

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: 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 hindrance 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 should never 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*

* Corresponding author: Janus Christian Jakobsen Tel.: +45 26186242.
E-mail address: jcyj@ctu.dk

20 1. INTRODUCTION

21

22 Observational studies, such as non-randomized cohort studies or patient series, are usually
23 viewed as producing results with less evidential weight compared to the results from
24 randomized clinical trials [1,2]. However, quite often clinicians argue that their clinical
25 experience sufficiently can assess the effects of some interventions [3] and some
26 publications state that observational studies can adequately validate intervention effects [4-
27 7]. Conducting observational studies require much less work and resources than conducting
28 randomized clinical trials, and randomized clinical trials are often perceived as bureaucratic
29 and difficult to conduct. Therefore, it is no surprise that many investigators choose
30 observational studies to try to assess intervention effects.

31

32 We will in the following paragraphs consider if randomized clinical trials always are
33 necessary and the best clinical study design to assess any kind of health-care intervention,
34 including drugs, medical devices, surgery, psychotherapy, etc.[8-12]. We are convinced that
35 Thomas C. Chalmers was correct when he stated that we should always randomize the first
36 patient [13]. However, we also acknowledge the difficulties that randomized clinical trials
37 may cause and that they too may show erroneous results. We will, therefore, in the second
38 part of the manuscript provide a list of the typical issues that represents a perceived or real
39 hindrance for the conduct of randomized clinical trials and we will suggest some remedies to
40 reduce these hindrances.

41

42 Randomized clinical trials cannot only assess the effects of many different forms of
43 experimental interventions, but also many different forms of control interventions, e.g., no
44 intervention, placebo, 'impure' placebo, nocebo, or an active control intervention (i.e., a
45 treatment backed by sufficient evidence). The latter trials compare the effects of two
46 interventions (so-called head-to-head trials or comparative intervention research). It is clear

47 that the inferences of the results from the different forms of trials differ accordingly. We will in
48 the following paragraphs use the term 'randomized clinical trials' as a collective term for all
49 kinds of trials, as we believe that the fundamental principles are similar regardless of type of
50 experimental intervention and control intervention. The fundamental construct of the
51 randomized clinical trial allows that any intervention using quantitative or qualitative
52 outcomes can be assessed using the same basic principles[14].

53

54 **2. METHODS AND RESULTS**

55

56 **2.1 Five arguments demonstrating the fundamental need of randomized clinical trials** 57 **to validate intervention effects**

58

59 **2.1.1 Development of interventions is a prospective process**

60

61 It is important to make the correct choice of study design before the initial assessment of a
62 new intervention. The optimal indication, effect size, and balance between harmful and
63 beneficial effects (see the paragraphs below) will remain unknown if randomized clinical
64 trials are not conducted before an intervention is implemented into clinical practice. We fully
65 agree with Thomas C. Chalmers when he in 1977 wrote that we should always randomize
66 the first patient[13]. Accordingly, when an investigator wants to assess if an intervention is
67 effective or not, an observational design should never be used for the initial assessment of
68 the intervention. We will in the paragraphs below consider if there are exemptions to this
69 rule.

70

71 Large well-conducted observational studies can sometimes provide useful information about
72 rare adverse events and intervention effects[15]. We acknowledge a few historical instances
73 where observational evidence validly have demonstrated benefits of new interventions (e.g.,

74 insulin for diabetic coma and ether for anaesthesia) [5]. However, we cannot a priori identify
75 such rare instances. It is only in retrospect it may be concluded that interventions have been
76 validly assessed by observational studies [5], and evidence based on observational
77 evidence will in most circumstances be uncertain [16-18]. Observational studies will often
78 either grossly overestimate or underestimate intervention effects and adjustment with
79 statistical analyses (logistic regression or propensity score) only seem to increase the
80 problem[18].If an intervention is implemented into clinical practicebased on observational
81 evidence and seems to work, it can be difficult to justify and to conduct randomized clinical
82 trials assessing the correct balance between benefits and harms. In this situation, we may
83 never know the 'true' balance between benefits and harms. If an intervention does not look
84 rewarding in an observational study we will likely stop further assessment of the intervention
85 and therefore risk 'throwing the baby out with the bath water'. Intervention research during
86 the development of drugs, devices, and other interventions are in essence a prospective
87 process and the correct research design has to be selected prospectively [19]. The correct
88 design ought to be the randomized clinical trial[13].

