Randomized clinical trials with run-in periods: frequency, characteristics and reporting

David Ruben Teindl Laursen[1–4]
Asger Sand Paludan-Müller[2]
Asbjørn Hróbjartsson[1,3,4]

[1]Centre for Evidence-Based Medicine Odense (CEBMO), Odense University Hospital, Odense, Denmark; [2]Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark; [3]Department of Clinical Research, University of Southern Denmark, Odense, Denmark; [4]Odense Patient data Explorative Network (OPEN), Odense University Hospital, Odense, Denmark

Background: Run-in periods are occasionally used in randomized clinical trials to exclude patients after inclusion, but before randomization. In theory, run-in periods increase the probability of detecting a potential treatment effect, at the cost of possibly affecting external and internal validity. Adequate reporting of exclusions during the run-in period is a prerequisite for judging the risk of compromised validity. Our study aims were to assess the proportion of randomized clinical trials with run-in periods, to characterize such trials and the types of run-in periods and to assess their reporting.

Materials and methods: This was an observational study of 470 PubMed-indexed randomized controlled trial publications from 2014. We compared trials with and without run-in periods, described the types of run-in periods and evaluated the completeness of their reporting by noting whether publications stated the number of excluded patients, reasons for exclusion and baseline characteristics of the excluded patients.

Results: Twenty-five trials reported a run-in period (5%). These were larger than other trials (median number of randomized patients 217 vs 90, *P*=0.01) and more commonly industry trials (11% vs 3%, *P*<0.01). The run-in procedures varied in design and purpose. In 23 out of 25 trials (88%), the run-in period was incompletely reported, mostly due to missing baseline characteristics.

Conclusion: Approximately 1 in 20 trials used run-in periods, though much more frequently in industry trials. Reporting of the run-in period was often incomplete, precluding a meaningful assessment of the impact of the run-in period on the validity of trial results. We suggest that current trials with run-in periods are interpreted with caution and that updates of reporting guidelines for randomized trials address the issue.

Keywords: run-in periods, lead-in periods, enrichment design, single-blind placebo, washout periods, research methodology

Background

Randomized clinical trials are generally considered the most reliable method for evaluating the effects of health care interventions. However, randomized trials vary in design, and different design characteristics may impact on internal validity (risk of bias), external validity (generalizability) and costs (both economical and logistical). One such design characteristic is a run-in period, which is a planned time period from formal patient enrollment to randomization that enables exclusion of certain patients, for example, if they experience harms (see Box 1).[1]

In trials with run-in periods, randomization may take place days or weeks after formal enrollment. During this post-inclusion pre-randomization period, all patients receive the same treatment, for example, a placebo ("placebo run-in"), the experimental

Correspondence: David Ruben Teindl Laursen
Centre for Evidence-Based Medicine Odense (CEBMO), Odense University Hospital, Kløvervænget 10, 13th Floor, Gate 112, DK-5000 Odense C, Denmark
Email David.Ruben.Teindl.Laursen@rsyd.dk

169

Box 1 Run-in periods in randomized clinical trials

- What is a run-in period?
  - A run-in period is a time period after inclusion, but before randomization, used to exclude certain patients. Other pre-randomization periods exist, for example, extended screening periods and washout periods. These different pre-randomization periods may overlap in purpose, design and terminology
- What types of run-in periods exist and which patients are excluded?
  - During the run-in period, all patients receive the same intervention, for example, active treatment, placebo treatment or no intervention. Patients are excluded due to, for example, noncompliance to treatment or data collection, non-response to treatment or response to placebo
- What are the reasons for using a run-in period?
  - By excluding certain patients, for example, noncompliers or placebo responders, run-in period may increase a study's power, that is, chance of detecting a potential treatment effect
- What other terms are used for a run-in period?
  - Similar terms used are lead-in periods, single-blind placebo periods and enrichment periods
- What potential problem does run-in period cause?
  - The use of a run-in period may affect external validity, by exclusion of patients from the clinical study population, as well as internal validity, by the risk of unblinding or exaggeration of the intention-to-treat effect estimate
- What is needed to assess the possible impact of a run-in period on trial results?
  - We propose that the study reader would want to study the number of excluded patients, reasons for exclusion and their baseline characteristics. These aspects should, therefore, be reported in the trial publications

drug ("active run-in") or observation only ("no-treatment" or "washout" run-in).[1,2]

The main rationale for a run-in period in a trial is to adjust the selection of patients for the post-randomization phase of the trial. The principal difference between standard screening by eligibility criteria in a trial and the procedures in a run-in phase is that the latter permits exclusions based on observations of patients' compliance or responses to trial interventions. Thus, for example, an active run-in period enables exclusion of patients who respond poorly to the experimental intervention. A placebo run-in period enables exclusion of patients who respond well to the placebo intervention. Both active and placebo run-in periods enable exclusion of patients who do not comply with trial procedures.[1,3]

The number of exclusions in run-in periods may be considerable. For example, in a trial of the effect of extended-

release opioid for lower back pain, 191 out of 459 enrolled patients (42%) were excluded during the active run-in period.[4] In another trial of the effect of aspirin and β-carotene on cardiovascular disease and cancer, 11,152 out of 33,223 patients (34%) were excluded during the run-in period.[5]

Run-in periods can affect the validity of a study's results. When patients are excluded during a run-in period, for example, due to harms or lack of response, the trial population may increasingly differ from the typical clinical patient population. The balance between benefits and harms of the trial intervention may appear more beneficial than if exclusions had not taken place.

Thus, interpretation of results from a trial with a run-in period is challenging as it involves careful consideration of how pre-randomization exclusions of patients could have affected trial validity. Such considerations presuppose access to relevant information on patient exclusions, typically through adequate reporting of the run-in period in the trial publication.

We, therefore, thought it relevant to study run-in periods in randomized clinical trials. Our objectives were to 1) assess the proportion of trial publications that report a run-in period; 2) characterize such trials and the types of run-in periods and 3) evaluate the completeness of reporting.

Materials and methods

This was an observational study of a random sample of PubMed-indexed trial publications.

Identification of trial publications

We identified publications indexed in PubMed as "randomized controlled trial" and published in 2014, and listed them in a random order using random.org.[6] One reference at a time, one author (DL) screened titles and abstracts (and full text if needed) to check if the publication reported a randomized clinical trial. If so, the same author read the full text of the trial publication and determined whether the trial used a run-in period according to our definition below. We continued the process until 25 trials with a run-in period were identified.

Randomized clinical trials were included if they had a parallel, crossover or split-body design (but excluded if they used cluster randomization). We considered trials to be "clinical" if they assessed the benefits or harms of a health care intervention. We planned to include trial publications written in all languages, using Google Translate as an aid.[7]

We operationally defined a "run-in period" as fulfilling either a main or a secondary criterion. The main criterion was that a trial publication had to 1) use one of the following terms: "run-in", "lead-in", "enrichment", "single-blind

placebo" or "baseline", indicating a time period after (explicitly reported) registration of trial patients, but before their randomization and 2) explicitly report possible exclusion of patients during this period due to non-response to experimental intervention, harms, response to control intervention, noncompliance to experimental or control intervention or noncompliance to data collection. The secondary criterion was that a trial publication had to 1) explicitly use the following terms: "run-in", "lead-in" or "enrichment", indicating a time period after (explicitly or implicitly reported) registration of trial patients, but before their randomization and 2) report no indications that investigator-driven patient exclusions due to response to, or compliance with, treatments were disallowed. The distinction between trials qualified as run-in trials according to the main and secondary criteria was used in sensitivity analyses (see Supplementary materials).

