993;7(6):505-510.
- 461 24. Kahrilas PJ, Lin S, Spiess AE, Brasseur JG, Joehl RJ, Manka M. Impact of
462 fundoplication on bolus transit across esophagogastric junction. *Am J*
463 *Physiol.* 1998;275(6 Pt 1):G1386-1393.
- 464 25. Wileman SM, McCann S, Grant AM, Krukowski ZH, Bruce J. Medical versus
465 surgical management for gastro-oesophageal reflux disease (GORD) in
466 adults. *Cochrane database of systematic reviews (Online).*
467 2010;(3):CD003243.
- 468 26. Cates C, Cates M (2008) Regular treatment with salmeterol for chronic asthma:
469 serious adverse events. *Cochrane Database of Systematic Reviews: John*
470 *Wiley & Sons, Ltd.*
- 471 27. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions.
472 *Cochrane DatabaseSystRev.* 2010;(1):CD003974.
- 473 28. World Medical Association. Declaration of Helsinki. Ethical principles for medical
474 research involving human subjects. *J Indian Med Assoc.* 2009;107(6):403-
475 405.
- 476 29. Bertele V, Banzi R, Gluud C, Garattini S. EMA's reflection on placebo does not
477 reflect patients' interests. *Eur J Clin Pharmacol.* 2012;68(5):877-879.
- 478 30. Garattini S, Bertele V, Banzi R. Placebo? no thanks, it might be bad for me! *Eur*
479 *J Clin Pharmacol.* 2013;69(3):711-714.
- 480 31. Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of
481 reported study design characteristics on intervention effect estimates from
482 randomized, controlled trials. *Ann Intern Med.* 2012;157(6):429-438.
- 483 32. Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical
484 evidence of bias in treatment effect estimates in controlled trials with
485 different interventions and outcomes: meta-epidemiological study. *BMJ.*
486 2008;336(7644):601-605.
- 487 33. Faurschou A, Olesen AB, Leonardi-Bee J, Haedersdal M. Lasers or light
488 sources for treating port-wine stains. *Cochrane database of systematic*
489 *reviews (Online).* 2011;11CD007152.
- 490 34. Boulaamane L, Boutayeb S, Errihani H. Bevacizumab based chemotherapy in
491 first line treatment of HER2 negative metastatic breast cancer: results of a

- Moroccan observational institutional study. BMC research notes. 2012;5(1):162.
35. 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;2732.
36. 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;1090.
37. Magnusson BP, Turnbull BW. Group sequential enrichment design incorporating subgroup selection. Stat Med. 2013;doi: 10.1002/sim.5738.
38. 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.
39. 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.
40. 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.
41. 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 Engl J Med. 2011;365(9):787-797.
42. Pincus T. Limitations of randomized clinical trials to recognize possible advantages of combination therapies in rheumatic diseases. Semin Arthritis Rheum. 1993;23(2 Suppl 1):2-10.
43. Cartwright N, Munro E. The limitations of randomized controlled trials in predicting effectiveness. J Eval Clin Pract. 2010;16(2):260-266.
44. 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.
45. Vist G, Bryant D, Somerville L, Birmingham 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 of Syst Rev. 2008;Jul 16;(3):MR000009.
46. 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.
47. 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.
48. Fleischmann R. Primer: establishing a clinical trial unit--obtaining studies and patients. Nat Clin Pract Rheumatol. 2007;3(8):459-463.
49. Fleischmann R. Primer: establishing a clinical trial unit - regulations and infrastructure. Nat Clin Pract Rheumatol. 2007;3(4):234-239.
50. 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.

51. Ohmann C, Bruns I, Wolff S. The Coordinating Centers for Clinical Trials (KKS) and the KKS Network: competence for clinical research. *Onkologie*. 2010;33 Suppl 711-15.
52. 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.
53. Gluud C, Sorensen TI. Organisation of a clinical trial unit—a proposal. *Fundam Clin Pharmacol*. 1998;12(3):298-305.
54. The UK Clinical Research Collaboration (UKCRC). <http://www.ukcrc-ctu.org.uk/>.
55. Nordic Trail Alliance (NTA). <http://www.nordforsk.org/no/programs/nordic-trial-alliance-nta>.
56. Demotes-Mainard J, Ohmann C. European Clinical Research Infrastructures Network: promoting harmonisation and quality in European clinical research. *Lancet*. 2005;365(9454):107-108.
57. Demotes-Mainard J, Ohmann C, Gluud C, Chêne 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.
58. 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. http://www.wellcome.ac.uk/stellent/groups/corporatesite/@sitestudioobjects/document_s/web_document/wtx052110pdf.
59. Nichol AD, Bailey M, Cooper DJ, Polar, Investigators EPO. Challenging issues in randomised controlled trials. *Injury*. 2010;41 Suppl 1S20-23.
60. Joffe S, Weeks JC. Views of American oncologists about the purposes of clinical trials. *J Natl Cancer Inst*. 2002;94(24):1847-1853.
61. 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.
62. 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.
63. Appelbaum P, Roth L, Lidz C, Benson P, Winslade W. False hopes and best data: consent to research and the therapeutic misconception. *The Hastings Center Report*. 1987;(17(2):20-4).
64. 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.
65. 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.
66. 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.