89

90 **2.1.2 Implementation of scientific results into clinical practice**

91

92 If an intervention offers more benefit than harm compared with previous treatment options, it
93 is an ethical obligation and hence necessary to get that intervention offered to as many
94 patients as possible, as fast as possible. In the discussion about choice of design for
95 assessing new interventions, investigators often claim that it is important to conduct a quick
96 observational study so it can reach the global market fast if 'proved' effective [20]. Many
97 medical devices have, for example, been implemented into clinical practice on the basis of
98 observational evidence alone[21]. However, if only observational evidence backs the
99 intervention it may be difficult to reach clinical consensus about a given intervention effect
100 because clinicians might rightly question the validity of such results[16-18]. It is much more

101 easy to reach clinical consensus based on results from randomized clinical trials preferably
102 assessed in systematic reviews ad modum those conducted according to The Cochrane
103 Collaboration Handbook [1]. Even if an intervention has an almost parachute-like beneficial
104 intervention effect [22], a fast way to the global market might be blocked if the intervention is
105 only assessed in observational studies. The results of properly conducted large randomized
106 clinical trials will be more readily accepted by more clinicians than results from observational
107 studies and will therefore probably offer a faster access to a larger market compared to
108 market penetration via an observational design.

109

110 **2.1.3 Balance between beneficial and harmful effects**

111

112 It is theoretically possible to quantify a beneficial intervention effect size via observational
113 evidence if the disease is stable and without any fluctuation in symptoms and if the
114 intervention effects are large enough to be recognized by 'observation'. However, very few
115 diseases show such stability and interventions with large easily observable effects occur
116 extremely rarely[14]. Most interventions have no beneficial effects or relatively small effects.
117 It is among the latter we shall find the interventions of tomorrow. Moreover, large 'surprising'
118 beneficial effects shown in observational studies may be due to random errors, systematic
119 errors, or confounding. Randomized clinical trials are, therefore, needed to assess when
120 potential beneficial effects outweigh the potential harmful effects. Randomization is able to
121 construct the perfect control, which, at baseline, becomes fully comparable to the
122 experimental group regarding all known and all unknown prognostic factors—provided that
123 the randomized groups become large enough. Without randomization and without an
124 appropriate control group it is often unclear if a change in symptoms is caused solely by an
125 intervention effect — or if some, or all, of the change is a natural fluctuation of the symptoms
126 (often a combination of 'regression towards the mean' and the natural fluctuation of the
127 symptoms). Observational studies including some kind of matched control group do not

128 provide valid information about effect sizes, because the participants in the control group will
129 almost never be fully comparable to the participants in the experimental group[18]. It is
130 therefore impossible to quantify and have an overview of the relative effect sizes via
131 observational evidence only (**Box 1**).

132

133 **BOX 1**

134

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.

135 Without an assessment of the balance between benefits and harms it is impossible to
136 assess the clinical significance of a preventive, prognostic, diagnostic, or therapeutic
137 intervention. It is important to use the appropriate control group of a randomized clinical trial
138 in order to make valid inferences. If a trial comparing the effects of two active interventions
139 shows no difference in effect it is not on the face of it clear whether the two interventions are
140 equally effective or equally ineffective. The interpretability of results from randomized trials

141 using placebo as control intervention will on the face of it in a similar way be unclear
142 because the placebo effects may be unknown. However, placebo has often very small
143 effects or no effects compared with no intervention[27] and placebo-controlled trials will
144 therefore often demonstrate the effects of the experimental intervention. Randomized clinical
145 trials assessing the effects of experimental interventions versus placebo are therefore in
146 general the optimal method to accurately assess effect sizes(**Table 1**).If effective treatments
147 exist, then such treatments may either be used as the control intervention or as basis
148 treatment for participants in all of the trial intervention groups, i.e., an experimental
149 intervention may then be assessed as an add-on intervention to one of the intervention
150 groups. Here The Declaration of Helsinki and medical regulatory agencies have been too
151 kind to the product and ignored the patient [28-30].

152

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

155

156

157 Studies have shown that observational studies compared to randomized clinical trials often
158 overestimate benefits and underestimate harms, i.e., produce biased results [16-18]. To
159 accurately and objectively assess the balance between benefits and harms, we need
160 randomized clinical trials with blinded outcome assessment. Blinded randomized clinical
161 trials compared to unblinded randomized clinical trials show significantly less biased results
162 [31,32]. A valid and unbiased assessment of benefits and harms are impossible to achieve in
163 an observational design where blinding usually is impossible.

