Journal of Clinical Epidemiology 124 (2020) 58–68

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Statistical models to predict recruitment in clinical trials were rarely used by statisticians in UK and European networks

Efstathia Gkioni[a,b,*], Susanna Dodd[a], Roser Rius[c], Carrol Gamble[a]

[a]*Department of Biostatistics, University of Liverpool, A Member of Liverpool Health Partners, Liverpool, UK*
[b]*Université de Paris, CRESS, INSERM, INRA, F-75004 Paris, France*
[c]*Department of Statistics and Operations Research, School of Mathematics and Statistics, BarcelonaTech (UPC), Barcelona, Spain*

Accepted 23 March 2020; Published online 27 March 2020

Abstract

Objective: Identify the current practice for recruitment prediction and monitoring within clinical trials.

Study Design and Setting: Chief investigators (CIs) were surveyed to identify data sources and adjustments made to support recruitment prediction. Statisticians were surveyed to determine methods and adjustments used when predicting and monitoring recruitment. Participants were identified from the National Institute for Health Research recently funded studies, the UK Clinical Research Collaboration registered Clinical Trial Units network or by the European Clinical Research Infrastructure Network.

Results: A total of 51 CIs (UK = 32, ECRIN = 19) and 104 statisticians (UK = 51, ECRIN = 53) were contacted. Response rates varied (CIs UK = 53% ECRIN = 32%; statisticians UK = 98% ECRIN = 36%).

Multiple data sources are used to support recruitment rates, most commonly audit data from multiple sites. Variation in individual site recruitment rates are frequently incorporated, but staggered site openings were featured more commonly among UK respondents. Simple prediction methods are preferred to rarely used statistical models. Lack of familiarity with statistical methods are barriers to their use with evidence needed to justify the time required to support their implementation.

Conclusion: Simplistic methods will continue as the mainstay of prediction; however, generation of evidence supporting the benefits of complex statistical models should promote their implementations. Multiple data sources to support recruitment prediction are being used, and further work on the quality of these data is needed. Pressure to be optimistic about recruitment rates for the trial to be attractive to funders was felt by a sizable minority.

Keywords: Recruitment; Prediction; Monitoring; Surveys; Clinical trials

1. Introduction

Clinical trials are a major financial investment with the time to recruit to the required sample size being a key driver of associated costs. Failure to successfully recruit clinical trial participants as predicted at the design stage has many negative consequences. These range from incurring increased costs and time to answer the clinical question of interest to abandoning the research with the question remaining unanswered. There may also be negative impact on the planning and roll out of future research.

Despite investment aimed at reducing difficulties in recruitment to clinical trials [1,2], there has been no demonstrable improvement and 45% of trials supported by two prestigious UK funding bodies, Health Technology Assessment and Medical Research Council (HTA & MRC), continue to fail to meet their original recruitment targets [3].

Authors' contributions: All authors have read and approved the final version of this manuscript. EG was one of the lead researchers on this project and was responsible for the protocol preparation, creation of surveys, data collection and analysis, and writing of this manuscript. SD is coinvestigator on this project and contributed to its design, creation of surveys, analysis, writing, and proofreading of this manuscript. RR is also coinvestigator and contributed to the creation of surveys and analysis of the results. CG is the principal investigator and is responsible for the project conception, design, creation of surveys, analysis, writing, and proofreading of this manuscript.

* Corresponding author. Department of Biostatistics, Institute of Translational Medicine, University of Liverpool, Block F/ Waterhouse Building, 1-5 Brownlow Street, Liverpool L69 3GL, UK. Tel: +44 (0) 151 794 9743.

E-mail address: e.gkioni@liverpool.ac.uk (E. Gkioni).

https://doi.org/10.1016/j.jclinepi.2020.03.012

What is new?

Key findings

•Statistical models to predict recruitment are rarely used. Uptake is limited by the absence of evidence regarding their benefit over more simplistic approaches. Predictions allowed for variation in site recruitment rates and staggered site openings, whereas seasonal variation and holiday periods were less frequently considered.

•Respondents in European network had greater awareness of statistical methods to predict recruitment in comparison with the UK respondents; however, numbers are small, and the European sample is subject to potential response bias.

•Approximately, one-third of UK and half of the European Clinical Research Infrastructure Network (ECRIN) respondents in the Chief Investigators' survey reported a need to be optimistic about the predicted recruitment rate, to be attractive to the funder.

What this adds to what is known?

•Prediction of recruitment in clinical trials continues to represent a formidable challenge. This survey identifies data sources and factors used to adjust recruitment rates. It identifies that simplistic approaches to predict recruitment are favored over more complex statistical models. Barriers to uptake of the statistical methods include complexity of their implementation and absence of evidence that the time taken to implement them will result in improving the accuracy of recruitment prediction.

What is the implication and what should change now?

•Evidence demonstrating superiority of statistical methods over simplistic methods needs to be established on a prospective cohort of trials. Subsequently, relevant software and training courses should be made available. Multiple data sources are being used, and further work on the quality of this data is needed. Balance of expectations between funders and applicants needs further exploration.

A prioritization exercise was undertaken to identify uncertainties related to trial recruitment as a focus for future methodological research [4]. Recruitment prediction was identified as a top 10 priority area. Despite the fact that every clinical trial will require such predictions to be made, little is known about how this is achieved either in terms of data sources or methods used. This is unlikely to change given that reporting requirements for recruitment within a main clinical trial article are minimal [5] and grant applications are generally not publicly available.