- 588 67. Karjalainen S, Palva I. Do treatment protocols improve end results? A study of
589 survival of patients with multiple myeloma in Finland. BMJ.
590 1989;299(6707):1069-1072.
- 591 68. Helfand M, Berg A, Flum D, Gabriel S, Normand S. Draft Methodology Report:
592 "Our Questions, Our Decisions: Standards for Patient-Centered Outcomes
593 Research" (Page 1-186). Patient-Centered Outcomes Research Institute.
594 2012;Published for Public Comment July 23 (Available at :
595 <http://www.pcori.org/assets/Preliminary-Draft-Methodology-Report.pdf>).
- 596 69. Higgins J. Convincing evidence from controlled and uncontrolled studies on the
597 lipid-lowering effect of a statin [editorial]. Cochrane Database of Systematic
598 Reviews.
599 2012;12(ED000049):[http://www.thecochranelibrary.com/details/editorial/3992221/](http://www.thecochranelibrary.com/details/editorial/3992221/Convincing-evidence-from-controlled-and-uncontrolled-studies-on-the-lipid-loweri.html)
600 [Convincing-evidence-from-controlled-and-uncontrolled-studies-on-the-lipid-](http://www.thecochranelibrary.com/details/editorial/3992221/Convincing-evidence-from-controlled-and-uncontrolled-studies-on-the-lipid-loweri.html)
601 [loweri.html](http://www.thecochranelibrary.com/details/editorial/3992221/Convincing-evidence-from-controlled-and-uncontrolled-studies-on-the-lipid-loweri.html).
- 602 70. Ioannidis J. Contradicted and initially stronger effects in highly cited clinical
603 research. JAMA. 2005;294(2):218-228.
- 604 71. Heiberg P. Studier over den statistiske undersøgelsesmetode som hjælpemiddel
605 ved terapeutiske undersøgelser [Studies on the statistical study design as an
606 aid in therapeutic trials]
607 ([http://www.jameslindlibrary.org/illustrating/records/studier-over-den-statistiske-](http://www.jameslindlibrary.org/illustrating/records/studier-over-den-statistiske-undersogelsesmetode-som-hjaelpemid/key_passages)
608 [undersogelsesmetode-som-hjaelpemid/key_passages](http://www.jameslindlibrary.org/illustrating/records/studier-over-den-statistiske-undersogelsesmetode-som-hjaelpemid/key_passages)). Bibl Laeger. 1897;891-
609 40.
- 610 72. Korn EL, McShane LM, Freidlin B. Statistical challenges in the evaluation of
611 treatments for small patient populations. Science translational medicine.
612 2013;5(178):178sr173.
- 613 73. The Library and Information Services Department, The Royal College of
614 Physicians of Edinburgh. James Lind Library. Available online at:
615 <http://www.jameslindlibrary.org/>. 2003.
- 616 74. Lundh A, Sismondo S, Lexchin J, Busuioac OA, Bero L. Industry sponsorship and
617 research outcome. Cochrane database of systematic reviews (Online).
618 2012;12MR000033.
- 619 75. Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, et al. The
620 number of patients and events required to limit the risk of overestimation of.
621 PLoS one. 2011;6e25491.
- 622 76. +AllTrials (avaialbe at: <http://www.alltrials.net>). All Trials Registered, All Results
623 Reported.
- 624 77. Gugiu PC, Gugiu MR. A critical appraisal of standard guidelines for grading
625 levels of evidence. Eval Health Prof. 2010;33(3):233-255.
- 626
627
628

Formatted: Line spacing: 1.5 lines

Formatted: Line spacing: 1.5 lines