164

165 **2.1.4 Patient relevant and clinically relevant outcomes**

166

167 Intervention effects on patient relevant and clinically relevant outcomes such as
168 psychological distress, quality of life, patient satisfaction, and pain are impossible to assess

169 accurately by 'observation' (**Box 2**). Such outcomes should be reported and assessed by the
170 patient and not by a clinician and are by nature subjective, fluctuating, and a placebo effect
171 can be significant [27]. Therefore, randomized clinical trials enabling blinding of all parties
172 (participants; investigators; health-care providers; outcome assessors; data managers;
173 statisticians; conclusion drawers) are mandatory to validly assess patient relevant and
174 clinically relevant outcomes [1].

175

176 **BOX 2**

177

A clinician can observe that laser intervention can reduce redness of a '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 QALY (quality adjusted life years) of the cancer patients[35]. These outcomes are impossible or difficult to assess only by clinical 'observation'.

178

179 **2.1.5 Indications for an intervention**

180

181 Most diseases have varying degrees of severity. When diseases are on the borderline
182 between severe and 'not severe', only randomized clinical trials can determine if we should
183 intervene or not. Randomized clinical trials are necessary to determine the most optimal
184 indication for an intervention — when to treat or when not to treat. We have illustrated this in
185 the two examples in **Box 3**. Randomized clinical trials, with low risk of bias, low risk of design
186 errors, and low risk of random errors can via prospectively planned subgroup analyses

187 suggest such indications [1,36]. However, because of concerns of multiplicity and of small
188 sample sizes often involved, subgroup analyses should be viewed only as hypothesis
189 generating exercises[37,38].If subgroup analyses show effect in only one or more of the
190 subgroups, then new confirmatory randomized clinical trials on these subgroups ought to be
191 conducted [39].

192

193 **BOX 3**

194

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 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].

195

196

197 **2.2 Typical hindrances for the conduct of randomized clinical trials and some**

198 **remedies to reduce these**

199

200 Conducting randomized clinical trials generally require more resources than conducting
201 observational studies. Researchers can be reluctant to conduct randomized clinical trials
202 because they are costly and time consuming. Lack of methodological and statistical know-
203 how can hinder the making of randomized clinical trials; it can be difficult to recruit enough
204 trial participants, etc. Typical misconceptions about the usefulness of results from
205 randomized clinical trials can also hinder that randomized trials are conducted. It is, e.g.,
206 often stated that trial populations are not representative of patients in the clinic [4,42,43].

207 Strict inclusion and exclusion criteria (e.g., the need of informed consent) are believed to put
208 together trial populations not representative of patients in the clinic. The ethically need of
209 informed consent can theoretically affect trial populations so they are different from the
210 everyday patients, but such fears are often overestimated [44,45]. Besides the need of
211 informed consent it is generally not necessary to use narrow criteria for selecting trial
212 participants, as this may impair the external validity [46]. We acknowledge all of these
213 difficulties regarding randomized clinical trials. Nevertheless, the establishment of academic
214 industry independent trial units with know-how about evidence-based medicine [47] can
215 lessen and solve some of the many problems conducting randomized clinical trials [48-53].
216 Furthermore, regional, national, international, and global research collaboration between trial
217 units and clinical sites (e.g., The European Clinical Research Infrastructures (ECRIN), The
218 UK Clinical Research Collaboration (UKCRC) Clinical Trials Units Network [54], and The
219 Nordic Trial Alliance (NTA)[55]) may reduce problems with recruitment of a sufficient number
220 of trial participants etc.[56,57]. Well-conducted multicentre clinical trials also offer better
221 external validity than well-conducted single centre trials. It must be recognized how much
222 health-care costs can be reduced if patient treatment becomes more effective through
223 evidence-based research. It has been calculated that investment in randomized clinical trials
224 usually gives a reasonable or high return on investment [58]. Politicians and decision makers
225 must be taught the key positions of the randomized clinical trial and of systematic reviews of
226 such trials in clinical intervention research.

227

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

231
232

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 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 [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.	Potential solutions: It may be 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 the experimental intervention group 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 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-up assessments of different kinds. It has been postulated that this might	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 strict treatment protocols in a randomized trial[1]. It is possible to randomize

specifically benefit trial participants (and hence the trial results) compared to patients in the clinic[4,60,61].	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.

