A rapid review indicated higher recruitment rates in treatment trials than in prevention trials

Cindy L. Cooper, Daniel Hind, Rosie Duncan, Stephen Walters, Adjoa Lartey, Ellen Lee, Mike Bradburn

The School of Health and Related Research (ScHARR), Clinical Trials Research Unit, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK

Abstract

Objectives: To test the hypothesis that the percentage of patients screened that randomize differs between prevention and therapy trials.

Study Design and Setting: Rapid review of randomized controlled trials (RCTs) identified through published systematic reviews in August 2013. Individually randomized, parallel group controlled RCTs were eligible if they evaluated metformin monotherapy or exercise for the prevention or treatment of type 2 diabetes. Numbers of patients screened and randomized were extracted by a single reviewer. Percentages were calculated for each study for those randomized: as a function of those approached, screened, and eligible. Percentages (95% confidence intervals) from each individual study were weighted according to the denominator and pooled rates calculated. Statistical heterogeneity was assessed using $I^2$.

Results: The percentage of those screened who subsequently randomized was 6.2% (6.0%, 6.4%; 3 studies, $I^2 = 100.0$%) for metformin prevention trials; 50.7% (49.9%, 51.4%; 21 studies, $I^2 = 99.6$%) for metformin treatment trials; 4.8% (4.7%, 4.8%; 14 studies, $I^2 = 99.9$%) for exercise prevention trials; and 43.3% (42.6%, 43.9%; 28 studies, $I^2 = 99.8$%) for exercise treatment trials.

Conclusion: This study provides qualified support for the hypothesis that prevention trials recruit a smaller proportion of those screened than treatment trials. Statistical heterogeneity associated with pooled estimates and other study limitations is discussed. © 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

Keywords: Prevention; Treatment; RCTs; Recruitment rates; Exercise; Screening failures; Consent rates; Eligibility

1. Introduction

It is well documented that inadequate recruitment poses a threat to the successful completion of randomized controlled trials (RCTs) [1]. Excessive optimism about the number of potentially eligible candidates who are available, or will need to be approached for screening, is a key contributory factor; “Lasagna’s Law” [2,3] and “Muench’s Third Law” [4] state, with tongue partly in cheek, that “in order to be realistic, the number of cases promised in any clinical study must be divided by a factor of at least ten.” Some have proposed a corollary to these laws—that the percent yield of those screened or initially contacted is related to the restrictiveness of the research protocol’s eligibility criteria and the motivation of patients to enroll [5,6]. It follows that when estimating the availability of participants for a trial, we should apply eligibility criteria carefully to patient records and make cautious estimates for randomization rates, based on previous studies that are analogous in terms of their population, interventions, and research burden [6].

However, the reporting of randomization rates is still variable, meaning data to guide yield estimates are not always readily available [7]. In a widely cited reference text, Spilker and Cramer [5] proposed that we should expect “1 in 5 [20-27%] screened patients to enroll if the trial offers benefit for an active medical problem,” and, “1 in 40 [typically 1-6%]… if the trial offers the possibility of disease prevention.” Their sample was small and unsystematic, being based on a convenience sample of 10 prevention and 9 treatment trials. A more recent review of 280 highly cited treatment trials published between 2002 and 2010 reported a mean nonenrolment rate of 40.1% (standard deviation:
23.7% [7]. While encouraging, this proportion may not generalize to other settings, most notably prevention trials, which were not represented in the analysis set. Clearly, there is a need to better establish a realistic recruitment rate for prevention trials because this has profound implications for how we design, cost, and manage such work, given the effort required in screening for eligibility [8,9].

To investigate whether trials investigating disease prevention do indeed recruit smaller percentages of those screened, we undertook a rapid review of published RCTs evaluating metformin monotherapy or exercise (alone or in combination with other lifestyle interventions) for the prevention or treatment of type 2 diabetes (T2D). We chose this sample frame because each intervention can be used for either the prevention or treatment of T2D. We hoped that as a result, any comparison we made would be controlled for the comparative appeal to patients of an intervention and reliably investigate instead the comparative ease of recruitment. As we discuss in the following sections, there are a number of assumptions in this proposition that may be open to question.

2. Methods

2.1. Literature search

Two separate searches were conducted to identify Cochrane and other systematic reviews, which had already selected RCTs evaluating the use of metformin or exercise for the prevention of, or treatment for, T2D. We used “search all text” operations in the Cochrane Database of Systematic Reviews with no restrictions on publication date. The first search was conducted on August 16, 2013, and used the terms “diabetes” and “metformin.” The second was conducted on August 27, 2013, and used the terms “diabetes” and “exercise.” Titles and abstracts were screened by one researcher. Systematic reviews evaluating the use of either metformin or exercise, for either the prevention of or treatment for T2D, were included. The systematic review articles were obtained, and the trials that they had deemed eligible for inclusion were compiled so as to exclude any duplicates. The original research articles were obtained.