If the trial publication referred to previous trial publications (eg, published protocols), we included information from these to determine whether a run-in period was reported or not. Thus, "trial publication" in this study means the index publication identified in our sample and previous journal publications on the same trial cited in the index publication.

Data extraction and processing

From the sample of 25 randomized clinical trials with a run-in period as well as a random selection of 100 trials without a run-in period, one author (DL) extracted descriptive data on publication and trial: language, clinical specialty (medical, surgery, others), trial design (eg, parallel, crossover), number of intervention arms, types of control interventions (eg, active or placebo), treatment class (pharmacological or non-pharmacological), number of patients randomized and industry involvement.

We operationally defined an "industry trial" as a trial in which a commercial company (eg, a drug or a device company) had participated in the trial design, and thus had potentially influenced the decision to use a run-in period. A trial was categorized as an industry trial if a drug or device company was listed as "responsible party" or "sponsor" in ClinicalTrials.gov (or in other trial registries if the trial publication mentioned these), if industry employees had participated in the design of the trial (according to, eg, the trial report), if industry authors were mentioned as co-authors or if the trial was mentioned as industry funded, but did not describe who had designed the study.

From the trial publications reporting a run-in period, the same author (DL) extracted the following additional data: intervention type during run-in period (eg, active, placebo, no intervention), duration of the run-in period, number of

patients enrolled and excluded in the run-in period, reasons for exclusion and characteristics at inclusion (the "baseline characteristics"). Furthermore, we noted the term used for run-in period (eg, "run-in" or "lead-in period") and purposes of the run-in period (eg, selecting patients compliant to treatment).

For the publications describing a trial with a run-in period, one author (DL) evaluated the completeness of reporting. We defined complete reporting of a run-in period as the unambiguous description of 1) the number of patients enrolled to the run-in period and the number of patients excluded during the run-in period (and, consequently, the number of patients randomized after the run-in period); 2) the reasons for exclusion of all patients during the run-in period and 3) the characteristics of excluded patients. For each paper, we also noted specifically which of our three criteria were met or not.

A second author (AP) independently repeated the data extraction, assessment of industry involvement and evaluation of completeness of reporting of run-in period. This was done for all 25 run-in trial publications and for 20 randomly selected trial publications not reporting a run-in period. Disagreements were settled by discussion.

Data analysis

We tabulated descriptive data as numbers (and percentages) or medians (and interquartile ranges [IQRs]). For trials with run-in periods, we summarized trial data, run-in period data and completeness of reporting. We also planned to compare patient characteristics for excluded and randomized patients, and we summarized reasons for using a run-in period and the terminology involved. We used Fisher's exact test or Mann–Whitney U test to compare characteristics of trials with and without run-in phases. The software used for data analysis was Microsoft Excel and the Real Statistics Resource Pack.[8]

We performed sensitivity analyses to study the robustness of our results to our definition of a "run-in period" and to our definition of "industry trials" (see Supplementary materials).

Results

Prevalence of trial publications reporting run-in periods

We screened 748 PubMed items and identified 470 randomized clinical trials in order to obtain 25 publications (5%) reporting a run-in period (Figure 1).[9–33]

Characteristics of trial publications reporting run-in periods

The trials with run-in periods were larger than trials without (median number of randomized patients 217 and 90, respec-
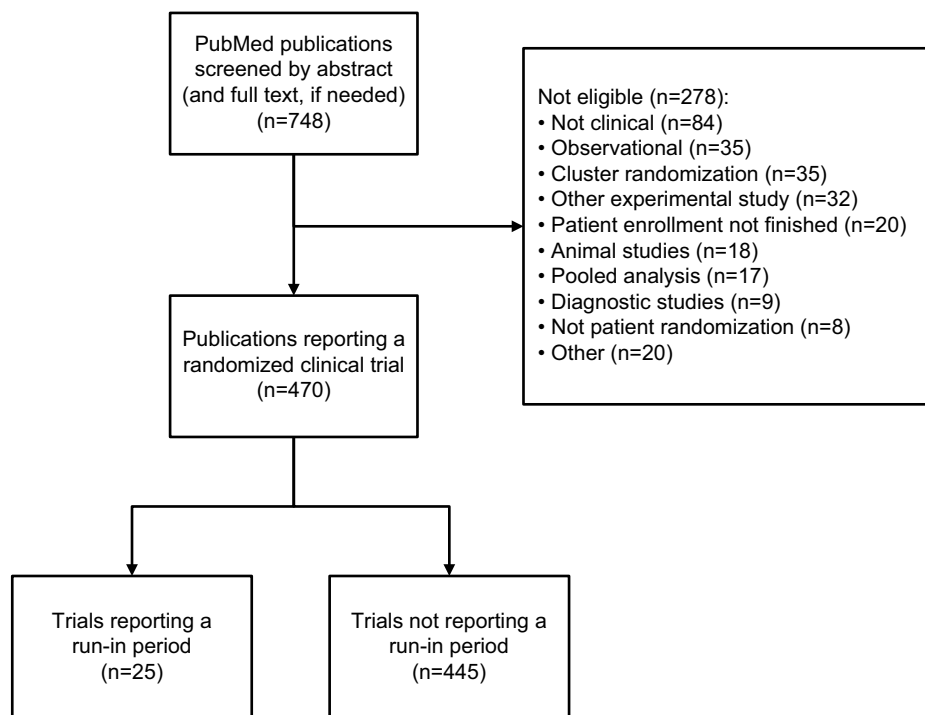
Figure 1 Flowchart of screening for randomized clinical trials and for run-in periods.ᵃ

Notes: ᵃWe screened PubMed publications from 2014 one by one in random order until we had obtained 25 randomized clinical trials reporting a run-in period. The PubMed query used was: "(randomized controlled trial[Publication Type]) AND ("2014/01/01"[Date - Publication]: "2014/12/31"[Date - Publication])".

tively, P=0.01). All the run-in trials were published in English. Six of the eight trials with non-pharmacological interventions studied dietary or lifestyle interventions (Tables 1 and 2).

Also, 16 of 25 trials with a run-in period (64%) were industry trials. This proportion was much higher than among trials without run-in periods (23 out of 82, 28%, P<0.01, the denominator 82 came from disregarding 18 out of 100 trials with unclear industry status; Table 1). Extrapolating the latter proportion of 28% from the 82 trials in the sample to all 445 trials without a run-in period, these would include ~125 industry trials. Thus, among all 470 randomized clinical trials with and without run-in periods, there would be a total of 141 industry trials (the sum of 16 and 125) and conversely 329 non-industry trials. It follows that an estimated proportion of publications reporting a run-in period was 11% among industry trials (16 out of 141) and 3% among non-industry trials (9 out of 329).