233
234

235
236

TABLE 2

Different types of control groups in randomized clinical trials					
Experimental intervention versus no intervention		Experimental intervention versus placebo, impure placebo*, 'active' placebo (nocebo)**, or a sham intervention		Experimental intervention versus 'treatment as usual'***	
Methodological strengths (+) and limitations (-)		Methodological strengths (+) and limitations (-)		Methodological strengths (+) and limitations (-)	
+	-	+	-	+	-
The beneficial and harmful effects of the experimental intervention can be shown 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').
<p style="text-align: center;">Co-interventions</p> <p>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</p>					

237
238
239
240
241
242
243
244

* 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

**** 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.

245
246
247
248

249
250

3. DISCUSSION

251

252 We believe that clinical experience and observational studies cannot and should not validate
253 the effects of new interventions. Observational studies can sufficiently assess associations
254 between certain interventions and outcomes, but the randomized clinical trials are always
255 needed to avoid falsely negating (type I error) or falsely confirming (type II error) the null
256 hypothesis and to assess causality between interventions and outcomes, i.e., randomized
257 clinical trials are needed to sufficiently validate intervention effects. Typical issues hindering
258 the conduct of trials can be overcome (Table 1).

259

260 A report from the Patient-Centered Outcomes Research Institute was recently published for
261 public comment [68]. This report claims that the use of observational studies to make causal
262 inference is potentially much stronger than it has been in the past [68], and similar
263 arguments are often published in highly esteemed journals [3-7,69]. We believe that the
264 fundamental construct of the observational studies limits the reliability of the results from
265 observational studies [16,18]. To assess if an intervention causes more benefit than harm
266 randomized clinical trials are, in practical terms, always needed. Deeks and colleagues have
267 in a comprehensive report compared results from randomized trials and observational
268 studies [18]. They showed that results from observational studies can be seriously
269 misleading and that adjusted results in observational studies may even appear more
270 misleading than unadjusted results [18]. Compared to small randomized clinical trials, small
271 observational studies often showed effects that were far from the 'true' intervention effect
272 [18]. Ioannidis and colleagues also observed that significant discrepancies do occur between
273 the results of randomized clinical trials and observational studies [16]— and that results from
274 observational studies are more often contradicted than results from randomized clinical trials
275 [70]. Observational studies can be the only possible option regarding assessment of very

276 rare adverse events, very late occurring effects, or of very long-term
277 interventions. Observational studies can also have their place when it is difficult to include
278 large enough sample sizes assessing extremely rare diseases or when lack of funds hinders
279 the conduct of randomized clinical trials. Observational studies can of course have their
280 place in such circumstances but their inferential power should always be considered
281 threatened by random errors, confounding by indication, unmeasured confounding, and
282 other systematic errors. Therefore, the randomized clinical trial would still in such
283 circumstances be the optimal design regardless of hindrances making them infeasible. It
284 may, as mentioned, be possible to present a few historical examples where intervention
285 effects have been sufficiently validated by observational evidence[5]. However, these
286 exceptions do not justify that observational evidence generally should be used prospectively
287 to validate intervention effects. As it has been clearly expressed by Heiberg already in 1897
288 and reiterated by others both before and since [71-73]— regarding the vast majority of
289 interventions randomized clinical trials are necessary to assess their effects.

290

291 We acknowledge that randomized clinical trials may also get intervention effects wrong.
292 However, the likelihood of this occurring decreases with increasing sample sizes of the trials,
293 number of outcomes (reducing the risks of random errors), as well as with improved quality
294 of the methodology (reducing the risks of systematic errors)[1,31,32,36,74,75]. Moreover,
295 the conduct of systematic reviews assessing all randomized clinical trials on an intervention
296 as conducted by The Cochrane Collaboration reduces these risks [1,36,74,75]. We therefore
297 need to invest more in education in clinical research as well as in infrastructures for clinical
298 research and for systematic reviewing of randomized clinical trials.

299

300 Another group of arguments also exposes the weaknesses of observational studies. For
301 observational studies we do not yet have requirements of making public peer-reviewed
302 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 do 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 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 rational clinicians 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]. 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 disagree with authors claiming that observational designs can be employed for assessing interventions.

