Types of randomized clinical trials published in PubMed over 40 years

MARIANA MARTÍNEZ-FRANCO, ANTONIO ROBERTO NIRTA-PÉREZ, JORGE HERNANDO DONADO-GÓMEZ. MEDELLÍN (COLOMBIA)

DOI: https://doi.org/10.36104/amc.2021.1884

Abstract

Introduction: there are several variations of randomized clinical trials. Trials can be classified by design as parallel, cross-over, factorial, assignment by group, n-of-1, paired, withdrawal, adaptive and pragmatic; and by purpose as superiority, non-inferiority and equivalence. Given this heterogeneity, the Consolidated Standards of Reporting Trials (CONSORT) were introduced in 1996 to provide a guideline for reporting randomized clinical trials.

Objective: to describe the publication tendency of the various types of randomized clinical trials over 40 years, with reference to the publication of CONSORT and its extensions.

Methods: the PubMed tool was used to search for randomized clinical trials published between 1979 and 2018, classifying them according to the varieties described.

Results: a total of 472,114 published articles were found; 90.2% did not report the type of design and 98.2% did not report the purpose. Among the articles that reported the variety of randomized clinical trial, the predominant design was cross-over (5.9%), followed by parallel groups (2.34%); while the most common purpose was superiority (0.84%). After the launch of CONSORT, there was an increased proportion of articles published with the following designs: parallel groups; difference in proportions 1.89 95% CI (1.1-2.7); paired 1.07 95% CI (0.2-1.9); and pragmatic 4.73 95% CI (4.4-5.1); and for the purpose of non-inferiority 5.97 95% CI (5.6-6.3).

Discussion: most articles on randomized clinical trials do not mention their type in the title and abstract. The proportion of articles that did, increased slightly after CONSORT was published. (Acta Med Colomb 2021; 46. DOI: https://doi.org/10.36104/amc.2021.1884).

Key words: clinical trial, epidemiological study design, publications, epidemiological study characteristics, assessment of therapeutic intervention results.

Introduction

Clinical trials are prospective studies in humans comparing the effects of an intervention with a control (1).

They date back to the eighteenth century, when Lind described the characteristics of what would be the first documented clinical trial, carried out with a small sample of scurvy patients (2).

Today, randomized controlled trials are considered to contribute the greatest level of evidence for healthcare interventions. Due to their growing popularity, multiple variants of the original design have emerged, now known as “parallel group” studies, which consist in selecting a sample of patients and randomly assigning them to one of two or more groups (3).

The other variants include:

• Cluster: used when individual assignment is not appropriate due to the possibility of intervention “contamination” among the participants. In this case, assignments are made by groups or “clusters,” which may be geographic zones, institutions or groups of patients seen by a physician (4, 5).

• Cross over: in this case, each participant acts as his/her own control, receiving the intervention and the comparator during different periods, with a washout time between each (4, 5, 19).

• N-of-1: a variant used when the available evidence is not useful for a particular patient. Therefore, a crossover study is performed with this patient in which he/she acts as his/her own control and receives the study intervention and comparator alternately and randomly, in several cycles. The entire population consists of a single patient and the study results cannot be extrapolated to other patients (6).

• Factorial: allows two or more interventions to be evaluated in a single study. The interventions are randomly assigned such that the participant may receive no intervention, only one, or perhaps all of them (4, 5, 19).
• **Paired:** in this design, the randomized unit is not an individual but rather an organ or body site, such that each participant may receive two or more treatments in different parts of the body. Other terms used to describe it are “within person”, “split body”, “split face”, “split mouth” or “contralateral” (7).

• **Withdrawal:** the aim of this variant is to evaluate the response to discontinuation or reduction of the dose of an intervention which has already been tolerated and has proven beneficial (4, 5).

• **Adaptive:** this is one of the most recent variants of randomized controlled trials, which consists in allowing variability in the participants, the size of the sample, the intervention and the outcomes according to the study’s preliminary results. This permissiveness must be stipulated in the research protocol (8).

• **Pragmatic:** these designs seek to maximize the generalizability of study results, and therefore they are carried out in “usual practice” settings, with a broad range of participants and taking into account clinically significant outcomes like mortality and morbidity. Their counterparts are explanatory trials, which are a more widely used and have much more controlled conditions in an effort to prove a causal hypothesis, decreasing the potential implementation capacity of this hypothesis (9).