In the 13 of 25 run-in period trials reporting relevant data, a median of 16% of enrolled patients were excluded during the run-in period (IQR: 5%–24%). In the 24 of 25 trials reporting relevant data, the median duration of the run-in period was 14 days (IQR: 11–28); for 20 trials, the duration of the run-in period was stated as a fixed number,

whereas 4 trials reported varying duration (eg, "7±3 days"). The intervention during the run-in period was, in most cases, no intervention (36%, 9 out of 25) or placebo (28%, 7 out of 25). No trial used the experimental treatment of the randomized phase as the run-in intervention (Table 3).

Common purposes of the run-in period were to ensure symptom stability (7 out of 25, 28%), for example, headache frequency over 1 month, and "baseline data collection" (5 out of 25, 2%). Most of the trials (20 out of 25, 80%) used the term "run-in period" (in some variation), whereas 3 trials (12%) used the term "baseline period" and 3 other trials (12%) used "lead-in period".

Completeness of reporting of run-in periods in trial publications

Two trials (8%) had complete reporting of run-in periods, because there were no exclusions during the run-in period in both cases. In 23 of the trials (92%), the reporting of run-in periods was incomplete according to our definition.

The main reason for incomplete reporting was that trials did not report the characteristics of excluded patients (22 out of 25 trials, 88%). Reporting for the two other aspects was also incomplete in many trials: 48% and 72% of trials did not

Table 1 General characteristics of randomized clinical trials with and without run-in periods[a]

Characteristic	Trials reporting a run-in period (n=25)	Trials not reporting a run-in period (n=100)	P-value
Number of randomized patients[b]	217 (133–502)	90 (46–354)	0.01*
Trial design			
Parallel	21 (84%)	90 (90%)	0.31
Crossover	4 (16%)	7 (7%)	
Split-body	0 (0%)	3 (3%)	
Number of treatment arms	2 (2–3)	2 (2–2)	0.13
Control group comparators[c]			
Active	15 (60%)	63 (63%)	0.82
Placebo	15 (60%)	18 (18%)	<0.01*
Standard therapy	2 (8%)	20 (20%)	0.24
No treatment	1 (4%)	6 (6%)	1.00
Treatment class[d]			
Pharmacological	17 (68%)	51 (51%)	0.18
Non-pharmacological	8 (32%)	49 (49%)	
Clinical specialty			
Medicine	19 (76%)	44 (44%)	0.02*
Surgery	3 (12%)	28 (28%)	
Other	3 (12%)	28 (28%)	
Language			
English	25 (100%)	91 (91%)	0.67
Chinese	0 (0%)	6 (6%)	
German	0 (0%)	1 (1%)	
Russian	0 (0%)	2 (2%)	
Industry status			
Industry trials	16 (64%)	23 (23%)	<0.01*
Not industry trials	9 (36%)	59 (59%)	
Unclear[e]	0 (0%)	18 (18%)	

Notes: [a]For categorical data, absolute numbers are shown, as well as percentages in parentheses and P-values from Fisher's two-sided exact test. For numerical data, medians are shown, as well as interquartile ranges in parentheses and P-values from the Mann–Whitney U test. The P-value is followed by an asterisk when P<0.05. For the trials without a run-in period, we extracted data from a random sample of 100 trials out of the total 445 trials. [b]The number of randomized patients was reported for 24 (96%) of the trials with a run-in period and 98 (98%) of the trials without a run-in period. [c]One trial could have more control group comparators. [d]Non-pharmacological trials include trials having more treatment arms where at least one of them is non-pharmacological. [e]Trials were classified as unclear if no information was available on study designers, funding, sponsorship, support or similar.

report number of excluded patients and exclusion reasons, respectively (Table 4).

Sensitivity analysis

Our main results were robust to variations in how we operationally defined "run-in period" and "industry trial" (see Supplementary materials).

Discussion

In a representative sample of randomized clinical trial publications, ~1 in 20 reported a run-in period, though in industry trials, the proportion was higher (1 in 10) than in non-industry trials (1 in 30). Trials with run-in periods were typically large industry trials with a placebo control group. A median of 16% of included patients were excluded during run-in periods of a median of 14 days. The run-in procedures differed in design and purpose, but in approximately nine of ten trial publications, the reporting on the run-in period was

too incomplete for a meaningful assessment of its potential impact on the trial result.

Strengths and challenges

To our knowledge, this is the first study of the characteristics and reporting of run-in periods in a random sample of randomized trials. It was based on contemporary trial publications indexed in PubMed, publications that a clinician may typically access.

Our sample size of 25 trials was chosen to provide an overview of the typical trials with a run-in period. A considerably larger sample size would have been required for a comprehensive overview of possible subgroup characteristics. Some clinical fields were not covered by our sample. For example, we screened ~30 psychiatry trials, but none of these reported a run-in period, even though run-in periods are often believed to occur frequently in psychiatry trials. A previous review has studied the impact of the placebo run-in

Table 2 Characteristics of randomized clinical trials with run-in periods

Trial	Subject: treatment	Industry or not	Study design	Study arms	Number of patients enrolled, excluded in run-in, randomized	Run-in type, duration	Comments on reporting of run-in period	Purpose of run-in period
Angelin et al[9]	Hypercholesterolemia: eprotirome	Industry	P	Two drug arms with different doses, one placebo arm	NR, NR, NR	Dietary lead-in, 4 weeks	Number of patients excluded, baseline characteristics and exclusion reasons not reported. Possibly, no exclusions during run-in period, but this is unclear	"Lead-in" to diet, washout
Bleecker et al[10]	Asthma: fluticasone furoate/vilanterol	Industry	P	Drug and co-drug vs drug vs placebo	730, 120, 610	Inhaled corticosteroid only, 4 weeks	Baseline characteristics and detailed exclusion reasons not reported	To ensure symptom stability. Baseline data collection. Minimize placebo response
Casabé et al[11]	Benign prostatic hyperplasia: tadalafil/finasteride	Industry	P	Drug and co-drug vs placebo and co-drug	NR, NR, 696	Placebo, 4 weeks	Number of excluded patients, baseline characteristics and exclusion reasons not reported	NR
Church et al[12]	COPD: umeclidinium and tiotropium	Industry	X	Six drug arms with different doses, one alternative drug arm, one placebo arm	163, 0, 163	NR, NR	No exclusions of patients during run-in. Reporting on exclusion reasons and baseline characteristics deemed complete	NR
De Boever et al[13]	Asthma: an anti-IL-13 mAb	Industry	P	Drug vs placebo	NR, NR, 198	Inhaled corticosteroid, 4 weeks	Number of excluded patients unclear (flowchart numbers do not add up). Baseline characteristics and detailed exclusion reasons not reported	Stable symptomatic disease
Diamond et al[14]	Endometriosis: elagolix	Industry	P	Two drug arms with different doses, one placebo arm	NR, NR, 155	Placebo, 4 weeks	Number of patients excluded, baseline characteristics and exclusion reasons not reported. Possibly, no exclusions during run-in period, but this is unclear	Minimize placebo effect

(Continued)

Table 2 (*Continued*)