330

331 **ACKNOWLEDGEMENTS**

332

333 We thank JørnWetterslev for helpful discussions.

334

335 **COMPETING INTERESTS**

336

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

339

340 **AUTHORS' CONTRIBUTIONS**

341

342 Both authors conceived the study, contributed to the writing of the paper, and are
343 guarantors.

344

345 REFERENCES

- 346 1. Higgins JPT, Green S. The Cochrane Handbook for Systematic Reviews of
347 Interventions, Version 5.1.0. The Cochrane Collaboration. 2011;Available
348 from <http://www.cochrane-handbook.org>.
349
- 350 2. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, et al.
351 GRADE guidelines 11-making an overall rating of confidence in effect
352 estimates for a single outcome and for all outcomes. *J Clin Epidemiol*.
353 2013;66(2):151-157.
- 354 3. Ulvenes LV, Aasland O, Nylenna M, Kristiansen IS. Norwegian physicians'
355 knowledge of and opinions about evidence-based medicine: cross-sectional
356 study. *PloS one*. 2009;4(11):e7828.
- 357 4. Vincent JL. We should abandon randomized controlled trials in the intensive care
358 unit. *Crit Care Med*. 2010;38(10 Suppl):S534-538.
- 359 5. Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials
360 unnecessary? Picking signal from noise. *BMJ*. 2007;334(7589):349-351.
- 361 6. Stolberg HO, Norman G, Trop I. Randomized controlled trials. *Am J Roentgenol*.
362 2004;183(6):1539-1544.
- 363 7. Hawkins NG, Sanson-Fisher RW, Shakeshaft A, D'Este C, Green LW. The
364 multiple baseline design for evaluating population-based research. *Am J*
365 *Prev Med*. 2007;33(2):162-168.
- 366 8. Gluud C, Kubiak C, Whitfield K, Byrne J, Huemer KH, Thirstrup S, et al. Typical
367 investigational medicinal products follow relatively uniform regulations in 10
368 European Clinical Research Infrastructures Network (ECRIN) countries.
369 *Trials*. 2012;1327.
- 370 9. Kramer DB, Xu S, Kesselheim AS. How does medical device regulation perform
371 in the United States and the European union? A systematic review. *PLoS*
372 *Med*. 2012;9(7):e1001276.
- 373 10. Feltham C (1997) Book: Which Psychotherapy?: Leading Exponents Explain
374 Their Differences (page 1-224): SAGE Publications Ltd.
- 375 11. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, et
376 al. No surgical innovation without evaluation: the IDEAL recommendations.
377 *Lancet*. 2009;374(9695):1105-1112.
- 378 12. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet*.
379 2011;377(9778):1693-1702.
- 380 13. Chalmers TC. Randomize the first patient. *N Engl J Med*. 1977;296(2):107.
- 381 14. Chalmers I, Milne I, Trohler U, Vandenbroucke J, Morabia A, Tait G, et al. The
382 James Lind Library: explaining and illustrating the evolution of fair tests of
383 medical treatments. *The journal of the Royal College of Physicians of*
384 *Edinburgh*. 2008;38(3):259-264.
- 385 15. Vandenbroucke JP. What is the best evidence for determining harms of medical
386 treatment? *CMAJ*. 2006;174(5):645-646.
- 387 16. Ioannidis JP, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, et al.
388 Comparison of evidence of treatment effects in randomized and
389 nonrandomized studies. *JAMA*. 2001;286(7):821-830.
- 390 17. Papanikolaou PN, Christidi GD, Ioannidis JP. Comparison of evidence on harms
391 of medical interventions in randomized and nonrandomized studies. *Can*
392 *Med Assoc J*. 2006;174(5):635-641.