In addition to the design variants, new purposes for randomized controlled trials have emerged over time. The classical purpose is *superiority*, which seeks to determine if an intervention is more effective than the reference. However, studies may also evaluate whether an intervention is not inferior, or whether both interventions are the same. These purposes correspond to *noninferiority* and *equivalence studies*, respectively (3, 19).

Due to the importance of randomized controlled trials today and the growing research in methodology, a group of experts has been developing a tool since 1996 to improve the quality of reporting and decrease its associated biases: the Consolidated Standards of Reporting Trials (CONSORT). The third version, the CONSORT 2010 statement, is currently in force: updated guidelines for reporting parallel group randomized controlled trials (10). Its extensions have also been published, with guidelines for publishing cluster designs available since 2004 (11, 12), noninferiority and equivalence since 2006 (13, 14), pragmatic since 2009 (9), N-of-1 since 2015 (15), and within person since 2017 (7).

This article seeks to describe the publication tendencies of the different randomized controlled trial variants over 40 years, from 1979 to 2018, and their relationship with the launching of CONSORT guidelines, specifically with regard to the guidelines for including the randomized controlled trial subtype in the article’s title and abstract.

**Materials and methods**

This was an observational study using the PubMed tool to search for randomized controlled trials published between 1979 and 2018.

First, the number of randomized controlled trials published during this period was identified, and then a search strategy was implemented for each of the following designs: parallel group studies, cluster studies, cross over studies, N-of-1 studies, factorial studies, paired studies, withdrawal studies, adaptive studies and pragmatic studies.

A bibliographic search was also performed based on the purpose of the clinical trials, considering the three existing possibilities: superiority, noninferiority and equivalence studies.

The search terms were selected based on the following criteria:

• For designs with an existent CONSORT extension, the CONSORT recommended term was used.

• For designs not yet included in CONSORT or for which no particular term is recommended for reporting, the study name was combined with the terms “study”, “design” and “trial”.

The search strategies employed may be found in the supplementary material (Annex 1).

The search results were reported separately by two of the authors (MMF and ARNP) and their concordance was assessed using the kappa coefficient. The data are presented as absolute and relative frequencies. A comparison of the data before and after the respective CONSORT publication was performed using the difference in proportions test. A p < 0.05 was considered to be statistically significant. The Epidat version 4.2 statistical package was used (22).

This study did not require ethics committee approval as the unit of analysis was published clinical trials.

**Results**

According to the bibliographic search performed between 1979 and 2018, a total of 472,114 articles labeled as “randomized controlled trials” were published, and 94.58% were published in English, followed by 1.7% in Chinese and 1.25% in German. The remaining languages like French, Spanish, Portuguese, Japanese, Italian and Russian accounted for less than 1%. A total of 90.2% (426,071) of these articles did not mention the design variant used in their titles or abstracts. The remaining percentage was distributed as follows: 28,246 (5.9%) were of the cross over variant, 11,064 (2.34%) were parallel groups, 2,056 (0.4%) were clusters, 1,796 (0.38%) were factorial, 1,548 (0.32%) were pragmatic, 847 (0.17%) were paired, 246 (0.05%) were adaptive, 141 (0.02%) were withdrawal and 99 (0.02%) were N-of-1 (Figure 1).

In the search by purpose, 98.2% (463,999) did not specify the purpose. In the remainder, there were 3,985 (0.84%) articles on superiority, 1,850 (0.39%) on noninferiority and 1,716 (0.36%) on equivalence randomized controlled trials (Figure 2). The distribution by five-year periods of the types of randomized controlled trials according to design and purpose may be found in Tables 1 and 2. Figures 3 and 4 show the publication tendency of the dif-
different clinical trial variants throughout the years, separated by five-year periods.

The proportion of published articles identified by the randomized controlled trial variant in their title or abstract following the launching of CONSORT and its respective extensions increased for parallel group, N-of-1, paired, and pragmatic designs, and for the purpose of noninferiority, but decreased for equivalence. These were statistically significant changes in all cases except for N-of-1 and equivalence (Table 3).

