Title: Pre-randomization run-in periods in randomized controlled trials of chronic diseases: a methodological study

Author names and affiliations:

1) David Collister[a,b] (DC)
Affiliation:
a) Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada
b) Population Health Research Institute, Hamilton, ON, Canada
Address: 3rd floor Marian Wing, 50 Charlton Avenue East, St. Joseph's Healthcare, Hamilton, ON, Canada L8N 4A6
Email: dtcollister@gmail.com
Job title: Clinical Scholar
2) Jennifer C. Rodrigues[a] (JCR)
Affiliation:
a) Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada
Address: 3rd floor Marian Wing, 50 Charlton Avenue East, St. Joseph's Healthcare, Hamilton, ON, Canada L8N 4A6
Email: jc.rodrigues@mail.utoronto.ca
Job title: Clinical Scholar
3) Lawrence Mbuagbaw[a] (LM)
Affiliation:
a) Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada
Address: H-321 Martha Wing, 50 Charlton Avenue East, St. Joseph's Healthcare, Hamilton, ON, Canada L8N 4A6
Email: mbuagblc@mcmaster.ca
Job title: Assistant Professor
4) PJ Devereaux[a,b] (PJD)
Affiliation:
a) Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada
b) Population Health Research Institute, Hamilton, ON, Canada
Address: Room 2C12, McMaster University, Faculty of Health Sciences, 1280 Main Street West, Hamilton, ON, Canada L8S 4K1
Email: PJ.Devereaux@phri.ca
Job title: Professor
5) Gordon Guyatt[a] (GG)
Affiliation:
a) Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada
Address: Room 2C12, McMaster University, Faculty of Health Sciences, 1280 Main Street West, Hamilton, ON, Canada L8S 4K1
Email: guyatt@mcmaster.ca

Job title: Distinguished Professor
6) William Herrington[c] (WH)
Affiliation:
c) Medical Research Council Population Health Research Unit at the University of Oxford
Address: Medical Research Council Population Health Research Unit at the University of Oxford, Clinical Trial Service Unit and Epidemiological Studies Unit, Richard Doll Building, Old Road Campus, Oxford, UK, OX3 7LF
Email: will.herrington@ndph.ox.ac.uk
Job title: Associate Professor
7) Michael Walsh[a,b] (MW)
Affiliation
a) Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada
b) Population Health Research Institute, Hamilton, ON, Canada
Address: 3rd floor Marian Wing, 50 Charlton Avenue East, St. Joseph's Healthcare, Hamilton, ON, Canada L8N 4A6
Email: lastwalsh1975@gmail.com
Job title: Associate Professor

Corresponding author:
David Collister
Address: 3rd floor Marian Wing, 50 Charlton Avenue East, St. Joseph's Healthcare, Hamilton, ON, Canada L8N 4A6
Phone: 1-204-295-7998
Email: dtcollister@gmail.com

Abstract:

Objective: To systematically review the epidemiology of pre-randomization run-in periods in randomized controlled trials (RCTs) of chronic diseases.
Study Design and Setting: Meta-epidemiologic study of all RCTs from the four highest impact medical journals from 2011 to 2016. Eligible trials included parallel RCTs that evaluated pharmacologic therapies in adults with chronic diseases with a minimum follow-up of 24 weeks.
Results: Of 262 eligible manuscripts, 48 (18.3%), representing 42 unique RCTs, included run-in periods. Run-in periods were most common in cardiovascular disease and diabetes trials. Of the 42 RCTs, in 22 patients received the experimental therapy, 15 placebo, 4 both (either sequentially or in combination) and one did not report the run-in period drug. The median run-in period duration was 28 days (Q1:Q3 14:66 days). Reasons for including a run-in period included ensuring eligibility criteria were met (18, 42.9%), excluding participants with non-adherence (18, 42.9%) and intolerances to therapy (15, 35.7%), and to standardize therapy prior to randomization (8, 19.0%). The median run-in completion rate was 77.4% (Q1:Q3 62.2:87.8%).
Conclusions: Run-in periods are uncommon in RCTs of chronic drug treatments and when used, their reporting is heterogeneous. Further research to improve the design, use and reporting of run-in periods is necessary.

Key words: run-in periods, randomized controlled trials, meta-epidemiology

Running title: Run-in periods in trials of chronic diseases

Abstract word count: 200

Manuscript word count: 3278

What Is New

Key findings

-Run-in periods are not commonly used in RCTs of pharmacologic therapies for chronic diseases in adults but are common in cardiovascular disease prevention and diabetes trials.

What this adds to what is known

-The majority of run-in periods are used to identify non-adherence, intolerability, standardize therapy and ensure that participants meet eligibility criteria.

What is the implication, what should change now

-The optimal duration of a run-in period is unknown; empirical studies are necessary in this area.