Trial	Subject: treatment	Industry or not	Study design	Study arms	Number of patients enrolled, excluded in run-in, randomized	Run-in type, duration	Comments on reporting of run-in period	Purpose of run-in period
Dodick et al[15]	Migraine: LY2951742, an mAb to calcitonin gene-related peptide	Industry	P	Drug vs placebo	367, 149, 218	No treatment, 28–38 days	Baseline characteristics and exclusion reasons not reported	Inclusion based on symptoms over 28 days. To ensure compliance to data collection
Fitzpatrick et al[16]	CVD in diabetes: intensive weight loss program	Not industry	P	Experimental lifestyle treatment vs standard lifestyle treatment	5,579, 434, 5,145	No treatment, 2 weeks	Baseline characteristics and explicit exclusion reasons not reported. Exclusions during run-in given as percentage not absolute number	Complete self-monitoring
Haab et al[17]	Overactive bladder: netupitant, a neurokinin-1 receptor antagonist	Industry	P	Three drug arms with different doses, one placebo arm	325, 79, 246	Placebo, 2 weeks	Baseline characteristics and exclusion reasons not reported	Inclusion based on symptoms. Baseline symptoms data collection
Halmos et al[18]	Irritable bowel syndrome: a diet low in FODMAPs or a typical Australian diet	Not industry	X	Experimental diet vs other control diet	45, 1, 44	No treatment, 1 week	One patient withdrew during baseline/run-in period. Missing baseline characteristics, but period was used for baseline characteristic collection. However, some baseline characteristics might have been reported	Baseline data collection
Hanhineva et al[19]	Impaired glucose metabolism: diet with whole grain, fatty fish and bilberries	Not industry	P	Two experimental diet arms, one control diet	NR, NR, 131	No treatment, 4 days	Number of patients excluded, baseline characteristics and exclusion reasons not reported. Maybe no exclusions during run-in, but unclear	Baseline data collection
Hoare et al[20]	Depression in HIV/AIDS: escitalopram	Industry	P	Drug vs placebo	NR, NR, 105	Placebo, 7±3 days	Number of patients excluded, baseline characteristics and exclusion reasons not reported. Maybe no exclusions during run-in, but unclear	To exclude early placebo responders. To ensure compliance to intervention

(*Continued*)

Table 2 (*Continued*)

Trial	Subject: treatment	Industry or not	Study design	Study arms	Number of patients enrolled, excluded in run-in, randomized	Run-in type, duration	Comments on reporting of run-in period	Purpose of run-in period
Julius et al[21]	Hypertension: candesartan	Industry	P	Drug vs placebo	NR, NR, 809	No treatment, 3 weeks	Number of excluded patients, baseline characteristics and separate exclusion reasons not reported	To ensure stable clinical condition: high blood pressure. Possibly to exclude of patients with adverse events
Laurent and Boutouyrie[22]	Hypertension: olmesartan	Industry	P	Three drug arms with different doses	202, 69, 133	Placebo, 2 weeks	Baseline characteristics and exclusion reasons not reported	NR
Maneechotesuwan et al[23]	Asthma: generic SFC	Not industry	X	Drug vs other drug	NR, NR, 51	Rescue medication only, 2 weeks	Number of patients excluded, baseline characteristics and exclusion reasons not reported. Maybe no exclusions during run-in, but unclear	Washout of previous medication
Marrero et al[24]	Type 2 diabetes: lifestyle intervention, metformin (and troglitazone)	Not industry	P	Drug plus standard treatment vs placebo plus standard treatment vs lifestyle (vs other drug, this arm was discontinued)	4,719, 900, 3,819	Placebo, 3 weeks	Baseline characteristics and separate exclusion reasons not reported	To ensure compliance to intervention. To ensure compliance to data collection
Martinez et al[25]	COPD: Internet-mediated walking program	Not industry	P	Active non-drug treatment vs waitlist control	307, 68, 239	No treatment, 1 week	Baseline characteristics not reported	Baseline data collection. To ensure compliance to data collection
Mugie et al[26]	Functional constipation in children: prucalopride	Industry	P	Drug vs placebo	NR, NR, 215	Laxative or similar if needed, 1–2 weeks	Number of excluded patients, baseline characteristics and separate exclusion reasons not reported	Documentation of symptoms. Washout

(*Continued*)

Table 2 (*Continued*)

Trial	Subject: treatment	Industry or not	Study design	Study arms	Number of patients enrolled, excluded in run-in, randomized	Run-in type, duration	Comments on reporting of run-in period	Purpose of run-in period
Oluleye et al[27]	Heart failure: irbesartan	Industry	P	Drug vs placebo	NR, NR, 4,128	Placebo, 1–2 weeks	Number of patients excluded, baseline characteristics and separate exclusion reasons not reported. Maybe no exclusions during run-in, but unclear	Stable clinical condition
Poulsen et al[28]	Nutrition: new Nordic diet	Not industry	P	Experimental diet vs control diet	190, 9, 181	Active control treatment, 1 week	Baseline characteristics not reported	Standardization to same diet
Reznik et al[29]	Type 2 diabetes: insulin pump treatment	Industry	P	Same drug in two different administration forms	495, 164, 331	Active control treatment, 2 months	Baseline characteristics not reported	Achieve optimum injection treatment. To ensure compliance to data collection
Saneei et al[30]	Childhood metabolic syndrome: DASH diet	Not industry	X	Experimental diet vs standard diet	60, 0, 60	No treatment, 2 weeks	Seemingly no exclusions	Baseline data collection. Learning data collection
Siproudhis et al[31]	Fecal incontinence: NRL001	Industry	P	Three drug arms with different doses, one placebo arm	NR, NR, 466	No treatment, 2 weeks	Design paper. No data reported even though study enrollment is finished	NR
van Gool et al[32]	Urinary incontinence: bladder training, oxybutynin	Not industry	P	Drug vs other treatment vs placebo	NR, NR, 97	No treatment, 3 months	Number of patients excluded, baseline characteristics and exclusion reasons not reported. Maybe no exclusions during run-in, but unclear	To ensure symptom stability: not regained continence
Zhu et al[33]	Hypertension: telmisartan/amlodipine	Industry	P	Drug and co-drug vs co-drug	381, 57, 324	Active control treatment, 6 weeks	Explicit baseline characteristics not reported. States that full analysis set was similar to treated set and to run-in set of patients, but data not shown	To exclude responders to A5 monotherapy lowering the blood pressure. To ensure compliance to intervention

Abbreviations: CVD, cardiovascular disease; mAb, monoclonal antibody; NR, not reported; P, parallel; SFC, salmeterol/fluticasone combination; X, crossover.