The evaluation of overall concordance in data gathering between researchers (MMF and ARNP) found a kappa coefficient (κ) of 0.88 95% CI (0.82-0.95), which indicates a very good performance according to Cohen’s criteria (21, 22).

Discussion

Throughout the 40 years of publications reviewed in PubMed, we found that more than 90% of the articles did not report the type of clinical trial presented, nor its purpose, in the title or abstract. However, we found a slight, but statistically significant, increase in the rate of reporting these characteristics in the years following the launching of CONSORT and its extensions. This indicates that the guidelines for publishing clinical trials dictated by CONSORT are beginning to be adopted by the scientific community, although, as has been shown in previous studies, an adequate adherence has not been achieved (16-18, 20).

On the other hand, a surprising finding is that most of the articles that reported their classification were cross over studies. The predominant design was expected to be parallel...
Table 1. Frequencies of types of randomized controlled trials published, by five-year period (by design).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>11,739</td>
<td>19,269</td>
<td>35,431</td>
<td>51,190</td>
<td>60,091</td>
<td>79,614</td>
<td>104,162</td>
<td>110,054</td>
</tr>
<tr>
<td>Parallel groups</td>
<td>47 (0.4)* (6.4)**</td>
<td>158 (0.8)* (9.9)**</td>
<td>595 (1.7)* (19.1)**</td>
<td>1,218 (2.4)* (26.3)**</td>
<td>1,670 (2.8)* (27.2)**</td>
<td>1,820 (2.3)* (23.6)**</td>
<td>2,350 (2.3)* (24.1)**</td>
<td>3,206 (2.9)* (26)**</td>
</tr>
<tr>
<td>Cluster</td>
<td>0 (0)* (0)**</td>
<td>0 (0)* (0)**</td>
<td>1 (0)* (0)**</td>
<td>74 (0.1)* (1.2)**</td>
<td>226 (0.3)* (2.9)**</td>
<td>608 (0.6)* (6.2)**</td>
<td>1,146 (1)* (9.3)**</td>
<td></td>
</tr>
<tr>
<td>Cross over</td>
<td>660 (5.6)* (90.3)**</td>
<td>1,386 (7.2)* (86.7)**</td>
<td>2,405 (6.8)* (77.1)**</td>
<td>3,145 (6.1)* (67.9)**</td>
<td>3,999 (6.7)* (65.1)**</td>
<td>4,986 (6.3)* (64.7)**</td>
<td>5,290 (5.4)* (57.2)**</td>
<td>6,075 (5.5)* (49.2)**</td>
</tr>
<tr>
<td>N-of-1</td>
<td>0 (0)* (0)**</td>
<td>2 (0)* (0.1)**</td>
<td>9 (0)* (0.3)**</td>
<td>17 (0)* (0.4)**</td>
<td>10 (0)* (0.2)**</td>
<td>23 (0)* (0.3)**</td>
<td>13 (0)* (0.1)**</td>
<td>25 (0)* (0.2)**</td>
</tr>
<tr>
<td>Factorial</td>
<td>17 (0.1)* (2.3)**</td>
<td>33 (0.2)* (2.1)**</td>
<td>85 (0.2)* (2.7)**</td>
<td>164 (0.3)* (3.5)**</td>
<td>219 (0.4)* (3.6)**</td>
<td>301 (0.4)* (3.9)**</td>
<td>464 (0.6)* (4.8)**</td>
<td>513 (0.5)* (4.2)**</td>
</tr>
<tr>
<td>Paired</td>
<td>3 (0)* (0.4)**</td>
<td>8 (0)* (0.5)**</td>
<td>9 (0)* (0.3)**</td>
<td>46 (0.1)* (1)**</td>
<td>85 (0.1)* (1.4)**</td>
<td>146 (0.2)* (1.9)**</td>
<td>251 (0.2)* (2.6)**</td>
<td>299 (0.3)* (2.4)**</td>
</tr>
<tr>
<td>Withdrawal</td>
<td>4 (0)* (0.5)**</td>
<td>7 (0)* (0.4)**</td>
<td>7 (0)* (0.3)**</td>
<td>11 (0)* (0.2)**</td>
<td>9 (0)* (0.1)**</td>
<td>21 (0)* (0.3)**</td>
<td>33 (0)* (0.3)**</td>
<td>49 (0)* (0.4)**</td>
</tr>
<tr>
<td>Adaptive</td>
<td>0 (0)* (0)**</td>
<td>2 (0)* (0.1)**</td>
<td>3 (0)* (0.1)**</td>
<td>2 (0)* (0)**</td>
<td>6 (0)* (0.1)**</td>
<td>21 (0)* (0.3)**</td>
<td>64 (0.1)* (0.7)**</td>
<td>148 (0.1)* (1.2)**</td>
</tr>
<tr>
<td>Pragmatic</td>
<td>0 (0)* (0)**</td>
<td>2 (0)* (0.1)**</td>
<td>6 (0)* (0.2)**</td>
<td>29 (0)* (0.6)**</td>
<td>72 (0.1)* (1.2)**</td>
<td>159 (0.2)* (2.1)**</td>
<td>395 (0.4)* (4)**</td>
<td>885 (0.8)* (7.2)**</td>
</tr>
<tr>
<td>No classification</td>
<td>11,008 (93.8)</td>
<td>17,671 (91.7)</td>
<td>32,311 (91.2)</td>
<td>46,557 (91)</td>
<td>53,947 (90.6)</td>
<td>71,911 (90.3)</td>
<td>94,394 (90.6)</td>
<td>97,708 (88.8)</td>
</tr>
</tbody>
</table>

