Recruitment failure in randomised controlled trials: a conundrum

Rikken JFW¹, Casteleijn RN¹, van der Weide MC¹, Ruben Duijnhoven¹, Mariëtte Goddijn¹, Ben Mol², van der Veen F¹, and Madelon van Wely¹

¹Amsterdam UMC Locatie AMC
²Monash University Monash eResearch Centre

March 16, 2024

Abstract

Objective: We aim to assess which variables are associated with recruitment failure of RCTs, leading to an extension of the study period.

Design: Nationwide cohort study.

Setting: A cohort of RCTs supported by the trial centre of the Dutch Consortium of Obstetrics and Gynaecology.

Population: We included 83 RCTs that recruited patients between March 1st 2003 and December 1st 2023.

Main outcome measures: Primary outcome was recruitment target not achieved within six months after the pre-planned recruitment period. Secondary outcomes were recruitment target not achieved within an extension period of at least twelve months and premature termination of the trial.

Results: In total, 46 of 83 RCTs (55%) did not achieve their targeted recruitment within the pre-planned study period with a maximal extension period of 6 months. The most relevant indicators for recruitment failure in multivariable risk prediction modelling were presence of a no-treatment arm, a compensation fee of less than 200 euros per included patient, funding of less than 350,000 euros, while a preceding pilot study lowered this risk.

Conclusions: We identified that the presence of a no-treatment arm, low funding and a low compensation fee per included patient were the most relevant risk factors for recruitment failure within the pre-planned period, while a preceding pilot study lowered this risk. Awareness of these indicators is important when designing future studies.

Introduction

Randomised controlled trials (RCTs) are considered to be the best strategy in evaluating the effectiveness of medical interventions and they maintain a dominant position in the hierarchy of medical evidence(1). RCT outcomes are most often adopted into (inter) national clinical guidelines and have great influence on daily routine clinical practice. Unfortunately, obtaining evidence from RCTs is often hampered by failure to recruit enough patients within the pre-planned study period, leading to premature termination of the trial or extension of the study period(2).

Premature termination due to poor recruitment has been estimated to occur in 9-10% of all RCTs(3-5). Variables that have been associated with poor recruitment are an overestimation of the number of eligible patients, a preference for one of the interventions by the patients, a high burden of the tested intervention for the patients, an unclear trial design, strict eligibility criteria, a lack of logistic support or a lack of funding(6-9).

While the variables that may result in poor recruitment leading to premature termination of the trial are well known, much less is known on variables related to recruitment failure within the pre-planned study period, leading to extension of the study period.

The one study to investigate this matter, explored factors associated with recruitment in a cohort of 114 multicentre RCTs in more than nine clinical areas, including cancer, cardiology and obstetrics & gynaecology (18 RCTs had a clinical area classified as ‘other’), and funded by two public bodies in the United Kingdom; the UK Medical Research Council (MRC) and the Health Technology Assessment (HTA) Programme(5).
RCTs that were funded by the MRC (as compared with the HTA) and were in the clinical area ‘cancer’, had better chances of good recruitment, which was a marginally statistically significant association. The vast heterogeneity of RCTs included in that study hampered the identification of other indicators associated with poor recruitment and did not allow the authors to provide useful advice for improvement.

A longer recruitment period may result in a shortage of resources possibly impacting the quality of the trial, limit the institutional capacity to start new RCTs, result in a trial that tries to answer a question that is no longer relevant, or result in premature termination of the study, thus hindering a conclusion with sufficient statistical power(10).

To assess factors that are associated with recruitment failure within the pre-planned study period, we performed a nationwide cohort study of RCTs within the setting of the Dutch Consortium of Obstetrics & Gynaecology in the Netherlands. Such knowledge may be instrumental in helping researchers, trial centres and funding agencies to prevent this type of recruitment failure.

Methods

Study design

This study was designed as a nationwide cohort study and included all multicentre RCTs carried out within the Dutch Consortium for Women’s Health Research, embedded within the professional society, i.e. Dutch Society of Obstetrics and Gynaecology (NVOG)(11). The Dutch Consortium for Women’s Health Research facilitated studies in obstetrics, gynaecology and reproductive medicine.

Within the consortium, participating clinical centres are both academic and non-academic hospitals. RCTs conducted within the Consortium are supported by a clinical trial centre (https://zorgevaluatiенederland.nl/), a multidisciplinary trial bureau with methodologists, data managers, contract managers and trial managers. The trial centre staff supports research groups by advising on the budget, logistics, methods, and ethics approval, developing electronic case record forms, performing contract management and monitoring, creating the interim reports for the data safety and monitoring board and providing advice on the statistical analyses. We constructed the manuscript according to the STROBE guideline(12).