1.1 Introduction:

Well designed and conducted randomized controlled trials (RCTs) provide the highest quality evidence to assess treatments but are time consuming and costly[1, 2]. A key issue affecting the statistical efficiency and results of RCTs is non-adherence to study interventions. This is because participants randomized to receive an intervention who do not actually receive the study intervention, even if it is effective, will not receive its benefits. The intervention effects for the population will therefore be diminished as will statistical power.

A pre-randomization run-in period was first used in the Physician's Health Study (PHS)[3] to help preserve statistical power by selecting participants likely to adhere to the trial's interventions [4, 5]. However, whether run-in periods actually achieve this goal has not been systematically assessed and the effects of a run-in period may be dependent on its design[6] (e.g. duration, method of classifying or identifying non-adherence, complexity as an additional study procedure, costs). Investigators may also use run-in periods for other reasons such as for additional screening procedures (e.g. standardizing background therapy or confirming eligibility criteria), to identify participants with intolerances to study interventions[7, 8] or those with responses or non-responses to a therapy[9] as well as individuals whom experience a placebo response[10], all of which may influence treatment effects. A run-in period may also affect generalizability of a trials results[11] by altering the trial's population.

Since there is no consensus on the design of run-in periods, we systematically identified and characterized run-in periods in contemporary RCTs evaluating long term, self-administered pharmacologic therapies for chronic diseases in adults where adherence is important.

1.2 Methods:

1.21 Objectives:
Our objectives are to describe the frequency, settings and purposes of run-in periods in RCTs published in high impact general medical journals and to compare those RCTs with and without run-in periods with regards their designs.

1.22 Design:
A meta-epidemiologic study of RCTs published in high impact journals indexed in PubMed.

1.23 Study selection:
PubMed was searched to identify RCTs published in the New England Journal of Medicine (NEJM), the Lancet, the Journal of the American Medical Association (JAMA) and the British Medical Journal (BMJ) using the randomized controlled trial medical subject headings (MeSH) term limiting from January 2011 until August 2016. We limited to high impact journals given their likelihood for publishing "positive" results[12] that we hypothesized would be associated with the use of run-in periods. No MeSH term for

"run-in period" exists. A run-in period was defined as a pre-randomization study procedure in which participants are exposed to a study drug prior to randomization for the purposes of excluding participants prior to randomization and includes "run-ins", "lead-ins" and "enrichment". A purposive sample was drawn by including all RCTs of any sample size addressing therapies for chronic diseases that met the following inclusion criteria: 1) parallel or factorial in design that could contain multiple arms but was not a cluster, stepped-wedge or crossover RCT, 2) compared pharmacologic therapies (active versus active) or a pharmacologic therapy to placebo (active versus placebo) whose administration was either orally, subcutaneously or inhaled but not intravenously (not self-administered so it is less dependent on adherence) and whose frequency of administration was equal to or more than once a day, 3) in the setting of a chronic disease [diseases of long duration and generally slow progression as defined by the World Health Organization[13]], 4) the trial duration was a minimum of 24 weeks (in which long term and potentially declining adherence is important), 5) was not a long term follow-up or extension study, 6) was not a secondary analysis, 7) at least half of the participants are adults greater or equal to 16 years of age.

Studies were uploaded in RefWorks[14]. Two reviewers (DC, JR) independently reviewed all full texts. If a factorial design was present, only one factorial was randomly chosen and included in analysis. All disagreements were resolved by a third reviewer (MW).

1.24 Data abstraction:

Two reviewers (DC, JCR) independently abstracted data for each study using the primary report, supplementary material and protocol. Abstracted data elements included the study author, journal, year, title, population, intervention(s), comparator(s), design (active or placebo controlled), speciality, disease, blinding, size, follow-up duration, adherence, sponsorship (industry or investigator), statistical significance of the primary outcome (defined as a p value less than 0.05) and the presence or absence of a run-in period. For studies with a run-in period, the following data elements were abstracted: the run-in period terminology (e.g. run-in, lead-in, enrichment), the run-in period duration, the reason(s) for run-in period, the type of run-in period (active versus placebo), the type of drug used during the run-in period, patient and physician blinding, the number of participants who completed the run-in period, reasons for run-in failures, the definition of adherence during the run-in period, whether or not the Consolidated Standards of Reporting Trials (CONSORT) diagram included the run-in period and reasons for post-randomization exclusions.

1.25 Statistical Analysis:

Descriptive statistics included frequency and percent, mean and standard deviation or medians and first and third quartiles (Q1:Q3). Studies with and without run-in periods were compared with regards to design, subspecialty, size, blinding, follow-up duration, and sponsorship using a two-sample t test (parametric) or Mann Whitney-U test (nonparametric) for continuous outcomes (normality was assessed using the Shapiro-Wilk normality test) or Chi square test for categorical outcomes. Run-in period duration was compared for each reason for the run-in periods using a Mann Whitney-U test. Run-in completion rates and the proportion of run-in failures due to non-adherence were

compared across run-in period duration (i.e. <1 week, >1-2 weeks, >2-4 weeks, >4 weeks-3 months and >3 months) using the Kruskal Wallis test. All analyses were performed using STATA (StataCorp. 2015. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP).