- 393 18. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, et al.
394 Evaluating non-randomised intervention studies. *Health Technol Assess*.
395 2003;7(27):iii-x, 1-173.
- 396 19. Chalmers I. Well informed uncertainties about the effects of treatments. *BMJ*.
397 2004;328(7438):475-476.
- 398 20. Cooper J. Randomized clinical trials for new surgical operations: square peg in a
399 round hole? *J Thorac Cardiovasc Surg*. 2010;140(4):743-746.
- 400 21. Huot L, Decullier E, Maes-Beny K, Chapuis FR. Medical device assessment:
401 scientific evidence examined by the French national agency for health - a
402 descriptive study. *BMC public health*. 2012;12585.
- 403 22. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to
404 gravitational challenge: systematic review of randomised controlled trials. *Int*
405 *J Prosthodont*. 2006;19(2):126-128.
- 406 23. Cuschieri A, Hunter J, Wolfe B, Swanstrom LL, Hutson W. Multicenter
407 prospective evaluation of laparoscopic antireflux surgery. Preliminary report.
408 *Surg Endosc*. 1993;7(6):505-510.
- 409 24. Kahrilas PJ, Lin S, Spiess AE, Brasseur JG, Joehl RJ, Manka M. Impact of
410 fundoplication on bolus transit across esophagogastric junction. *Am J*
411 *Physiol*. 1998;275(6 Pt 1):G1386-1393.
- 412 25. Wileman SM, McCann S, Grant AM, Krukowski ZH, Bruce J. Medical versus
413 surgical management for gastro-oesophageal reflux disease (GORD) in
414 adults. *Cochrane database of systematic reviews (Online)*.
415 2010;(3):CD003243.
- 416 26. Cates C, Cates M (2008) Regular treatment with salmeterol for chronic asthma:
417 serious adverse events. *Cochrane Database of Systematic Reviews*: John
418 Wiley & Sons, Ltd.
- 419 27. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions.
420 *Cochrane DatabaseSystRev*. 2010;(1):CD003974.
- 421 28. World Medical Association. Declaration of Helsinki. Ethical principles for medical
422 research involving human subjects. *J Indian Med Assoc*. 2009;107(6):403-
423 405.
- 424 29. Bertele V, Banzi R, Gluud C, Garattini S. EMA's reflection on placebo does not
425 reflect patients' interests. *Eur J Clin Pharmacol*. 2012;68(5):877-879.
- 426 30. Garattini S, Bertele V, Banzi R. Placebo? no thanks, it might be bad for me! *Eur*
427 *J Clin Pharmacol*. 2013;69(3):711-714.
- 428 31. Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of
429 reported study design characteristics on intervention effect estimates from
430 randomized, controlled trials. *Ann Intern Med*. 2012;157(6):429-438.
- 431 32. Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical
432 evidence of bias in treatment effect estimates in controlled trials with
433 different interventions and outcomes: meta-epidemiological study. *BMJ*.
434 2008;336(7644):601-605.
- 435 33. Faurschou A, Olesen AB, Leonardi-Bee J, Haedersdal M. Lasers or light
436 sources for treating port-wine stains. *Cochrane database of systematic*
437 *reviews (Online)*. 2011;11CD007152.
- 438 34. Boulaamane L, Boutayeb S, Errihani H. Bevacizumab based chemotherapy in
439 first line treatment of HER2 negative metastatic breast cancer: results of a