Table 3 Characteristics of the run-in period in the 25 trials reporting such a period

Characteristic	Measure[a]
Patient flow[b]	
Number of patients enrolled to run-in period	325 (190–495)
Number of patients excluded during run-in period	69 (9–149)
Percentage of patients excluded during run-in period	16% (5%–24%)
Duration of run-in period in days[c]	14 (11–28)
Intervention during the run-in period	
Control (placebo)	7 (28%)
Control (active)	3 (12%)
Experimental intervention	0 (0%)
Other intervention[d]	5 (20%)
No intervention	9 (36%)
Unclear[e]	1 (4%)
Main purposes of the run-in period[f]	
Extended screening, symptom stability	7 (28%)
Baseline data collection	5 (20%)
Compliance to data collection	3 (12%)
Exclusion of responders to non-active treatment	2 (8%)
Unclear	8 (32%)

Notes: [a]For categorical data, absolute numbers are shown, as well as percentages in parentheses. For numerical data, medians are shown, as well as interquartile ranges in parentheses. [b]Patient flow numbers were reported for 13 out of 25 trials (52%). [c]Run-in period duration was reported for 24 out of 25 trials (96%). [d]For example, inhaled corticosteroid in an asthma trial where this treatment was not an intervention arm during the randomized phase. [f]Some trials reported more than one rationale of the run-in period. Here, we present the purpose most prominently mentioned. For example, four additional trials had compliance to data collection or compliance to run-in intervention noted as a secondary purpose.

Table 4 Completeness of reporting on the run-in periods[a]

Characteristic	Number (percentage)
Overall level of reporting	
Complete reporting[b]	2 (8%)
Incomplete reporting[c]	23 (92%)
Reports number of patients excluded during run-in period	
Yes	13 (52%)
No	12 (48%)
Reports reasons for exclusion of patients	
Yes	7 (28%)
No	18 (72%)
Reports baseline characteristics for excluded patients	
Yes	2 (8%)
No	22 (88%)
Unclear[d]	1 (4%)

Notes: [a]Absolute numbers are shown, as well as percentages in parentheses. A trial run-in period was completely reported if the publication clearly described 1) the number of excluded patients, 2) reasons for exclusion and 3) baseline characteristics of the excluded patients. [b]Two publications reported trials with run-in periods, but stated to have no exclusions during this phase. They were, therefore, deemed to be completely reported on all three criteria. [c]One publication was a rationale and design paper reporting on a trial where patient recruitment and randomization was completed. The publication did not present information on the trial's 2-week run-in period. We deemed reporting to be incomplete for all three criteria. [d]Data were not presented, but excluded and randomized patients were stated to be similar.

period on the efficacy of antidepressants. Sixty-seven of the 141 trials included (48%) used a placebo run-in period.[34] Our sample size would also be too small for the detection of small or modest differences between trials with and without run-in periods. We did, however, detect a clear difference between trials with and without a run-in period with respect to industry status.

Our study addressed reporting in trial publications, and not the frequency of conducted but unreported run-in periods. Twelve of the 25 trials from our study were reported in multiple publications, and in 2 cases, we found examples of unreported run-in periods.

We have not investigated whether run-in periods are reported in more detail in other formats than trial publications, for example, protocols, trial registers, study reports or regulatory agency documents. We chose to focus on trial publications as this is by far the most accessible and accessed format for communicating trial findings.

Other similar studies

Two previous reviews have addressed run-in periods in trials of patients with chronic pain.[35,36] The reviews analyzed randomized clinical trials using "enriched enrollment", a variant of the active run-in-period where patients were randomized if they tolerated and responded to active treatment. The first review described characteristics of the trials and enrichment, including discontinuation rate. In the eight included trials, the average discontinuation rate during the enrichment phase was 35% (compared to our median of 16%).[35] The second review identified many of the same trials as the first one.[36] Our study sample did not include active run-in periods that may cause even more frequent patient exclusions due to harms and lack of response.

In our study, we were not able to obtain the characteristics of the excluded patients during the run-in period for any of the 25 trials. We are not aware of reviews of run-in trials that have compared excluded and randomized patients, but publications on single trials have been published. These indicate that the characteristics of excluded patients and randomized patients may be similar or quite different, depending on the study.[37–41] It would be relevant for the clinician to inspect the characteristics closely in order to relate the trial population to their own patients.

We did not investigate the impact of run-in periods on post-randomization attrition rates, partly because attrition was reported poorly. However, in a previous review for depression, trials with run-in periods did not seem to lower the attrition rates.[42]

Reviews of interventions for depression, weight loss and chronic pain trials reported that run-in periods also did not seem not to alter the effect sizes.[34,42–45] These empirical results are somewhat at odds with the theoretical reasoning behind using run-in periods. Possible explanations for the unexpected results could be unusual clinical settings, low run-in exclusion rates and low statistical power in the trials or in the reviews in question. Further adequately powered empirical review studies would be interesting.

Mechanisms and perspectives

Approximately 20,000 new randomized clinical trial publications are listed each year in PubMed, so we estimate that about 1,000 trial publications yearly report a run-in period (see Supplementary materials). The impact of such trials is larger than reflected solely by their number because they tend to be comparatively larger industry trials which inform decisions made by physicians and regulatory authorities more often than smaller non-industry trials.

From the perspective of trial logistics, a run-in period is used to make a trial more statistically efficient, that is, better at detecting a presumed effect of an intervention. Assuming a moderately effective intervention, a trial with a run-in phase will need fewer patients to reach a statistically significant result, if 1) patient attrition is reduced, 2) non-adherence to experimental intervention is reduced, 3) missed appointments and resulting lack of data is reduced and 4) fewer patients participate who respond well to placebo or poorly to the experimental interventions.

Similarly, a run-in phase will fit a trial with the restricted objective to evaluate the effect of an intervention under ideal conditions, that is, an "explanatory" or a "proof-of-concept" trial assessing "efficacy", and not a "pragmatic" trial assessing clinical "effectiveness" under conditions close to the expected standard clinical situation.[46–48] Thus, a run-in period will tend to improve the sensitivity of the instrument used to detect a treatment effect in compliant patients under a non-standard clinical situation.

However, from the perspective of users of information derived from clinical trials – patients, physicians, authors of systematic reviews and clinical guidelines, and policymakers – the most relevant information are the estimates of the treatment effect sizes applicable to the clinically relevant patient population. A run-in period resulting in pre-randomization exclusion of patients may, if the exclusions are not clearly reported, generally not facilitate such information.

Some proponents argue that active run-in periods can actually imitate the clinical practice of closely monitoring patients when they start a new therapy,[49] and that the relevant effect estimate is the one deriving from compliant patients experiencing minimal harms. However, this is a problematic comparison for three reasons. First, the typical clinical monitoring of patients starting a new therapy is often fairly informal and will often differ considerably from the stricter monitoring in a clinical trial. Second, most clinicians cannot reliably predict or detect noncompliant patients.[1] Third, the risk of harmful effects or anticipated intention-to-treat effect is relevant for those patients who start on the drug. In other words, if patients are excluded in a trial with a run-in period due to noncompliance, a clinician will have considerable difficulty in identifying and treating the patient group for whom a treatment effect has been shown. The problem with applicability also applies with regards to harms of the intervention. In active or placebo run-in periods, the occurrence of harms may be difficult to interpret when no comparison arm is present, and the exclusion of patients who experience harms may underestimate their clinically relevant occurrence.[45,50,51] The more efficient a run-in phase is for excluding a specific category of patients, the less directly clinically applicable the trial result may tend to be.