2.2. Study selection

Individually randomized, parallel group controlled trials were eligible for inclusion if they allocated to one arm either metformin monotherapy (insulin and additional dietary advice permitted) or exercise (including physical activity, advice on either exercise or physical activity, behavior change interventions, and supervised exercise). For the analysis of metformin, we excluded studies where metformin was used in combination with other pharmacotherapies. For the analysis of exercise, we excluded any study where the methods described the arms without mention of exercise. The stated reason for the intervention in any eligible trial had to be either prevention of diabetes or the treatment for T2D in adults. Studies that compared metformin monotherapy (as one arm) with exercise (as another arm) were included. We excluded preventive studies in which the intervention focused on the prevention of gestational diabetes and polycystic ovary syndrome (PCOS). We excluded therapeutic studies which recruited the following populations: type 1 diabetes, gestational diabetes, PCOS, or any other population which was not T2D. We excluded pediatric studies. We excluded cluster trials, crossover trials, and nonrandomized controlled studies. We excluded trials that were not published in English.

2.3. Data extraction

Each research article was read by the reviewer, and the data were extracted into a standardized Excel form, including the details about the population and intervention studied, whether a CONSORT diagram was published and the recruitment metrics from the eligibility criteria. The CONSORT statement proposes that researchers report how many people were assessed for eligibility and excluded based on ineligibility or refusal of consent; we recorded whether these variables were reported.

Where a CONSORT diagram was absent, attempts were made to extract data from the text. Where data were absent from the article, but a previous article relating to the study was cited, this article was retrieved and screened for data. Where this was not the case, or failed, authors were contacted to obtain missing data. Where author contact failed, we estimated the absolute numbers from percentages, where provided, rounding up to the nearest whole number.

2.4. Analysis

We defined those “approached” as those invited to screen or where screening was undertaken based on records, the number of records to which researchers attempted to
apply the eligibility criteria. We defined “screened” as any procedure applied to determine trial eligibility, either through interaction with an individual (eg, interviews) or their records (eg, chart reviews). As very few articles contained comprehensive data on how study candidates were screened, data on the character of screening were not extracted. We defined “eligible” as satisfying the trial inclusion and exclusion criteria. Crude rates were calculated for each study for those randomized (numerator) as a function of those approached, screened, and eligible (denominators). The overall percentages were derived by calculating the various rates from each individual study and weighting them according to the denominator [10]. Pooled rates, with 95% confidence intervals, were then calculated for each outcome. We used $I^2$ to measure the amount of between-study variation in conversion (recruitment) rates, which could not be explained by the play of chance alone (statistical heterogeneity). By convention, $I^2$ values of 25%, 50%, and 75% indicate low, moderate, and high levels of statistical heterogeneity [11].

3. Results

3.1. Systematic reviews

Fig. 1 presents a flow diagram of the study selection process. The searches initially identified 220 systematic reviews related to metformin and 183 related to exercise. After screening the title, abstract, and keywords for inclusion criteria, we included 10 systematic reviews about the use of metformin in the prevention ($n = 4$ [12–15]) and treatment ($n = 6$ [16–21]) of diabetes. We included 21 systematic reviews on the use of exercise in the prevention ($n = 6$ [22–27]) and treatment ($n = 15$ [28–42]) of diabetes. The median date of publication for included systematic reviews was 2009 (range 2001–2013).

3.2. Randomized controlled trials

The systematic reviews contained the following numbers of unique trials: metformin prevention, $n = 33$; metformin treatment, $n = 107$; exercise prevention, $n = 28$; and exercise treatment, $n = 146$. After data extraction, 6 preventative and 69 therapeutic RCTs evaluating metformin and 18 preventative and 87 therapeutic RCTs were retrievable, found to be eligible, and were included in the analysis (see Supplementary Tables 1–4 for eligible studies and Tables 5–8 for ineligible studies—Web only Appendix at www.jclinepi.com). There were no duplicate studies (those which had both exercise and metformin arms) in the analysis of treatment studies; there were two duplicate studies in the analysis of prevention studies (The Diabetes Prevention Programme, $n = 3,819$ [43,44] and Indian Diabetes Prevention Programme, $n = 531$ [45]). Across the whole study, 125 unique studies were excluded from the review because the full article was irretrievable, permanently or within the timescale of the study ($n = 41$); they were ineligible study designs ($n = 21$); they had the wrong intervention ($n = 31$); they had the wrong population ($n = 30$); they were not published in English ($n = 2$). The studies included in the review were published between 1985 and 2011. A small number of treatment studies were published before the initial CONSORT statement (1996 [46]—see Table 1), and the great majority of the exercise studies were published after the first revised statement (2001 [47]).