There was no patient involvement in this study. Given the nature of the study, ethics approval was not considered necessary.

1.3 Results:

Of 262 eligible manuscripts [New England Journal of Medicine (NEJM) 136, Lancet 64, Journal of the American Medical Association (JAMA) 57, British Medical Journal (BMJ) 5], 48 (18.3%) included run-in periods (Figure 1). RCTs with pre-randomization run-in periods were larger (2369 participants, Q1:Q3 358:9651 vs 761 participants, Q1:Q3 300:2463, P=0.006) and had longer follow-up (2 years, Q1:Q3 1:3.85 vs 1.41 years, Q1:Q3 0.875:2.4, P=0.042) than those without run-in periods but had similar design, blinding, sponsorship and statistical significance of their primary outcomes (Table 1).

Of the 48 manuscripts with run-in periods, 4 were excluded as they had duplicate run-in period data due to their factorial designs [Heart Outcome Prevention Evaluation-3 (HOPE-3)[15-17], Outcome Reduction with an Initial Glargine Intervention (ORIGIN)[18, 19], PHS-2[20, 21]]. One study was excluded because its run-in period occurred post-randomization[22] and one study was excluded because it was later retracted due to fraudulent results[23]. Therefore, the analysis included 42 RCTs with run-in periods (Table 2 and Supplemental Table 1 for a detailed summary).

Run-in periods were more common in chronic treatments for cardiovascular disease and diabetes: 16/50 (32%) of cardiovascular disease trials, 12/32 (37.5%) of coronary artery disease trials, 3/11 (27.2%) of heart failure trials and 9/20 (45%) of diabetes trials included run-in periods. Run-in periods were uncommon in oncology, infectious diseases and benign hematology trials. For example, 2/37 (5.4%) of cancer trials included run-in periods (where they were used in the setting of primary prevention and not chemotherapy), as did 1/23 (4.3%) of human immunodeficiency virus trials (in the context of multivitamin and/or selenium supplementation and not highly active anti-retroviral therapy). None of the 8 venous thromboembolism trials (which evaluated chronic anticoagulation) included run-in periods.

Of the 42 RCTs with run-in periods, 21 (50%) used the term run-in period, 10 (23.8%) used the term run-in phase, 4 (9.5%) used run-in variant terms together including period/phase/stage, 1 (2.4%) used the term lead-in, 2 (4.8%) used other terms and 4 (9.5%) did not refer to their run-in periods specifically. The majority (22, 52.3%) used active therapies during the run-in periods, including renin angiotensin aldosterone system inhibitors, inhalers and lipid lowering therapies. Other trials' run-in periods used a placebo design (15, 35.7%), active and placebos either sequentially or in combination (4, 9.5%) and one study did not report the run-in period drug (1, 2.4%). In 15 of 34 trials

(44.1%) patients were blinded, as were physicians in 3 of 34 trials (8.8%) –other trials did not report on blinding (Table 2).

The median duration of a run-in period was 28 days (Q1:Q3 14:66 days; minimum 7 to maximum 365); 36 RCTs included one run-in period, 4 two sequential periods, and 2 RCTs had three run-in periods. Reasons for including a run-in period were as follows (more than one reason could apply to a RCT): 18 (42.9%) to ensure that eligibility criteria were met prior to randomization; 18 (42.9%) to exclude participants with non-adherence; 15 (35.7%) to exclude participants with intolerance to therapy; 8 (19.0%) to standardize therapy prior to randomization and 2 (4.8%) to identify apparent non-responders. Run-in period duration did not differ between RCTs done for each of these reasons (P>0.05 for all comparisons).

The median percentage of participants who completed the run-in period and were subsequently randomized was 77.4% (Q1:Q3 62.2:87.8%) with a range of 38.2%[24] to 97.1%[25]. Run-in completion rate did not significantly vary by run-in duration (Table 3). The run-in completion rate was significantly lower for trials that used them for eligibility purposes (median 0.60, Q1:Q3 0.46:0.77) compared to those that did not use them for this reason (median 0.86, Q1:Q3 0.73:0.91) (P=0.003 by Mann-Whitney U test) with no significant difference for other reasons (i.e. P>0.05 for non-adherence, intolerance, standardize therapy, response).

In the 18 studies in which run-in periods were performed to identify and exclude participants with non-adherence, the duration of run-in period was not associated with run-in completion rate or the proportion of run-in failures due to non-adherence (Supplemental Table 2). However, this analysis included only 13 studies that reported run-in completion rates and 9 studies that specifically reported run-in failures due to non-adherence. Definitions for adherence during the run-in period were variable and ranged from at least 50% to at least 85.7% or between 80-120% of study medications using self-reported adherence or counting. We did not compare run-in period duration with post-randomization adherence or post-randomization exclusions for non-adherence because the necessary information (i.e. adherence at a set time point following randomization) was not reported in a sufficient number of studies.