A run-in period may also impact directly on the internal validity of a trial. Half of the patients in trials with an active or placebo run-in period change study intervention at randomization, either from placebo to experimental or vice versa. This enables them to directly compare experimental and control interventions and increases the risk of bias due to unblinding.[52] In our sample, this problem was relevant in 6 out of 25 trials (24%). Furthermore, selection of highly compliant patients who tolerate treatment well may result in a lower post-randomization attrition rate and lower loss of outcome data. This may inflate the estimated effect of assignment of treatment, the intention-to-treat effect.[53] In 7 of 25 trials (12%) in our sample, patient compliance was one of the purposes of the run-in design.

A run-in period may thus impact on both the external and internal validity of a trial. Furthermore, the run-in period may be considered as an example of "bias by design". Bias is usually understood as synonymous with internal validity, but "bias by design" is a broader concept incorporating aspects of external validity, and refers to design features which increase the chance of detecting an effect at the cost of clinical applicability. Other potential examples of bias by design are selection of an inadequate comparator,[54–57] short trial duration,[54,57,58] selection of clinically irrelevant outcome measures[54,59] and narrow inclusion criteria.[54,60,61] Bias by design has been suggested as one possible mechanism for why industry trials tend to

have more favorable conclusions and outcomes than non-industry trials.[62]

Disagreement may exist as to if and when a run-in period affects external and internal validity. However, the proper assessment of the impact of a run-in period on both internal and external validity of a trial relies substantially on adequate reporting. We document that reporting of run-in periods in trial publications is generally inadequate, in line with similar findings for randomization, blinding and attrition.[63–66]

Implications

Ideally, a reader of a trial publication wants to be able to apply a reliable trial result to an identifiable group of patients. Thus, it is important that the exclusion process in a run-in period is transparent and that any excluded patients are described in sufficient detail. At present, this is far from the case in the vast majority of trials.

One suggestion for improvement is to include reporting of run-in periods in the next revision of the CONSORT (Consolidated Standards of Reporting Trials) statement. The present version CONSORT 2010 on the reporting of trial publications does not offer advice on how to report the use of run-in periods.[67,68] We suggest that trials with a run-in period could report this in an adjusted CONSORT flow diagram (eg, interposed between screening and randomization in the current CONSORT flow diagram) and include information on the number of excluded patients and reasons for exclusions. We also suggest that trial publications report the baseline characteristics of excluded patients (eg, in a table).

The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 statement on reporting trial protocols briefly states that a protocol should report run-in periods when describing intervention dosing schedules and in the participant timeline.[69,70] One option is to expand that section for the revised version of the statement.

While awaiting reporting guideline updates and improved reporting of run-in periods in trial publications, we suggest that results from trials with run-in periods are always interpreted cautiously with respect to external validity and, in many cases, also with respect to internal validity.

Conclusion

The frequency of randomized clinical trials with run-in periods was, on average, ~5%, but three times as frequent in industry trials as compared to non-industry trials. The run-in procedures differed in design and purpose, but a median of 16% of the included patients were excluded during the run-in periods of a median of 14 days. In approximately nine out of ten trial publications, the reporting on the run-in period was too incomplete for a meaningful assessment of its potential impact on the trial results. We suggest that updates of reporting guidelines for randomized trials address run-in periods. We propose the minimum information needed for complete reporting, and we recommend that results from trials with run-in periods are interpreted cautiously with respect to both internal and external validity.

Data sharing statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

AH and DRTL conceived and designed the study. DRTL and ASP performed the data extraction. DRTL and AH performed the data analysis. All authors contributed to drafting or revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Pablos-Méndez A, Barr RG, Shea S. Run-in periods in randomized trials: implications for the application of results in clinical practice. *JAMA*. 1998;279(3):222–225.
2. Talley NJ, Janssens J, Lauritsen K, Rácz I, Bolling-Sternevald E. Eradication of *Helicobacter* pylori in functional dyspepsia: randomised double blind placebo controlled trial with 12 months' follow up. *BMJ*. 1999;318(7187):833–837.
3. Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH). Draft guidance for industry: enrichment strategies for clinical trials to support approval of human drugs and biological products. Available from: http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm332181.pdf. Published December 2012, Accessed April 9, 2017.
4. Hale M, Khan A, Kutch M, Li S. Once-daily OROS hydromorphone ER compared with placebo in opioid-tolerant patients with chronic low back pain. *Curr Med Res Opin*. 2010;26(6):1505–1518.
5. Lang JM, Buring JE, Rosner B, Cook N, Hennekens CH. Estimating the effect of the run-in on the power of the physicians' health study. *Stat Med*. 1991;10(10):1585–1593.
6. RANDOM.ORG. Integer generator. Available from: https://www.random.org/integers/. Accessed November 22, 2018.
7. Google Translate. Available from: https://translate.google.com/. Accessed November 22, 2018.

8. Performing real statistical analysis using Excel. Available from: http://www.real-statistics.com/. Accessed November 29, 2018.

9. Angelin B, Kristensen JD, Eriksson M, et al. Reductions in serum levels of LDL cholesterol, apolipoprotein B, triglycerides and lipoprotein(a) in hypercholesterolaemic patients treated with the liver-selective thyroid hormone receptor agonist eprotirome. *J Intern Med*. 2015;277(3):331–342.

10. Bleecker ER, Lötvall J, O'Byrne PM, et al. Fluticasone furoate-vilanterol 100-25 mcg compared with fluticasone furoate 100 mcg in asthma: a randomized trial. *J Allergy Clin Immunol Pract*. 2014;2(5):553–561.

11. Casabé A, Roehrborn CG, da Pozzo LF, et al. Efficacy and safety of the coadministration of tadalafil once daily with finasteride for 6 months in men with lower urinary tract symptoms and prostatic enlargement secondary to benign prostatic hyperplasia. *J Urol*. 2014;191(3):727–733.

12. Church A, Beerahee M, Brooks J, Mehta R, Shah P. Dose response of umeclidinium administered once or twice daily in patients with COPD: a randomised cross-over study. *BMC Pulm Med*. 2014;14(1):2.

13. De Boever EH, Ashman C, Cahn AP, et al. Efficacy and safety of an anti-IL-13 mAb in patients with severe asthma: a randomized trial. *J Allergy Clin Immunol*. 2014;133(4):989–996.

14. Diamond MP, Carr B, Dmowski WP, et al. Elagolix treatment for endometriosis-associated pain: results from a Phase 2, randomized, double-blind, placebo-controlled study. *Reprod Sci*. 2014;21(3):363–371.

15. Dodick DW, Goadsby PJ, Spierings EL, Scherer JC, Sweeney SP, Grayzel DS. Safety and efficacy of LY2951742, a monoclonal antibody to calcitonin gene-related peptide, for the prevention of migraine: a Phase 2, randomised, double-blind, placebo-controlled study. *Lancet Neurol*. 2014;13(9):885–892.

16. Fitzpatrick SL, Jeffery R, Johnson KC, et al. Baseline predictors of missed visits in the look AHEAD study. *Obesity*. 2014;22(1):131–140.

17. Haab F, Braticevici B, Krivoborodov G, Palmas M, Zufferli Russo M, Pietra C. Efficacy and safety of repeated dosing of netupitant, a neurokinin-1 receptor antagonist, in treating overactive bladder. *Neurourol Urodyn*. 2014;33(3):335–340.