Study population

We included finalized multicentre RCTs supported by the clinical trial centre and performed within the Dutch Consortium for Women’s Health Research, between March 1st 2003 and December 1st 2023. We excluded studies with an observational design, single centre RCTs, RCTs initiated outside the Netherlands, RCTs with a cluster or parallel study design, RCTs that never actually started, RCTs in which inclusion of patients was still ongoing and RCTs prematurely discontinued for other reasons than poor recruitment, for example due to safety issues after an interim analysis.

Outcome measures

Primary outcome was recruitment target not achieved within 6 months after the pre-planned recruitment period. These RCTs were defined as RCTs with recruitment failure. The pre-planned recruitment period was documented by the principal investigator before the start of the trial. Secondary outcomes included recruitment target not achieved within an extension period of at least 12 months and premature termination of the trial (defined as stopping with including patients before the recruitment target was achieved). All studies that recruited during the COVID-19 pandemic received 6 months extension of their recruitment period.

In all RCTs, we collected information on variables with a potential effect on recruitment failure, identified after a scoping review. We recorded variables at five levels; patient, doctor, participating centre, study organisation and study design (Appendix 1).

Statistical analysis
For the primary outcome, we used the planned recruitment period as documented in the General Assessment and Registration form, a form that needs to be submitted to the ethical committee before actual start of the study. If we could not get access to this form, we retrieved this information from the main investigator and/or used the data mentioned in the protocol of the study. The actual recruitment period was calculated as the time between the first and last inclusion date.

We checked the continuous potential indicators with spline curve analysis. We dichotomised on basis of the spline curve and used the median when the spline suggested a straight line. We used logistic regression to evaluate the association between potential indicators of recruitment failure and expressed these as odds ratios (OR) with corresponding 95% confidence intervals (CI).

To further explore the most relevant risk factors for recruitment failure multivariable risk prediction modelling was done by using both forward and backward stepwise logistic regression (entry p=0.2 and exclusion p=0.1).

We used SPSS® (IBM 2019, USA) software for all statistical analyses (version 25).

Ethics approval

Our study focussed on logistics and design issues and did not include patients as study participants. Consequently, we did not need ethical approval for this study.

Transparency statement

All authors had full access to all the data in the study and the corresponding author had final responsibility for the decision to submit for publication. The manuscript is an honest, accurate and transparent account of the study being reported, no important aspects of the study have been omitted, and any discrepancies from the study as originally planned have been explained.

Role of the funding source

This study was supported by a small departmental grant of the Centre for Reproductive Medicine, Amsterdam University Medical Centres, location AMC.

Public and patient involvement

No patients or members of the public were involved in this study since the study did not concern patients directly.

Results

Between March 1st 2003 and December 1st 2023 189 studies started recruitment and were assessed for eligibility. Of these, 106 studies did not fulfil our inclusion criteria, such that in total 83 RCTs were included in the analyses (Figure 1). Characteristics of the included studies are summarized in Table 1. Fifteen RCTs did not have funding at all (18%). A more detailed list of all RCTs can be found as supplementary file Appendix 2(13-89).

Primary and secondary outcomes

In total, 46 of 83 RCTs (55%) did not achieve their targeted recruitment within the pre-planned study period with a maximal extension period of 6 months (Table 2). Recruitment was not achieved within the pre-planned study period with a maximal extension period of 12 months in 41 RCTs (49%). Of these 41 RCTs, 29 studies had a total recruitment period of up to five years, and 12 RCTs finished their recruitment within five to ten years.

Nineteen RCTs (23%) stopped prematurely due to recruitment issues. Of these 19 RCTs, four studies reached 0 to 10% of their recruitment target, six studies 10 to 20%, two studies 20 to 30%, five studies 30 to 60% and two studies reached 70 to 80% of their planned recruitment target.

The mean recruitment period was 50 months (range 12-96 months) for RCTs with recruitment failure versus 31 months (range 12-91 months) for RCTs without recruitment failure. Twenty-two RCTs had a recruitment
period of over 48 months. The actual absolute recruitment rate was 4.5 inclusions per month in RCTs with recruitment failure compared to 18.5 inclusions per month in RCTs without recruitment failure (p<0.001).