In 28 of 42 (66.7%) RCTs, the run-in period was included in the CONSORT diagram with run-in failures reported separately from screening. Reasons for run-in failure included withdrawal of consent (23, 54.8%), adverse events (15, 35.7%), failure to meet eligibility criteria (15, 35.7%), loss to follow-up (12, 28.6%), non-adherence (9, 21.4%), abnormal laboratory results (7, 16.7%), protocol deviation (7, 16.7%), death (6, 14.2%), side effects (5, 12.0%), and other (20, 47.6%). In studies in which they were reported, withdrawal of consent, non-adherence and lost to follow-up, (all patient related behaviors) were responsible for 22.9% (Q1:Q3 14.4:27.7%), 26.0% (Q1:Q3 8.9:38%) 4.9% (Q1:Q3 2.1-8.2%) of run-in failures respectively. In studies in which they were reported, screening failure, abnormal laboratory results and side effects (all markers of tolerability) were responsible for 34.5% (Q1:Q3 11.5:70.3%), 50.5% (Q1:Q3 25.7:66.7%) and 6.0% (Q1:Q3 3.0:14.5%) of run-in failures respectively.

1.4 Discussion:

1.41 Principal findings:

Of the over 250 RCTs addressing self-administered pharmacologic therapies for chronic diseases in adults published in high impact journals between 2011 and 2016, only one-fifth included a pre-randomization run-period. Run-in periods were most common in cardiovascular disease and diabetes trials and were used most frequently to ensure eligibility criteria prior to randomization and to exclude participants who were either non-adherent or intolerant to therapy.

1.42 Previous studies:

One other study has reported the epidemiology of run-in periods in a sample of contemporary RCTs[26]. The authors randomly sampled RCTs published in 2014 until they identified 25 with run-in periods from a total of 470 trials. They also compared RCTs with and without run-in periods and similarly found that run-ins were more common in larger trials and reporting was frequently incomplete, particularly in terms of the reasons for run-in failures and the characteristics of patients that failed the run-in period, an issue not unique to trials with run-in periods[27]. However, their results differed from our study in the reasons for using a run-in period because their study's definition of a run-in period was slightly different than ours since it included trials with an extended screening period that did not include an exposure to an intervention. This study also did not focus on high-impact journals, trials of chronic diseases and self-administered drugs, so its' run-in period duration was shorter and the proportion of run-in failures was smaller than in our study. Lastly, it found that run-in periods were more common in industry sponsored trials which we did not.

1.43 Mechanism

Few studies have empirically evaluated or simulated the effect of a run-in period on a trial's power. In the PHS trial[5], 2/3 of participants completed an 18 week run-in period consisting of active aspirin and beta-carotene placebo. If the trial had included patients that failed the run-in and assumed that these individuals took no study medication (zero adherence) or took the same amount study medication as participants who did not fail run-in, neither case would not have significantly changed the power to detect a large treatment effect (relative risk=0.60) since power was 96% and 99% in these two scenarios as compared to 98% with the run-in period. However, for a more modest and realistic treatment effect (relative risk=0.70), power dropped substantially to 79% if it is assumed that those who failed the run-in had very poor adherence. This finding is supported by the Cholesterol Reduction in Seniors Program (CRISP) trial in which those non-adherent to the placebo run-in medication were still randomized. The non-adherence during the placebo run-in did identify a group of participants who were much more likely to be non-adherent during study follow-up, but because only about 15% of participants were non-adherent during run-in, the true effect of lovastatin on low density lipoprotein was only underestimated by a small amount by their inclusion[28]. Whether these findings are generalizable to other chronic disease populations and interventions is uncertain due

to likely variability in background adherence and the impact of non-adherence on treatment effects[29].

In diseases where non-adherence is common, such as end-stage kidney disease (ESKD), non-adherence could have substantial effects on power. For example, in the EValuation Of Cinacalcet Hydrochloride to Lower CardioVascular Events) EVOLVE[30] trial, the calculated power decreased from approximately 90% (assuming a 10% drop-out in the cinacalcet group and a 10% drop-in in the placebo group) to 54% due to a drop-out rate of 62% (due to non-adherence and intolerance) and 19.8% drop-in rate (in addition to an overall lower expected number of events). Whether power could have been improved with a run-in period is unknown but in the Study of Heart and Renal Protection (SHARP)[31] which had a similar ESKD population and a 6-week placebo run-in period, adherence was marginally better at 71% at the study midpoint of 2.5 years. Schechtman et al.[6] showed in their simulation study that run-in periods are most likely to be useful if the randomized proportion of participants is not high, there are many zero adherers and if partial-adherers have a limited response to therapy.