18. Halmos EP, Christophersen CT, Bird AR, Shepherd SJ, Gibson PR, Muir JG. Diets that differ in their FODMAP content alter the colonic luminal microenvironment. *Gut*. 2015;64(1):93–100.

19. Hanhineva K, Lankinen MA, Pedret A, et al. Nontargeted metabolite profiling discriminates diet-specific biomarkers for consumption of whole grains, fatty fish, and bilberries in a randomized controlled trial. *J Nutr*. 2015;145(1):7–17.

20. Hoare J, Carey P, Joska JA, Carrara H, Sorsdahl K, Stein DJ. Escitalopram treatment of depression in human immunodeficiency virus/acquired immunodeficiency syndrome: a randomized, double-blind, placebo-controlled study. *J Nerv Ment Dis*. 2014;202(2):133–137.

21. Julius S, Egan BM, Kaciroti NA, Nesbitt SD, Chen AK; TROPHY Investigators. In prehypertension leukocytosis is associated with body mass index but not with blood pressure or incident hypertension. *J Hypertens*. 2014;32(2):251–259.

22. Laurent S, Boutouyrie P; Vascular Mechanism Collaboration. Dose-dependent arterial destiffening and inward remodeling after olmesartan in hypertensives with metabolic syndrome. *Hypertension*. 2014;64(4):709–716.

23. Maneechotesuwan K, Assawabhumi J, Rattanasaengloet K, Suthamsmai T, Pipopsuthipaiboon S, Udompunturak S. Comparison between the effects of generic and original salmeterol/fluticasone combination (SFC) treatment on airway inflammation in stable asthmatic patients. *J Med Assoc Thai*. 2014;97(Suppl 3):S91–S100.

24. Marrero D, Pan Q, Barrett-Connor E, et al. Impact of diagnosis of diabetes on health-related quality of life among high risk individuals: the Diabetes Prevention Program outcomes study. *Qual Life Res*. 2014;23(1):75–88.

25. Martinez CH, Moy ML, Nguyen HQ, et al. Taking healthy steps: rationale, design and baseline characteristics of a randomized trial of a pedometer-based Internet-mediated walking program in veterans with chronic obstructive pulmonary disease. *BMC Pulm Med*. 2014;14(1):12.

26. Mugie SM, Korczowski B, Bodi P, et al. Prucalopride is no more effective than placebo for children with functional constipation. *Gastroenterology*. 2014;147(6):1285–1295.

27. Oluleye OW, Rector TS, Win S, et al. History of atrial fibrillation as a risk factor in patients with heart failure and preserved ejection fraction. *Circ Heart Fail*. 2014;7(6):960–966.

28. Poulsen SK, Due A, Jordy AB, et al. Health effect of the New Nordic Diet in adults with increased waist circumference: a 6-mo randomized controlled trial. *Am J Clin Nutr*. 2014;99(1):35–45.

29. Reznik Y, Cohen O, Aronson R, et al. Insulin pump treatment compared with multiple daily injections for treatment of type 2 diabetes (OpT2mise): a randomised open-label controlled trial. *Lancet*. 2014;384(9950):1265–1272.

30. Saneei P, Hashemipour M, Kelishadi R, Esmaillzadeh A. The Dietary Approaches to Stop Hypertension (DASH) diet affects inflammation in childhood metabolic syndrome: a randomized cross-over clinical trial. *Ann Nutr Metab*. 2014;64(1):20–27.

31. Siproudhis L, Jones D, Shing RN, Walker D, Scholefield JH; Libertas Study Consortium. Libertas: rationale and study design of a multicentre, Phase II, double-blind, randomised, placebo-controlled investigation to evaluate the efficacy, safety and tolerability of locally applied NRL001 in patients with faecal incontinence. *Colorectal Dis*. 2014;16(Suppl 1):59–66.

32. van Gool JD, de Jong TP, Winkler-Seinstra P, et al. Multi-center randomized controlled trial of cognitive treatment, placebo, oxybutynin, bladder training, and pelvic floor training in children with functional urinary incontinence. *Neurourol Urodyn*. 2014;33(5):482–487.

33. Zhu D, Gao P, Holtbruegge W, Huang C. A randomized, double-blind study to evaluate the efficacy and safety of a single-pill combination of telmisartan 80 mg/amlodipine 5 mg versus amlodipine 5 mg in hypertensive Asian patients. *J Int Med Res*. 2014;42(1):52–66.

34. Trivedi MH, Rush H. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*. 1994;11(1):33–43.

35. Katz N. Enriched enrollment randomized withdrawal trial designs of analgesics: focus on methodology. *Clin J Pain*. 2009;25(9):797–807.

36. McQuay HJ, Derry S, Moore RA, Poulain P, Legout V. Enriched enrolment with randomised withdrawal (EERW): time for a new look at clinical trial design in chronic pain. *Pain*. 2008;135(3):217–220.

37. Hudmon KS, Chamberlain RM, Frankowski RF. Outcomes of a placebo run-in period in a head and neck cancer chemoprevention trial. *Control Clin Trials*. 1997;18(3):228–240.

38. Perkovic V, Joshi R, Patel A, Bompoint S, Chalmers J; ADVANCE Collaborative Group. ADVANCE: lessons from the run-in phase of a large study in type 2 diabetes. *Blood Press*. 2006;15(6):340–346.

39. Fukuoka Y, Gay C, Haskell W, Arai S, Vittinghoff E. Identifying factors associated with dropout during prerandomization run-in period from an mHealth physical activity education study: the mPED trial. *JMIR Mhealth Uhealth*. 2015;3(2):e34.

40. Ulmer M, Robinaugh D, Friedberg JP, Lipsitz SR, Natarajan S. Usefulness of a run-in period to reduce drop-outs in a randomized controlled trial of a behavioral intervention. *Contemp Clin Trials*. 2008;29(5):705–710.

41. Walline JJ, Jones LA, Mutti DO, Zadnik K. Use of a run-in period to decrease loss to follow-up in the Contact Lens and Myopia Progression (CLAMP) study. *Control Clin Trials*. 2003;24(6):711–718.

42. Greenberg RP, Fisher S, Riter JA. Placebo washout is not a meaningful part of antidepressant drug trials. *Percept Mot Skills*. 1995;81(2):688–690.

43. Lee S, Walker JR, Jakul L, Sexton K. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? A meta-analytic evaluation. *Depress Anxiety*. 2004;19(1):10–19.

44. Affuso O, Kaiser KA, Carson TL, et al. Association of run-in periods with weight loss in obesity randomized controlled trials. *Obes Rev*. 2014;15(1):68–73.

45. Furlan A, Chaparro LE, Irvin E, Mailis-Gagnon A. A comparison between enriched and nonenriched enrollment randomized withdrawal trials of opioids for chronic noncancer pain. *Pain Res Manag*. 2011;16(5):337–351.

46. Kramer MS, Shapiro SH. Scientific challenges in the application of randomized trials. *JAMA*. 1984;252(19):2739–2745.

47. Bailey KR. Generalizing the results of randomized clinical trials. *Control Clin Trials*. 1994;15(1):15–23.

48. Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials*. 2009;10(1):37.