**Potential indicators of recruitment failure**

The association of the potential indicators with RCTs with recruitment failure i.e. RCTs that did not achieve their recruitment target within the pre-planned study period with a maximal extension period of 6 months, is shown in Table 3.

Indicators associated with higher chances on recruitment failure were presence of a no-treatment arm, having a design with more than two arms, funding, a compensation fee of less than 200 euros per included patient, funding of less than 350,000 euros and having more than four inclusion criteria. An indicator associated with lower chances on recruitment failure was a preceding pilot study. The most relevant indicators for recruitment failure in multivariable risk prediction modelling were presence of a no-treatment arm (OR 4.95, 95% CI 1.18 to 20.80), a compensation fee of less than 200 euros per included patient (OR 2.90, 95% CI 1.02 to 8.25), funding of less than 350,000 euros (OR 2.99, 95% CI 1.05 to 8.51), while a preceding pilot study lowered the risk for treatment failure (OR 0.21, 95% CI 0.05 to 0.83).

When we compared the 41 RCTs that did not achieve their recruitment target within the pre-planned study period with a maximal extension period of 12 months, with the 42 RCTs that completed recruitment within that period, the described associations with treatment failure remained comparable in direction and size.

The most relevant indicators for stopping prematurely were the absence of a preceding pilot study and having a no-treatment arm. None of the 19 RCTs that stopped prematurely had performed a pilot study (0%), compared to 17 of the 62 RCTs that completed recruitment (27%). Ten of the 19 RCTs that stopped prematurely had a no-treatment arm (52%), compared to eight of the 64 RCTs that completed recruitment (12.5%) (OR 6.13, 95% CI 1.98 to 19.06).

**Discussion**

**Main findings**

In this nationwide cohort study, 46 of 83 included RCTs (55%) did not achieve their recruitment target within the pre-planned study period with a maximal extension period of six months. RCTs that had a no-treatment arm, low funding and low financial compensation per included patient were at risk to experience this type of recruitment failure, while a preceding pilot study lowered this risk. Upon extension of the pre-planned study period from six to twelve months, 41 RCTs (49%) still did not achieve the pre-planned recruitment target. Nineteen RCTs (23%) were stopped prematurely because of recruitment issues.

**Strengths and limitations**

Our study has a number of strengths. First, we investigated recruitment failure in 83 RCTs embedded within the Dutch Consortium for Women’s Health Research – and thus within one discipline - with support and monitoring by the clinical trial centre. This allowed us to standardize several important aspects, like trial management and logistics, data collection and data monitoring. Second, we were able to assess all indicators with a potential association with poor recruitment as described in literature; type of investigation, placebo-controlled study, treatment versus no treatment, whether the intervention was new or only available in the trial, whether the study was blinded or if there were any competing RCTs, number of study arms, number of inclusion and exclusion criteria, whether a pilot study was performed, number of participating centres and funding and compensation per included patient.

The main limitation of our study is the number of trials. Obviously, if we could have accessed an even larger cohort of trials, we might have been able to identify more potential indicators for recruitment failure. A further limitation may be that within our study we focussed on objective indicators, such as trial logistics and design issues. Other aspects, like patients’ or practitioners’ perspectives, which may affect recruitment as well were beyond the scope of our study.
Interpretation

The design of a no-treatment arm where treatment is standard clinical practice was associated with recruitment failure. This design is particularly relevant, since we may be over-treating patients while we are actually in equipoise on whether the intervention is effective at all. Possibly, in this design specifically, the preference of the doctor or patient might play a role in the laborious recruitment. A no treatment arm was also associated with stopping prematurely, supporting its relevance as a risk factor. In our study ten (52%) of 19 RCTs that stopped prematurely had a no-treatment arm where in current clinical practice treatment is expected.

Two typical examples of RCTs with such a design that stopped prematurely were a trial that compared intrauterine insemination (IUI) with expectant management in couples with unexplained subfertility, and a trial that compared immediate delivery with temporizing management in women between 27+5 and 33+5 weeks of gestation admitted for early-onset severe preeclampsia with or without HELLP syndrome(33, 81).