3.3. Data completion

Table 2 lists the variation in the reporting of recruitment metrics in the data set. Of the metformin studies, 3 of 6 prevention studies (50%) and 9 of 69 treatment studies (13%) provided adequate data to determine our primary outcome, the percentage of those randomized from those screened. Of the exercise studies, 14 of 18 prevention studies (78%)
and 29 of 87 treatment studies (33%) provided adequate data to determine the percentage of those randomized from those screened. Of 123 included studies published after 2001, when the revised CONSORT statement was published [47], the following numbers reported CONSORT-required items: flow diagram, 55 (45%); number screened, 60 (49%); number ineligible, 77 (63%); and number withholding consent, 46 (37%).

### 3.4. Conversion rates

Supplementary Tables 1–4 at www.jclinepi.com show the individual study data which contributed to the meta-analyses, summarized in Table 3 and Figs. 2–4. The randomization rate as a percentage of those approached was 2.6% (2.5%, 2.7%; 2 studies, $I^2 = 99.3\%$) for metformin prevention trials; 36.4% (35.0%, 37.9%; 2 studies, $I^2 = 99.8\%$) for metformin treatment trials; 1.9% (1.9%, 2.0%; 8 studies, $I^2 = 99.8\%$) for exercise prevention trials; and 16.4% (16.1%, 16.8%; 10 studies, $I^2 = 99.3\%$) for exercise treatment trials (Fig. 2). The primary outcome, randomization rate as a percentage of those screened, was 6.2% (6.0%, 6.4%; 3 studies, $I^2 = 100.0\%$) for metformin prevention trials; 50.7% (49.9%, 51.4%; 21 studies, $I^2 = 99.6\%$) for metformin treatment trials; 4.8% (4.7%, 4.8%; 14 studies, $I^2 = 99.9\%$) for exercise prevention trials; and 43.3% (42.6%, 43.9%; 28 studies, $I^2 = 99.8\%$) for exercise treatment trials (Fig. 3). The randomization rate as a percentage of those eligible was 48.6% (47.6%, 49.6%; 3 studies, $I^2 = 98.3\%$) for metformin prevention trials; 75.6% (74.2%, 76.9%; 9 studies, $I^2 = 99.8\%$) for metformin treatment trials; 68.5% (67.9%, 69.1%; 12 studies, $I^2 = 99.8\%$) for exercise prevention trials; and 70.4% (69.7%, 71.2%; 29 studies, $I^2 = 99.8\%$) for exercise treatment trials (Fig. 4).

### 4. Discussion

The percentage of people randomized from those screened averaged 6% and 5% in prevention studies evaluating metformin monotherapy and exercise, respectively, but 51% and 43% in studies evaluating the same interventions for treatment. There were very high levels of statistical heterogeneity associated with all pooled estimates, and this, together with limitations in the review methods and conduct, implies that findings must be treated cautiously. Nonetheless, the magnitude of this difference does suggest that recruitment into trials evaluating preventive and therapeutic interventions is different.

This work aimed to test a hypothesis, generated over twenty years ago, about the relative difficulty of recruiting to prevention and treatment trials [5]. We felt it was important to test this hypothesis as our own experience tells us that many funders and researchers expect similar randomization rates across the two study categories. Our work is the most systematic evaluation of this question of which we are aware and will raise awareness that researchers should not typically expect to achieve in prevention studies the high randomization rates identified as typical by Humphreys et al. [7] in contemporary treatment trials. Our work also recognizes some potential sources of variation by evaluating randomization rates for both drug and nondrug interventions in comparable populations. Future studies should aim to control for this source of potential variation where possible. However, our review has a number of limitations.