1.44 Unanswered questions and future research

Given that the median run-in completion rate was 77.4% (Q1:Q3 62.2:78.8%) in our sample, most trials with run-in periods appeared to effectively identify and exclude participants for their intended reasons. However, some studies had excellent run-in completion rates such as the study by Marso et al. where only 2.9% of participants failed the 2 week run-in period due to <50% adherence with subcutaneous placebo injections or withdrawal suggesting the run-in period was either unnecessary or the threshold to define non-adherence was too liberal[25]. There are still many basic questions regarding the optimal use and design of run-in periods. For example, whether longer duration run-in periods or more or less stringent methods of identifying non-adherence during the run-in improve the selection of participants who will adhere to chronic therapies is uncertain. If longer or more complicated run-in designs do not improve participant selection, they will add unnecessary costs and complexity and possibly affect the feasibility of the overall trial. Because non-adherence has potentially dramatic effects on the efficiency of a RCT and run-in periods are a potentially inexpensive and straightforward mechanism to identify participants who will be adherent, it is important to determine their optimal characteristics.

The trade-off between potentially increased statistical power and loss of generalizability needs to be considered by investigators that are contemplating using a run-in period in their trial designs. The degree to which a trial's population is modified by excluding non-adherent participants, those with intolerability or for other reasons during a run-in period presumably impacts how clinicians interpret results and use them in their practice both on a population and individual patient level. How this actually impacts shared decision making across clinicians and specific populations, interventions and outcomes is not known.

1.45 Strength and weaknesses of the study

Our study is notable in the breadth of its search, the completeness of its full text review to ensure all RCTs with a run-in period were included, and the detailed collection of data from supplemental materials and protocols. While our study includes almost twice as many trials with run-in periods as prior reviews, it is still relatively small and restricted to high impact general medical journals. The use of high impact journals in particular may result in misestimating the prevalence of their use or their association with adherence. However, if anything, RCTs published in high impact journals are more likely to have "positive" results and, if run-in periods do in fact improve statistical efficiency, they are therefore more likely to have employed run-in periods. A further limitation is our study's uncontrolled nature. Prior reviews comparing trials of a single class of treatments with and without a run-in period (RCTs of dipeptidyl peptidase-4 inhibitors[32] and statins[33]) did not demonstrate an effect of run-in periods on treatment efficacy or safety but these comparisons were likely underpowered. A meta-analysis of 101 randomized placebo-controlled antidepressant trials for major depressive or bipolar disorder did not demonstrate that run-in periods used to exclude placebo responders reduce post-randomization placebo response rate, drug response rate or drug-placebo differences[34]. Having said that, run-in periods may certainly improve power but the impact of run-in periods on effect sizes is unknown. However, it reasonable to assume that trials with run-in periods for non-adherence increase the measured effect size of an intervention if there is a linear relationship between adherence and efficacy without a ceiling effect . Run-in periods for tolerability are likely more complicated as toxicity or side effects not leading to intervention discontinuation may be linked to efficacy as is in the case of angiotensin-converting enzyme inhibitors or chemotherapy[35].

1.46 Meaning of the study

In summary, only one-fifth of high-impact RCTs include a pre-randomization run-period and the details reported are heterogeneous. We recommend that all trials that use a pre-randomization run-in period include it in the trial's CONSORT diagram and explicitly report the number of participants that enter the run-in, complete the run-in and the specific reasons for run-in failures. We recommend people designing RCTs of self-administered therapies for chronic diseases in adults consider the degree of non-adherence and intolerance likely for their population and intervention and consider using a run-in period to improve power if these are likely to be common post-randomization. Pilot trials may offer a unique opportunity to gain insight on these issues as well as optimize the run-in period of the later definitive trial by defining the amount of time needed to identify non-adherence and intolerance.

.

Author contributions: David Collister: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing-Original draft preparation, Visualization **Jennifer C. Rodrigues:** Validation, Writing- Reviewing and Editing **Lawrence Mbuagbaw:** Methodology, Writing- Reviewing and Editing **PJ Devereaux:** Methodology, Writing- Reviewing and Editing **Gordon Guyatt:** Methodology, Writing- Reviewing and Editing **William Herrington:** Writing- Reviewing and Editing **Michael Walsh:** Conceptualization, Methodology, Writing- Reviewing and Editing, Supervision

Declarations of interest: none

Acknowledgements: none

Role of the funding source: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

Research data: The research data is available upon request by with approval from the corresponding author

Table 1: Characteristics of included studies with and without pre-randomization run-in periods from NEJM, Lancet, JAMA and BMJ 2011-2016