*Percentage of total.  
**Percentage of effective articles (not including those with "no classification").

groups, since this was the initial model, although the cross over design is known to be used more in basic sciences than in clinical sciences (3). However, one possibility is that most of the unclassified articles are parallel group designs, which, being the baseline model for the other designs, is assumed to be the design when an alternative is not specified.

As expected, the publication tendency of all randomized controlled trials increased progressively over the years. A greater variety in these studies was introduced around 1990, with a growing report of cluster, N-of-1, paired, adaptive, withdrawal, pragmatic and factorial studies. Prior to this date, only cross over and parallel group variants were reported as such.

Superiority clinical trials continue to increase over time and account for most of the articles that report their purpose. The articles on noninferiority randomized controlled trials appeared around the year 2000, and they are also increasingly reported, while equivalence studies have tended to stabilize since then.

Study weaknesses include the fact that the complete text of the articles was not reviewed, and only one bibliographic database was consulted. However, PubMed is the most important bibliographic database worldwide. In addition, duplicates were not controlled for, and thus the actual number of randomized controlled trials is probably overestimated. However, with the existing tools for identifying

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>11739</td>
<td>19269</td>
<td>35431</td>
<td>51190</td>
<td>60091</td>
<td>79614</td>
<td>104162</td>
<td>110054</td>
</tr>
<tr>
<td>Superiority</td>
<td>96 (0.8)*</td>
<td>121 (0.6)*</td>
<td>170 (0.5)*</td>
<td>223 (0.4)*</td>
<td>345 (0.6)*</td>
<td>536 (0.7)*</td>
<td>958 (0.9)*</td>
<td>1536 (1.4)*</td>
</tr>
<tr>
<td>Noninferior</td>
<td>0 (0)*</td>
<td>0 (0)*</td>
<td>0 (0)*</td>
<td>0 (0)*</td>
<td>33 (0.1)*</td>
<td>270 (0.3)*</td>
<td>659 (0.6)*</td>
<td>888 (0.8)*</td>
</tr>
<tr>
<td>Equivalence</td>
<td>6 (0.5)*</td>
<td>21 (0.1)*</td>
<td>65 (0.2)*</td>
<td>182 (0.4)*</td>
<td>338 (0.6)*</td>
<td>342 (0.4)*</td>
<td>363 (0.3)*</td>
<td>399 (0.4)*</td>
</tr>
<tr>
<td>Unclassified</td>
<td>11637 (99.1)</td>
<td>19127 (99.3)</td>
<td>35196 (99.3)</td>
<td>50785 (99.2)</td>
<td>59375 (98.8)</td>
<td>78466 (98.5)</td>
<td>102182 (98)</td>
<td>107231 (97.4)</td>
</tr>
</tbody>
</table>

*Percentage of total.
**Percentage of effective articles (not including “unclassified” articles).