Rapid reviews aim for explicit and rigorous method but with concessions on the gold standard systematic review process, usually necessitated by resource constraints. Such concessions, such as the use of simpler search strategies used in our project, are considered legitimate by guidelines, but other sources of potential bias are more problematic [48]. For instance, we used a single reviewer to select and extract data from studies, with recourse to other team members only when perceived problems arose. Using two reviewers reduces the possibility that relevant reports are discarded and results in fewer errors than when selection

### Table 1. Distribution of included studies by date, relative to CONSORT publications

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Metformin as treatment $n = 69$</td>
<td>14</td>
<td>22</td>
<td>33</td>
<td>1986</td>
<td>2011</td>
</tr>
<tr>
<td>Metformin as prevention $n = 6$</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>1998</td>
<td>2005</td>
</tr>
<tr>
<td>Exercise as treatment $n = 87$</td>
<td>6</td>
<td>8</td>
<td>73</td>
<td>1985</td>
<td>2011</td>
</tr>
<tr>
<td>Exercise as prevention $n = 18$</td>
<td>0</td>
<td>4</td>
<td>14</td>
<td>1997</td>
<td>2010</td>
</tr>
</tbody>
</table>

[^1]: Indicates CONSORT recommended metric.

[^2]: Indicates $n$ or % published from 2002.

### Table 2. Completeness of the data set

<table>
<thead>
<tr>
<th>Presence of...</th>
<th>Metformin</th>
<th>Exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSORT diagram</td>
<td>Prevention, % $[N = 6 (3)]$</td>
<td>Treatment, % $[N = 69 (33)]$</td>
</tr>
<tr>
<td>Number approached</td>
<td>33 (67)</td>
<td>22 (67)</td>
</tr>
<tr>
<td>Number screened$^a$</td>
<td>33 (67)</td>
<td>3 (9)</td>
</tr>
<tr>
<td>Number eligible$^a$</td>
<td>50 (67)</td>
<td>13 (39)</td>
</tr>
<tr>
<td>Number refused consent$^a$</td>
<td>0 (0)</td>
<td>10 (30)</td>
</tr>
<tr>
<td>Number randomized$^a$</td>
<td>100 (100)</td>
<td>100 (100)</td>
</tr>
</tbody>
</table>
and data extraction are performed by a single author [49,50]. By combining the eligible RCTs of several overlapping systematic reviews, it is unlikely that we failed to identify significant numbers of eligible studies. However, as we have documented, the need to truncate the review process, together with the inaccessibility of some trial reports, meant that 41 studies were irretrievable by the time of analysis, introducing unquantifiable bias and undermining our findings. Most of these reports (33 of 41) related to the exercise treatment analysis, meaning this aspect of the study is at greater risk of bias. In addition, where primary research studies were retrieved and eligible, the quality of reporting was poor; even trials published after the revised CONSORT statement of 2001 [47], frequently omitted recommended items, such as the numbers of study candidates assessed for eligibility, excluded on eligibility grounds and refusing consent.

In testing the hypothesis that fewer of those screened will randomize in prevention studies than in treatment studies, we have taken a crude approach, categorizing together studies with often striking methodological differences. For instance, the presence of an active control is thought to increase recruitment [51]. The restrictiveness of an RCT’s eligibility criteria is widely thought to inhibit accrual as a function of those screened [2,6,8,52]. And, targeted mass mail outs based on database searches are thought to be a more efficient method of recruitment than opportunistic approaches during medical consultations for individually randomized prevention trials in chronic conditions [53]. We did not collect data on any of these or many other variables, which might conceivably affect recruitment rates [52]. Additionally, we have standardized study metrics in such a way that we could combine trials with a simple, single-stage screening process with others that have complex multistage processes. It is thought that those with multistage screening tend to lose larger numbers overall, with candidates withdrawing between each successive screen [8]. In our analysis, we have also combined studies with different pathways to randomization. The basic assumption is that recruitment happens by approaching candidates before screening, after which eligibility is established: out of choice or necessity, the researchers approach people first before inviting them for screening, as in The Indian Diabetes Prevention Programme [45]. On the other hand, some studies use routinely collected clinical data to screen for eligibility before approaching people to consider study participation, as in the Västerbotten Intervention Programme [54]. Not all those screened as eligible by a study team using their records, without their direct involvement, will always be approached. There were no instances in our sample where researchers reported both numbers approached and numbers screened and where numbers screened were larger than numbers approached. Nonetheless, the variation in the patient selection pathway is one likely source of heterogeneity. CONSORT 2010 does not mandate detailed information about the screening phase so it is unlikely that reporting of this data will improve in the near term. This means that published reports may not

<table>
<thead>
<tr>
<th>Table 3. Summary of conversion rates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Randomized as a percentage of:</strong></td>
</tr>
<tr>
<td>Those screened (estimate by Spilker and Cramer [5])</td>
</tr>
<tr>
<td><strong>This study</strong></td>
</tr>
<tr>
<td>Those approached (%)</td>
</tr>
<tr>
<td>Those screened (%)</td>
</tr>
<tr>
<td>Those eligible (%)</td>
</tr>
</tbody>
</table>

Fig. 2. Randomization rate as a percentage of those approached.