	No run-in period (n=214)	Run-in period (n=48)	P value
Placebo controlled: n (%) Active controlled: n (%) Usual care: n (%)	146 (68.2%) 63 (29.4%) 5 (2.34%)	34 (70.8%) 14 (29.2%) 0 (0%)	0.560
Number of patients randomized (median) (Q1:Q3)	761 (300:2463)	2369 (358:9651)	0.006
Duration (median) (years) (Q1:Q3)	1.41 (0.875:2.4)	2 (1:3.85)	0.042
Patient blinding: n (%) Open label: n (%) Both: n (%)	170 (79.4%) 40 (18.7%) 4 (1.9%)	43 (89.6%) 4 (8.3%) 1 (2.1%)	0.222
Industry: n (%) Investigator: n (%)	151 (70.6%) 63 (29.4%)	40 (83.3%) 8 (16.7%)	0.072
Cardiovascular: n (%) -atrial fibrillation: n (%) -coronary artery disease: n (%) -heart failure: n (%) -other: n (%)	34 (15.9%) 5 (14.7%) 20 (58.8%) 8 (23.5%) 1 (2.9%)	16 (33.3%) 0 (0%) 12 (75.0%) 3 (18.8%) 1 (6.3%)	0.003
Oncology: n (%)	37 (17.3%)	2 (4.2%)	
Infectious Diseases: n (%)	37 (17.3%)	1 (2.1%)	
Endocrinology: n (%) -Diabetes: n (%) -other: n (%)	19 (8.9%) 11 (57.9%) 8 (42.1%)	11 (22.3%) 9 (81.8%) 2 (18.1%)	
Pulmonary: n (%) -asthma: n (%) -chronic obstructive pulmonary disease: n (%) -other: n (%)	24 (11.2%) 5 (20.8%) 4 (16.7%) 15 (62.5%)	6 (12.5%) 1 (16.7%) 4 (66.7%) 1 (16.7%)	
Hematology: n (%)	11 (5.1%)	0 (0%)	
Nephrology: n (%)	12 (5.6%)	5 (10.4%)	
Neurology: n (%)	13 (6.1%)	2 (4.1%)	
Post-randomization adherence monitoring: n (%)	97 (45.3%)	25 (52.0%)	0.396
Post-randomization adherence in active arm (continuous): n (%) (median, Q1:Q3)	30 (14.0%) (0.89, 0.81-0.95)	13 (27.0%) (0.86, 0.68-0.99)	0.832
Post-randomization adherence in active arm (binary): n (%) (median, Q1:Q3)	31 (14.5%) (0.89, 0.63-0.99)	5 (10.4%) (0.81, 0.69-0.93)	0.064
Statistically significant primary outcome*: n (%)	147 (68.7%)	30 (62.5%)	0.408

Note: NEJM = New England Journal of Medicine, JAMA = Journal of the American Medical Association, BMJ= British Medical Journal, Q1 = 1st quartile, Q3 = 3rd quartile
*p<0.05 for primary outcome or any co-primary outcome for non-inferiority or superiority

Table 2: Pre-randomization run-in period characteristics (n=42)

Characteristic	Category	Number
Terminology	Run-in period: n (%)	21 (50%)
	Run-in phase: n (%)	10 (23.8%)
	Run-in variant: n (%)	4 (9.5%)
	Lead-in: n (%)	1 (2.4%)
	Other*: n (%)	2 (4.8%)
	Not referred to: n (%)	4 (9.5%)
Duration	≤7 days: n (%)	4 (9.52%)
	>7-14 days: n (%)	8 (19.05%)
	>14-28 days: n (%)	11 (26.19%)
	>28-90 days: n (%)	11 (26.19%)
	>90 days: n (%)	8 (19.05%)
Blinding	patient: n (%)	15/34 (44.12%)
	physician: n (%)	3/34 (8.825%)
Agent	placebo: n (%)	15 (35.71%)
	RAAS inhibitor: n (%)	6 (14.29%)
	Inhaler: n (%)	3 (7.14%)
	lipid lowering agents: n (%)	3 (7.14%)
	antiplatelet agents: n (%)	2 (4.76%)
	other: n (%)	8 (19.04%)
	placebo with other: n (%)	4 (9.52%)
	missing: n (%)	1 (2.38%)
Cardiovascular (n=12)	Coronary artery disease: n	8
	Heart failure: n	3
	Hypertension: n	1
Infectious disease (n=1)	HIV: n	1
Oncology (n=2)	Cancer: n	1
	Colorectal cancer: n	1
Endocrinology (n=10)	Diabetes: n	8
	FHH: n	1
	Obesity: n	1
Pulmonary (n=5)	Asthma: n	1
	COPD: n	4
Rheumatology (n=1)	Retroperitoneal fibrosis: n	1
Geriatrics (n=1)	Alzheimer's dementia: n	1
Nephrology (n=5)	Chronic kidney disease: n	1
	Diabetic nephropathy: n	2
	IgA nephropathy: n	1
	Kidney transplant: n	1
Neurology (n=2)	Stroke: n	1
	Restless legs syndrome: n	1
Ophthalmology (n=2)	Blindness: n	1
	Macular degeneration: n	1
Other (n=1)	Smoking cessation: n	1