Not very surprisingly, the lack of funding and compensation fee per included patient was associated with recruitment failure. Twelve studies with recruitment failure had no funding at all, compared with three studies without recruitment failure. In combination with our outcome that extending the recruitment period from six to twelve months did not increase the numbers of RCTs that reached their pre-planned sample size, this has important clinical, logistic and financial consequences. RCTs may reach their recruitment target, but in 12 RCTs in our study, recruitment took up to ten years. It implies that when recruitment is doomed to fail, it may reach its required sample size in the end, but at the expense of a lot of endurance and extra funding by a willing sponsor. On the other hand, RCTs can still be of extreme clinical importance if the research question is – and remains – relevant. This is shown by a trial that investigated low-molecular-weight heparin in women with recurrent pregnancy loss and inherited thrombophilia, which took 7,5 years to recruit, but results were eagerly awaited and eventually published in a high impact journal(15).

A preceding pilot study lowers recruitment failure, while a study design with more than two arms or more than four inclusion criteria might increase the chance of recruitment failure, although with a wide confidence interval due to small numbers. We think that a preceding pilot study helps to notice and resolve potential issues before start of the actual study, while a study design with more than two arms or more than four inclusion criteria could result in an overly complex recruitment process. In a review of the literature on factors limiting the quality and progress of RCTs not hampered by recruitment failure, a straightforward study protocol and data collection as well as careful planning were also identified as key factors for completion(90).

A competing study was not associated with a lower chance on recruitment failure, which is the opposite of what we expected. We hypothesize that when more RCTs in the same field are recruiting patients at the same time, clinicians are more aware of the possibility of including patients in a particular RCT, or when one RCT recruits rapidly, this might be “contagious” for the other RCTs.

It is important to note that our results should not withhold clinicians from conducting RCTs on these research questions. Investigating the efficacy and safety of treatments and providing robust evidence can be of the utmost importance. Although it is known that the results of randomized and nonrandomized studies have a good correlation, nonrandomized studies tend to show larger treatment effects, and thus observational studies can be good adjunct to RCTs, but they cannot replace them(91, 92). More importantly, our study shows that also RCTs with recruitment that takes many years answer highly relevant clinical questions and can truly make a big difference in the clinical field. Principal investigators, sponsors and all who are participating in an RCT should be aware of the indicators associated with poor recruitment, and that with dedication and persistence the RCT could be successfully completed and published.

Conclusion

To conclude, RCTs with a no-treatment arm, low funding, low financial compensation per included patient are more likely to experience recruitment failure, while a preceding pilot study lowers this chance. We propose that investigators and grant providers consider these issues before the actual start of the study, to improve
the chances of recruitment success. If a relevant trial is destined to have a suspected long recruitment period, it seems wise to ponder on the question whether to start the trial, or to accept a longer recruitment period with all its consequences.

**Acknowledgements:**

The authors would like to thank all the clinical and principal investigators of the included RCTs for cooperating in data collection, and Maya Kruijt, policy advisor Zorgevaluatie Nederland, for her advice.

**Disclosure of interest:**

MG reports department research and educational grants from Guerbet and Ferring (location VUmc) outside the submitted work. BWM reports grants from NHMRC, personal fees from ObsEva, personal fees from Merck KGaA, personal fees from Guerbet, personal fees from iGenomix, outside the submitted work.

**Contribution to Authorship:**

JFWR, MvW, MCW and RGD conceived the study. JFWR and RC did the scope review, selected the potential indicators, and collected the data. Differences of opinion and questions regarding the data were resolved with MvW. JFWR was responsible for the data. JFWR, RC and MvW analysed the data. JFWR, MvW, MG and FvdV drafted the manuscript, supported by BWM. All authors contributed to the critical revision of the paper and approved the final manuscript.

**Reference list:**


67. Detollenaere RJ, den Boon J, Stekelenburg J, IntHout J, Vierhout ME, Kluivers KB, van Eijndhoven HW. Sacrospinous hysteropexy versus vaginal hysterectomy with suspension of the uterosacral ligaments in
women with uterine prolapse stage 2 or higher: multicentre randomised non-inferiority trial. BMJ (Clinical research ed). 2015;351:h3717.


71. van de Laar R, Kruitwagen RF, Zusterzeel PL, Van Gorp T, Massuger LF. Correspondence: Premature Stop of the SOCceR Trial, a Multicenter Randomized Controlled Trial on Secondary Cytoreductive Surgery: Netherlands Trial Register Number: NTR3337. Int J Gynecol Cancer. 2017;27(1):2.


Table 2.docx available at [https://authorea.com/users/756128/articles/726997-recruitment-failure-in-randomised-controlled-trials-a-conundrum](https://authorea.com/users/756128/articles/726997-recruitment-failure-in-randomised-controlled-trials-a-conundrum)

Hosted file


Hosted file


Hosted file