- 440 Moroccan observational institutional study. BMC research notes.
441 2012;5(1):162.
- 442 35. Montazeri A. Health-related quality of life in breast cancer patients: a
443 bibliographic review of the literature from 1974 to 2007. J Exp Clin Cancer
444 Res. 2008;2732.
- 445 36. Keus F, Wetterslev J, Gluud C, van Laarhoven CJ. Evidence at a glance: error
446 matrix approach for overiewing available evidence. BMC Med Res
447 Methodol. 2010;1090.
- 448 37. Magnusson BP, Turnbull BW. Group sequential enrichment design incorporating
449 subgroup selection. Stat Med. 2013;doi: 10.1002/sim.5738.
- 450 38. Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G.
451 Subgroup analyses in randomised controlled trials: quantifying the risks of
452 false-positives and false-negatives. Health Technol Assess. 2001;5(33):1-56.
- 453 39. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI.
454 How to use an overview. Evidence-Based Medicine Working Group. JAMA.
455 1994;272(17):1367-1371.
- 456 40. Jackson LS, Davis JW, Kaups KL, Sue LP, Wolfe MM, Bilello JF, et al.
457 Percutaneous tracheostomy: to bronch or not to bronch-that is the question.
458 J Trauma. 2011;71(6):1553-1556.
- 459 41. Stiell IG, Nichol G, Leroux BG, Rea TD, Ornato JP, Powell J, et al. Early versus
460 later rhythm analysis in patients with out-of-hospital cardiac arrest. N Engl J
461 Med. 2011;365(9):787-797.
- 462 42. Pincus T. Limitations of randomized clinical trials to recognize possible
463 advantages of combination therapies in rheumatic diseases. Semin Arthritis
464 Rheum. 1993;23(2 Suppl 1):2-10.
- 465 43. Cartwright N, Munro E. The limitations of randomized controlled trials in
466 predicting effectiveness. J Eval Clin Pract. 2010;16(2):260-266.
- 467 44. West J, Wright J, Tuffnell D, Jankowicz D, West R. Do clinical trials improve
468 quality of care? A comparison of clinical processes and outcomes in patients
469 in a clinical trial and similar patients outside a trial where both groups are
470 managed according to a strict protocol. Qual Saf Health Care.
471 2005;14(3):175-178.
- 472 45. Vist G, Bryant D, Somerville L, Birmingham T, Oxman A. Outcomes of patients
473 who participate in randomized controlled trials compared to similar patients
474 receiving similar interventions who do not participate. Cochrane Database of
475 Syst Rev. 2008;Jul 16;(3):MR000009.
- 476 46. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized
477 controlled trials published in high-impact general medical journals: a
478 systematic sampling review. JAMA. 2007;297(11):1233-1240.
- 479 47. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence
480 based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-72.
- 481 48. Fleischmann R. Primer: establishing a clinical trial unit--obtaining studies and
482 patients. Nat Clin Pract Rheumatol. 2007;3(8):459-463.
- 483 49. Fleischmann R. Primer: establishing a clinical trial unit - regulations and
484 infrastructure. Nat Clin Pract Rheumatol. 2007;3(4):234-239.
- 485 50. Seiler CM, Wentz MN, Diener MK, Frohlich BE, Buchler MW, Knaebel HP.
486 Center for clinical studies in a surgical department--an approach for more
487 evidence-based medicine. Contemp Clin Trials. 2006;27(3):211-214.

- 488 51. Ohmann C, Bruns I, Wolff S. The Coordinating Centers for Clinical Trials (KKS)
489 and the KKS Network: competence for clinical research. *Onkologie*. 2010;33
490 Suppl 711-15.
- 491 52. Gluud C, Sorensen TI. New developments in the conduct and management of
492 multi-center trials: an international review of clinical trial units. *Fundam Clin*
493 *Pharmacol*. 1995;9(3):284-289.
- 494 53. Gluud C, Sorensen TI. Organisation of a clinical trial unit--a proposal. *Fundam*
495 *Clin Pharmacol*. 1998;12(3):298-305.
- 496 54. The UK Clinical Research Collaboration (UKCRC). <http://www.ukcrc-ctu.org.uk/>.
- 497 55. Nordic Trail Alliance (NTA). [http://www.nordforsk.org/no/programs/nordic-trial-](http://www.nordforsk.org/no/programs/nordic-trial-alliance-nta)
498 [alliance-nta](http://www.nordforsk.org/no/programs/nordic-trial-alliance-nta).
- 499 56. Demotes-Mainard J, Ohmann C. European Clinical Research Infrastructures
500 Network: promoting harmonisation and quality in European clinical research.
501 *Lancet*. 2005;365(9454):107-108.
- 502 57. Demotes-Mainard J, Ohmann C, Gluud C, Chène G, Fabris N, Garattini S.
503 European clinical research infrastructures network meeting report: towards
504 an integration of clinical research infrastructures in Europe. *Int J Pharm Med*.
505 2005;19(1):43-45.
- 506 58. Health Economics Research Group, Office of Health Economics, RAND Europe.
507 Medical Research: What's it worth? Estimating the economic benefits from
508 medical research in the UK.
509 [http://www.wellcome.ac.uk/stellent/groups/corporatesite/@sitestudioobjects/document](http://www.wellcome.ac.uk/stellent/groups/corporatesite/@sitestudioobjects/document/s/web_document/wtx052110pdf)
510 [s/web_document/wtx052110pdf](http://www.wellcome.ac.uk/stellent/groups/corporatesite/@sitestudioobjects/document/s/web_document/wtx052110pdf).
- 511 59. Nichol AD, Bailey M, Cooper DJ, Polar, Investigators EPO. Challenging issues
512 in randomised controlled trials. *Injury*. 2010;41 Suppl 1S20-23.
- 513 60. Joffe S, Weeks JC. Views of American oncologists about the purposes of clinical
514 trials. *J Natl Cancer Inst*. 2002;94(24):1847-1853.
- 515 61. Davis S, Wright PW, Schulman SF, Hill LD, Pinkham RD, Johnson LP, et al.
516 Participants in prospective, randomized clinical trials for resected non-small
517 cell lung cancer have improved survival compared with nonparticipants in
518 such trials. *Cancer*. 1985;56(7):1710-1718.
- 519 62. Gross CP, Krumholz HM, Van Wye G, Emanuel EJ, Wendler D. Does Random
520 Treatment Assignment Cause Harm to Research Participants? *PLoS Med*.
521 2006;3(6):e188.
- 522 63. Appelbaum P, Roth L, Lidz C, Benson P, Winslade W. False hopes and best
523 data: consent to research and the therapeutic misconception. *The Hastings*
524 *Center Report*. 1987;(17(2):20-4).
- 525 64. Middleton S, McElduff P, Ward J, Grimshaw JM, Dale S, D'Este C, et al.
526 Implementation of evidence-based treatment protocols to manage fever,
527 hyperglycaemia, and swallowing dysfunction in acute stroke (QASC): a
528 cluster randomised controlled trial. *Lancet*. 2011;378(9804):1699-1706.
- 529 65. Becker A, Held H, Redaelli M, Chenot JF, Leonhardt C, Keller S, et al.
530 Implementation of a Guideline for Low Back Pain Management in Primary
531 Care - A Cost-Effectiveness Analysis. *Spine (Phila Pa 1976)*.
532 2011;37(8):701-710.
- 533 66. Shorr AF, Micek ST, Jackson WL, Jr., Kollef MH. Economic implications of an
534 evidence-based sepsis protocol: can we improve outcomes and lower costs?
535 *Crit Care Med*. 2007;35(5):1257-1262.