Reason for run-in period	adherence: n (%) eligibility: n (%) tolerability: n (%) response: n (%) standardize therapy: n (%)	18 (42.86%) 18 (42.86%) 15 (37.71%) 8 (19.05%) 2 (4.76%)

Note: RAAS = renin-angiotensin aldosterone system = angiotensin converting enzyme inhibitor, angiotensin II receptor blocker, spironolactone, direct renin inhibitor, neprilysin inhibitor, lipid lowering agents = statin, ezetimibe, niacin, laropiprant, placebo with other = metformin, naltrexone-buproprion, AREDS formulation, inhalers/oral corticosteroid, other = ivabradine, risperidone, kidney transplant induction immunosuppression, prednisone, tasimelteon, varenicline, ARB/thiazide/statin, RAAS/statin, HIV=human immunodeficiency virus, FHH=familial homozygous hypercholesterolemia, COPD=chronic obstructive pulmonary disease

*=phase A, open phase

Table 3: Pre-randomization run-in period completion rates by run-in period duration (N=29 of 42 studies)

Run-in period duration	≤1 weeks n=2	>1-2 weeks n=2	>2-4 weeks n=10	>4 weeks-3 months n=8	>3 months n=7
Completion rate median (Q1:Q3)	0.93 (0.92:0.94)	0.71 (0.57:0.86)	0.82 (0.57:0.86)	0.80 (0.65:0.87)	0.60 (0.48:0.71)

Note: Q1 = 1st quartile, Q3 = 3rd quartile, P=0.0824 by Kruskal Wallis test

References:

1. Sertkaya A, Wong HH, Jessup A, Beleche T. Key cost drivers of pharmaceutical clinical trials in the United States. *Clin Trials*. Apr 2016;13(2):117-26. doi:10.1177/1740774515625964
2. Hind D, Reeves BC, Bathers S, et al. Comparative costs and activity from a sample of UK clinical trials units. *Trials*. May 2 2017;18(1):203. doi:10.1186/s13063-017-1934-3
3. Steering Committee of the Physicians' Health Study Research G. Final report on the aspirin component of the ongoing Physicians' Health Study. *The New England journal of medicine*. 1989;321(3):129-135. doi:10.1056/NEJM198907203210301 [doi]
4. Pablos-Mendez 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. doi:jsc71017 [pii]
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. *Statistics in medicine*. 1991;10(10):1585-1593.
6. Schechtman KB, Gordon ME. A comprehensive algorithm for determining whether a run-in strategy will be a cost-effective design modification in a randomized clinical trial. *Statistics in medicine*. 1993;12(2):111-128.
7. Probstfield Jl WDHBSICLDHGMH-lHJHJJMYJB. The studies of left ventricular dysfunction (SOLVD) experience with a placebo run-in phase. *Controlled Clin Trials 10:334, 1989*.
8. Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med*. Mar 21 1991;324(12):781-8. doi:10.1056/nejm199103213241201
9. Boessen R, Heerspink HJ, De Zeeuw D, Grobbee DE, Groenwold RH, Roes KC. Improving clinical trial efficiency by biomarker-guided patient selection. *Trials*. 2014;15:103-6215-15-103. doi:10.1186/1745-6215-15-103 [doi]
10. Rosenkranz GK. Remarks on designs enriching for placebo non-responders. *Clinical trials (London, England)*. 2016;13(3):338-343. doi:10.1177/1740774515625186 [doi]
11. Berger VW, Rezvani A, Makarewicz VA. Direct effect on validity of response run-in selection in clinical trials. *Controlled clinical trials*. 2003;24(2):156-166. doi:S0197245602003161 [pii]
12. Ivanov A, Kaczkowska BA, Khan SA, et al. Review and Analysis of Publication Trends over Three Decades in Three High Impact Medicine Journals. *PLoS One*. 2017;12(1):e0170056. doi:10.1371/journal.pone.0170056
13. Shanthi M, et al. *Global status report on noncommunicable diseases 2014. World Health Organization.*:1-280. http://www.who.int/nmh/publications/ncd-status-report-2014/en/
14. *RefWorks*. ProQuest;
15. Lonn EM, Bosch J, Lopez-Jaramillo P, et al. Blood-Pressure Lowering in Intermediate-Risk Persons without Cardiovascular Disease. *The New England journal of medicine*. 2016;374(21):2009-2020. doi:10.1056/NEJMoa1600175 [doi]
16. Yusuf S, Lonn E, Pais P, et al. Blood-Pressure and Cholesterol Lowering in Persons without Cardiovascular Disease. *The New England journal of medicine*. 2016;374(21):2032-2043. doi:10.1056/NEJMoa1600177 [doi]
17. Yusuf S, Bosch J, Dagenais G, et al. Cholesterol Lowering in Intermediate-Risk Persons without Cardiovascular Disease. *The New England journal of medicine*. 2016;374(21):2021-2031. doi:10.1056/NEJMoa1600176 [doi]