49. Franciosa JA. Commentary on the use of run-in periods in clinical trials. *Am J Cardiol*. 1999;83(6):942–944.

50. Hewitt DJ, Ho TW, Galer B, et al. Impact of responder definition on the enriched enrollment randomized withdrawal trial design for establishing proof of concept in neuropathic pain. *Pain*. 2011;152(3):514–521.

51. Schroll JB, Penninga EI, Gøtzsche PC. Assessment of adverse events in protocols, clinical study reports, and published papers of trials of orlistat: a document analysis. *PLoS Med*. 2016;13(8):e1002101.

52. Leber PD, Davis CS. Threats to the validity of clinical trials employing enrichment strategies for sample selection. *Control Clin Trials*. 1998;19(2):178–187.

53. Higgins J, Sterne J, Savović J. A revised tool for assessing risk of bias in randomized trials. *Cochrane Database Syst Rev*. 2016;10(Suppl 1):29–31.

54. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?" *Lancet*. 2005;365(9453):82–93.

55. Mann H, Djulbegovic B. Comparator bias: why comparisons must address genuine uncertainties. *J R Soc Med*. 2013;106(1):30–33.

56. Rochon PA, Gurwitz JH, Simms RW, et al. A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Intern Med*. 1994;154(2):157–163.

57. Safer DJ. Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. *J Nerv Ment Dis*. 2002;190(9):583–592.

58. Pincus T. Rheumatoid arthritis: disappointing long-term outcomes despite successful short-term clinical trials. *J Clin Epidemiol*. 1988;41(11):1037–1041.

59. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. *Trials*. 2017;18(1):122.

60. Rothwell PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials*. 2006;1(1):e9.

61. Charlson ME, Horwitz RI. Applying results of randomised trials to clinical practice: impact of losses before randomisation. *Br Med J (Clin Res Ed)*. 1984;289(6454):1281–1284.

62. Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev*. 2017;2:MR000033.

63. Hopewell S, Dutton S, Yu LM, Chan AW, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ*. 2010;340:c723.

64. Mills EJ, Wu P, Gagnier J, Devereaux PJ. The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials*. 2005;26(4):480–487.

65. Hewitt C, Hahn S, Torgerson DJ, Watson J, Bland JM. Adequacy and reporting of allocation concealment: review of recent trials published in four general medical journals. *BMJ*. 2005;330(7499):1057–1058.

66. Altman DG, Moher D, Schulz KF. Improving the reporting of randomised trials: the CONSORT Statement and beyond. *Stat Med*. 2012;31(25):2985–2997.

67. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.

68. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.

69. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med*. 2013;158(3):200–207.

70. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ*. 2013;346:e7586.

Supplementary materials

Sensitivity analyses
Methods
To study how sensitive our results were to our definition of a run-in period, we evaluated the consequences of a narrower and a broader definition. In comparison with the original definition, the narrower definition only consisted of the main criterion and not the secondary criterion. The broader definition added the terms "washout" to the main and the secondary criterion and "screening" to the main criterion. It also included trials with unclear reporting of the formal time of enrollment, for example, a trial where patients completed a "2-week run-in period to confirm they met the criteria before enrollment".

To study how sensitive our results were to our definition of "industry trial", we defined a broader definition (but not a narrower definition as we did not consider that meaningful). This definition included trials that reported any support, financial or otherwise, from a commercial company (eg, a company provided the study drug free of charge). To study the impact of trials with unclearly reported industry status, we also performed a sensitivity analysis where half of the unclear trials were categorized as industry trials and the other half as not industry trials. Finally, we also presented industry status for the subgroup of pharmacological trials separately.

We also studied how robust our calculation of the frequency of run-in trials among industry trials and non-industry trials was to variation. We calculated the 95% CI of the proportion of industry trials among our sample of trials without run-in periods. Then, we performed the calculations with these upper and lower bounds.

Results
Our finding of incomplete reporting of run-in periods, and our estimation of proportion of trials reporting run-in periods, was robust to variations in how we operationally defined "run-in period". Nine out of the original 25 trials complied with a narrower definition of the term and 11 additional trials would have been included had we adopted a broader definition. So, our estimation of proportion of trials with run-in periods varied from 9 out of 470 (2%) to 36 out of 470 (8%). The two trials without participant exclusions (ie, with complete reporting) during the run-in period were not among the nine publications complying with the narrower definition, nor did any of the additional eleven trials complying with the broader definition present any of the three aspects required for complete reporting. Therefore, in all relevant scenarios, reporting of run-in periods remained clearly incomplete.

Our assessment of industry involvement of the design of the trial was also robust to unclear categorization in 18 of 100 trials with no run-in period and to our operational definition of industry trial. In the subgroup of trials with pharmacological interventions, industry trials were even more drastically overrepresented (94% vs 39% in trials with and without run-in periods, respectively; see Table S1).

The proportion of industry trials among trials without a run-in period was 28% (23 out of 82, 95% CI =19%–39%). Performing our calculations with the upper bound would reveal that 8.4% of the industry trials and 3.2% of the non-industry trials used a run-in period (using the lower bound: 16.0% and 2.4%, respectively). Therefore, at the least, run-in trials seemed to occur 2.6 times as often in industry trials as in non-industry trials.

Table S1 Industry status in trials with and without run-in periods[a]

Type of analysis	Trials reporting a run-in period (n=25)	Trials not reporting a run-in period (n=100)	P-value
Main analysis			
Industry trials	16 (64%)	23 (23%)	<0.01*
Not industry trials	9 (36%)	59 (59%)	
Unclear	0 (0%)	18 (18%)	
Sensitivity analysis, broader definition			
Industry trials	22 (88%)	41 (41%)	<0.01*
Not industry trials	3 (12%)	38 (38%)	
Unclear	0 (0%)	21 (21%)	
Subgroup analysis, pharmacological trials			
Industry trials	16 (94%)	18 (35%)	<0.01*
Not industry trials	1 (6%)	28 (55%)	
Unclear	0 (0%)	5 (10%)	

Notes: [a]Absolute numbers are shown, as well as percentages in parentheses and P-values from Fisher's two-sided exact test. The P-value is followed by an asterisk when $P<0.05$. For the trials without a run-in period, we extracted data from a random sample of 100 trials out of the total 445 trials. Trials were classified as unclear in the main analysis if no information was available on study designers, funding, sponsorship, support or similar. In a sensitivity analysis, we reclassified half of the trials with unclear industry status as industry trials and the other half as not industry trials. In that analysis, the difference between the two groups remained significant ($P<0.01$). With the broader definition, trials were classified as industry trials if they received any support, financial or otherwise, from a commercial company.

Number of randomized clinical trials in PubMed per year

On November 23, 2018, we performed a search in PubMed for randomized clinical trials in the years 2013–2017 using the search string: "(Randomized Controlled Trial[ptyp]) AND ("2013/01/01"[Date - Publication] : "2017/12/31"[Date - Publication])". This yielded 122,408 hits, or an average of 24,482 trial publications per year. A conservative number that excluded, for example, misclassifications and duplicates could then be 20,000. In our study sample, the proportion of publications reporting on trials with a run-in period was ~5%, yielding around 1,000 publications per year.