Fig. 3. Randomization rate as a percentage of those screened (primary outcome). Red vertical lines indicate the percentages expected by Spilker and Cramer [5] (2.5% and 20%).
provide a sound basis for recruitment projection. Those seeking to improve their reporting should consider the CONSORT guidance (those screened and numbers ineligible and withholding consent put in the flow diagram) as a minimum standard for the flow diagram. CONSORT 2010 notes that “it will sometimes be useful or necessary to adapt the structure of the flow diagram to a particular trial” [55]. This means researchers may, and we would encourage them to, describe in detail the methods and results of multistage screening processes in the flow diagram, where the information will not contribute to a manuscript’s word count.

In planning this review, we made certain arbitrary choices in how to set its scope. It is likely that different recruitment dynamics may be observed with the use of different populations (pediatric rather than adult; PCOS vs. T2D), interventions (combination rather than monotherapy), or study designs (crossover or cluster rather than individually randomized). Perhaps, implicit in the statement by Spilker and Cramer’s [5] aforementioned is the idea, also expressed by others, that the motivation for participation in prevention trials is different to that in treatment trials. Spilker and Cramer are not alone in assuming that participation in treatment trials is often driven by a desire for symptom relief, whereas if there is any clinical reward for an essentially healthy person’s participation in prevention trials, it is comparatively distant in time [56,57]. This is a rather simplistic view, and the categories may not always be rigidly distinct. In those studies included in a systematic review [58] from which adequate data could be derived, four trials evaluating chemopreventive agents for the prevention of colorectal cancer in high-risk populations recruited significantly higher percentages (median 71%; range 65% to 89%) of those screened than eight trials in the general population (median 13%; range 2% to 61%) [59]. Although this reflects a range of factors, it might indicate that the perceived risk of a hazard emerging in the future increases the chances of an individual recruiting to a prevention trial. Our prevention studies focus on prediabetes, a population which is arguably at high risk of developing a chronic health condition and already effectively medicalized [60]. It follows that the recruitment rates observed in our study may not be transferrable to studies recruiting people in the general population, or in at-risk populations, where the likelihood or potential impact of a hazard is taken less seriously by members of that population. Further research may usefully investigate whether primary prevention trials with better recruitment rates involve target populations with more severe risk factors.

Although we were testing a specific published hypothesis [5], it is questionable whether those randomized as a percentage of those screened is the most important statistic for all researchers to consider. Arguably, the randomization rate as a function of those eligible has more utility, especially if the concern is motivation to recruit. In particular, some prevention trials target individuals who are not routinely identified by the health system. As a result, the research team will have to screen larger volumes of people to establish a pool of eligible study candidates than in a treatment trial, where the broad diagnosis is already known, even if eligibility criteria may cause attrition at the margins. The reader will note from Table 3 that differences in those randomized as a function of those eligible are not as marked between prevention and treatment trials as the differences observed in this study’s other outcomes. This observation might undermine any implication, which it is not clear whether Spilker and Cramer intended, treatment trials have higher enrollment rates because “the trial offers benefit for an active medical problem” [5]. However, the most interesting to us about the hypothesis was not the question of motivation, but the resource implications that stem from the necessity to approach and screen larger numbers of candidates.

5. Conclusion

Larger prevention studies are unlikely to randomize as many of those screened as in a typical treatment trial. Those planning prevention trials should be aware of other sources of variation in randomization rates, including the design of the patient selection pathway and the perceived impact of the risk factors, which make the target population eligible for the trial and the value they place on disease prevention. Future research might seek to replicate this work in other populations and settings and collect more data on design and other covariates.

Acknowledgments

The authors acknowledge and thank Ed Clein and Helen Wakefield for their support for this work. The study was not funded by any specific grant. The authors also thank Janet Blodgett of the US Veterans Administration and Keith
Humphreys of Stanford University for the provision of unpublished data. The authors gratefully acknowledge the contribution of two anonymous peer reviewers to the improvement of the article.

Authors’ contributions: All authors contributed to the study. D.H. and C.L.C. contributed to the design of the study and interpretation of the findings. A.L. undertook preliminary work which informed the design of the study, as part of her Masters dissertation. R.D. retrieved the articles, extracted, and collated the data. D.H. and C.L.C. checked assumptions and decisions made on inclusion of articles. S.W. advised on statistical analysis. M.B. advised on analysis and presentation of the results. M.B. and E.L. undertook the data analysis. D.H. wrote the first draft of the article. All authors contributed to critically revising drafts of the article. All authors have approved the final manuscript. D.H. had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Supplementary data
Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2014.10.007.

References