18. Investigators OT, Gerstein HC, Bosch J, et al. Basal insulin and cardiovascular and other outcomes in dysglycemia. *The New England journal of medicine*. 2012;367(4):319-328. doi:10.1056/NEJMoa1203858 [doi]
19. Investigators OT, Bosch J, Gerstein HC, et al. N-3 Fatty Acids and Cardiovascular Outcomes in Patients with Dysglycemia. *The New England journal of medicine*. 2012;367(4):309-318. doi:10.1056/NEJMoa1203859 [doi]
20. Sesso HD, Christen WG, Bubes V, et al. Multivitamins in the prevention of cardiovascular disease in men: the Physicians' Health Study II randomized controlled trial. *Jama*. 2012;308(17):1751-1760. doi:10.1001/jama.2012.14805 [doi]
21. Gaziano JM, Sesso HD, Christen WG, et al. Multivitamins in the prevention of cancer in men: the Physicians' Health Study II randomized controlled trial. *Jama*. 2012;308(18):1871-1880. doi:10.1001/jama.2012.14641 [doi]
22. Altenburg J, de Graaff CS, Stienstra Y, et al. Effect of azithromycin maintenance treatment on infectious exacerbations among patients with non-cystic fibrosis bronchiectasis: the BAT randomized controlled trial. *Jama*. 2013;309(12):1251-1259. doi:10.1001/jama.2013.1937 [doi]
23. Jamal SA, Hamilton CJ, Eastell R, Cummings SR. Effect of nitroglycerin ointment on bone density and strength in postmenopausal women: a randomized trial. *Jama*. 2011;305(8):800-807. doi:10.1001/jama.2011.176 [doi]
24. Castro M, King TS, Kunselman SJ, et al. Effect of vitamin D3 on asthma treatment failures in adults with symptomatic asthma and lower vitamin D levels: the VIDA randomized clinical trial. *Jama*. 2014;311(20):2083-2091. doi:10.1001/jama.2014.5052 [doi]
25. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *The New England journal of medicine*. 2016;375(4):311-322. doi:10.1056/NEJMoa1603827 [doi]
26. Laursen DRT, Paludan-Muller AS, Hrobjartsson A. Randomized clinical trials with run-in periods: frequency, characteristics and reporting. *Clin Epidemiol*. 2019;11:169-184. doi:10.2147/clep.s188752
27. Hopewell S, Hirst A, Collins GS, Mallett S, Yu LM, Altman DG. Reporting of participant flow diagrams in published reports of randomized trials. *Trials*. Dec 5 2011;12:253. doi:10.1186/1745-6215-12-253
28. Davis CE, Applegate WB, Gordon DJ, Curtis RC, McCormick M. An empirical evaluation of the placebo run-in. *Controlled clinical trials*. 1995;16(1):41-50. doi:019724569400027Z [pii]
29. Brittain E, Wittes J. The run-in period in clinical trials. The effect of misclassification on efficiency. *Controlled clinical trials*. 1990;11(5):327-338.
30. Investigators ET, Chertow GM, Block GA, et al. Effect of cinacalcet on cardiovascular disease in patients undergoing dialysis. *The New England journal of medicine*. 2012;367(26):2482-2494. doi:10.1056/NEJMoa1205624 [doi]
31. Baigent C, Landray MJ, Reith C, et al. The effects of lowering LDL cholesterol with simvastatin plus ezetimibe in patients with chronic kidney disease (Study of Heart and Renal Protection): a randomised placebo-controlled trial. *Lancet (London, England)*. 2011;377(9784):2181-2192. doi:10.1016/S0140-6736(11)60739-3 [doi]
32. Fralick M, Avorn J, Franklin JM, Abdurrob A, Kesselheim AS. Application and impact of run-in studies. *J Gen Intern Med*. May 2018;33(5):759-763. doi:10.1007/s11606-018-4344-7
33. Fralick M, Avorn J, Franklin JM, Bartsch E, Abdurrob A, Kesselheim AS. Application and Impact of Run-In Studies for the Evaluation of Statin Efficacy and Safety. *J Gen Intern Med*. Jun 2018;33(6):792-794. doi:10.1007/s11606-018-4349-2

34. Trivedi MH, Rush H. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*. 1994;11(1):33-43. doi:10.1038/npp.1994.63 [doi]
35. Bakris GL, Weir MR. Angiotensin-converting enzyme inhibitor-associated elevations in serum creatinine: is this a cause for concern? *Archives of Internal Medicine*. 2000;160(5):685-693.