- 536 67. Karjalainen S, Palva I. Do treatment protocols improve end results? A study of
537 survival of patients with multiple myeloma in Finland. BMJ.
538 1989;299(6707):1069-1072.
- 539 68. Helfand M, Berg A, Flum D, Gabriel S, Normand S. Draft Methodology Report:
540 "Our Questions, Our Decisions: Standards for Patient-Centered Outcomes
541 Research" (Page 1-186). Patient-Centeret Outcomes Research Institute.
542 2012;Published for Public Comment July 23 (Available at :
543 <http://www.pcori.org/assets/Preliminary-Draft-Methodology-Report.pdf>).
- 544 69. Higgins J. Convincing evidence from controlled and uncontrolled studies on the
545 lipid-lowering effect of a statin [editorial]. Cochrane Database of Systematic
546 Reviews.
547 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)
548 [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)
549 [loweri.html](http://www.thecochranelibrary.com/details/editorial/3992221/Convincing-evidence-from-controlled-and-uncontrolled-studies-on-the-lipid-loweri.html).
- 550 70. Ioannidis J. Contradicted and initially stronger effects in highly cited clinical
551 research. JAMA. 2005;294(2):218-228.
- 552 71. Heiberg P. Studier over den statistiske undersøgelsesmetode som hjælpemiddel
553 ved terapeutiske undersøgelser [Studies on the statistical study design as an
554 aid in therapeutic trials]
555 ([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)
556 [undersogelsesmetode-som-hjaelpemid/key_passages](http://www.jameslindlibrary.org/illustrating/records/studier-over-den-statistiske-undersogelsesmetode-som-hjaelpemid/key_passages)). Bibl Laeger. 1897;891-
557 40.
- 558 72. Korn EL, McShane LM, Freidlin B. Statistical challenges in the evaluation of
559 treatments for small patient populations. Science translational medicine.
560 2013;5(178):178sr173.
- 561 73. The Library and Information Services Department, The Royal College of
562 Physicians of Edinburgh. James Lind Library. Available online at:
563 <http://www.jameslindlibrary.org/>. 2003.
- 564 74. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and
565 research outcome. Cochrane database of systematic reviews (Online).
566 2012;12MR000033.
- 567 75. Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, et al. The
568 number of patients and events required to limit the risk of overestimation of.
569 PloS one. 2011;6e25491.
- 570 76. +AllTrials (avaible at: <http://www.alltrials.net>). All Trials Registered, All Results
571 Reported.
- 572 77. Gugiu PC, Gugiu MR. A critical appraisal of standard guidelines for grading
573 levels of evidence. Eval Health Prof. 2010;33(3):233-255.
574
575
576