Table 3. Difference in proportions in the pattern of publication before and after CONSORT.

<table>
<thead>
<tr>
<th>Type of study</th>
<th>Proportion* (%)</th>
<th>Difference in proportions (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-CONSORT</td>
<td>4.061/17.759 (22.86)</td>
<td>1.89 (1.1 - 2.7)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Post-CONSORT</td>
<td>7.003/28.284 (24.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-of-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-CONSORT</td>
<td>84/38.939 (0.21)</td>
<td>0.06 (-0.00 - 0.002)</td>
<td>0.27</td>
</tr>
<tr>
<td>Post-CONSORT</td>
<td>15/9.773 (0.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-CONSORT</td>
<td>800/44.410 (1.80)</td>
<td>1.07 (0.2 - 1.9)</td>
<td>0.002</td>
</tr>
<tr>
<td>Post-CONSORT</td>
<td>47/1.633 (2.87)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-CONSORT</td>
<td>323/25.599 (1.26)</td>
<td>4.73 (4.4 - 5.1)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Post-CONSORT</td>
<td>1,225/20.444 (5.99)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noninferiority</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-CONSORT</td>
<td>153/20.754 (0.73)</td>
<td>5.97 (5.6 - 6.3)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Post-CONSORT</td>
<td>1,697/25.289 (6.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equivalence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-CONSORT</td>
<td>802/20.754 (3.86)</td>
<td>-0.25 (-0.6 - 0.1)</td>
<td>0.16</td>
</tr>
<tr>
<td>Post-CONSORT</td>
<td>914/25.289 (3.61)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Number of articles by type of design / number of effective articles published.

duplicate publications, this occurs less and less frequently, and it does not justify the lack of adherence to CONSORT found. Publications without an English abstract were not considered either, since they could not be analyzed by the researchers. However, these currently represent a minority of the studies. As a strength, few publications like this one were found. One of the most important of these is the 2012 Cochrane review (16) which evaluated the effect of adherence to CONSORT on the completeness of clinical trial reports, and whose conclusion was that the quality of reporting did improve in journals which adopted CONSORT. The review by Susvirkar et al. in 2018 (18) evaluated the adherence of clinical trial articles published in two high-impact journals (the Journal of the American Medical Association and the British Medical Journal) to the CONSORT guidelines, which, although close to 80%, showed poor adherence to identification of the article as a clinical trial in the title and a description of its design, among others.

No articles were found evaluating the publication of clinical trials by type of design, and therefore this could be a starting point for new studies on the topic, as well as serve to incentivize adherence to the CONSORT recommendations, improving the quality of scientific production.

References


22. Epidat: programa para análisis epidemiológico de datos. Versión 4.2, julio 2016. Consellería de Sanidade, Xunta de Galicia, España; Organización Panamericana de la salud (OPS-OMS); Universidad CES, Colombia.

**Supplementary material**

- **Randomized controlled trial**: randomized controlled trial, filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Parallel groups**: parallel group[Title/Abstract], filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Cluster**: cluster randomised[Title/Abstract] filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Cross over**: filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **N-of-1**: n of 1[Title/Abstract], filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Factorial**: factorial trial[Title/Abstract] OR factorial study[Title/Abstract] OR factorial design[Title/Abstract] OR factorial randomised[Title/Abstract], filter: Randomized Controlled Trial, Date of publication from 1979/01/01 to 2018/12/31.
- **Withdrawal**: withdrawal study[Title/Abstract] OR withdrawal trial[Title/Abstract] OR withdrawal design[Title/Abstract] filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Adaptive**: adaptive randomised[Title/Abstract] OR adaptive trial[Title/Abstract] OR adaptive study[Title/Abstract] OR adaptive design[Title/Abstract]; filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Pragmatic**: pragmatic trial[Title/Abstract] OR pragmatic randomised[Title/Abstract] OR pragmatic study[Title/Abstract] OR pragmatic design[Title/Abstract]; filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Superiority**: superiority[Title/Abstract]; filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Noninferiority**: noninferiority[Title/Abstract]; filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.
- **Equivalence**: equivalence[Title/Abstract]; filter: Randomized Controlled Trial, Publication date from 1979/01/01 to 2018/12/31.