To determine current practice within the United Kingdom and Europe, we undertook a survey of Chief Investigators (CIs) and a survey of statisticians across a UK and a European network.

2. Methods

2.1. Design

EG led the design of each survey with input from all coauthors. Questions targeted data sources and methods used for recruitment prediction, identifying team members contributing to the process, and awareness and implementation of the statistical models (statisticians survey only). Multiple choice answers were informed by relevant publications about statistical models [6,7] and other approaches that could be used for recruitment prediction and monitoring, as well as by a number of factors that impact recruitment rate that should be considered [8–11].

2.1.1. Chief investigators' survey

The CIs' survey targeted those collaborating within UK and European research infrastructures. The survey aimed to be brief to maximize return rates collecting information not available from publicly available sources covering data sources used to predict recruitment and how these were applied to trial and site requirements. The survey was reviewed within the study team before circulation across the UK and the European Clinical Research Infrastructure Network (ECRIN, https://www.ecrin.org/).

A prize equivalent to £75 in vouchers was offered as an incentive to participation. The full list of survey questions and the invitation email are provided in the Additional file 1.

2.1.1.1. UK chief investigators. UK CIs of recently funded clinical trials were surveyed as identified from the National Institute for Health Research (NIHR) journals library (https://www.journalslibrary.nihr.ac.uk/programmes/hta/, searched in May and October 2018). The NIHR is the largest funder of health and care research within the United Kingdom. To be eligible for inclusion projects were required to be randomized with the trial status listed as ''waiting to start''. CI details were obtained from the projects' website and each contacted by EG via email, containing an invitation to participate and the survey attached as a Word document. CIs were also given the option of delegating completion to a trial team member. If no response was obtained within 2 weeks then a reminder email was sent and a further reminder followed 3 weeks later, within which we gave them the option to answer the survey via a phone call. A final attempt was made to contact nonrespondents by phone to ensure the correct contact details.

2.1.1.2. European Clinical Research Infrastructure Network chief investigators. We surveyed CIs working in collaboration with European Clinical Research Infrastructure Network (ECRIN), a nonprofit distributed infrastructure that supports the conduct of multinational clinical research in Europe. The ECRIN European Correspondents (EuCos), based within each member or observer country, distributed the survey within their respective countries. The survey context and purpose was explained to the EuCos via a Web-based meeting before contacting CIs of ECRIN-supported studies via email. The email contained an invitation to participate and the survey as an attached Word document. Two reminders were sent to the EuCos requesting recirculation of the survey.

2.1.2. Statisticians' survey

We surveyed statisticians within UKCRC-registered CTUs and ECRIN. The aim of the survey was to establish current practice, and knowledge and implementation of available statistical models. The survey was reviewed within the study team and piloted with a senior statistician before circulation across the UKCRC-registered CTU Statistics Group and ECRIN. A prize equivalent to £75 in vouchers was offered as an incentive to participation.

The full list of questions for the online survey and the invitation email are provided in the Additional file 1.

2.1.2.1. UK CTU Statisticians. The UKCRC-registered CTU secretariat distributed the survey via email to the nominated senior statistician within each registered CTU. The email contained an invitation to participate and a link to an online survey, which was constructed using SelectSurvey.NET (http://selectsurvey.net/). The statistician could discuss responses within the wider statistical team of their CTU, but only a single response per CTU was required. Email reminders were sent after 1, 2, and 4 weeks. Nonrespondents also had the opportunity to respond by completion of a paper copy of the survey distributed during the UKCRC-registered CTU Statistics Operational Group Network statisticians meeting in October 2018.

2.1.2.2. ECRIN Statisticians. The EuCos at ECRIN circulated the email invitation with the link to the online survey to the statisticians identified within their national network. The same procedure was followed as for the CIs' survey with the EuCos sending two reminders.

2.2. Analysis

Quantitative data from closed-ended questions were analyzed using RStudio, version 3.5.0 [12]. Due to the restricted sample sizes, statistical testing was not planned and results are reported as frequencies and percentages. EG, SD, and CG reviewed responses to open-ended questions identifying themes within the free text answers and categorized them in groups.

3. Results

3.1. Chief investigators' survey

The CIs' survey was conducted between 24 October 2018 and 30 November 2018 within the United Kingdom and between 18 October 2018 and 8 March 2019 within ECRIN with results summarized in Table 1. A total of 32 studies were identified as eligible for inclusion in the UK cohort, and 17 responses were received (53%) from the CIs contacted. Two CIs completed the survey twice each allowing for the multiple trials which they led as the CI. Nineteen studies were identified via the ECRIN EuCos with six responses (32%) received.

The data source most commonly used to predict trial recruitment was audit data from across multiple sites with the impact of specific eligibility criteria being the most frequently adjusted factor (Table 1, Question 1). Although no respondents reported adjusting for ethnic minorities, one respondent elaborated that not adjusting for this factor negatively impacted their predictions.

Allowing for variation in recruitment rates at individual sites was also common (13/17 UK, 6/6 ECRIN) with comments supporting the need for this practice based on variation in patient numbers and knowledge of site research activity infrastructure and experience. The majority of UK respondents (15/17, 88%) did not assume that all sites would be open for the same length of time in comparison with only one of the six ECRIN respondents (17%). Free text responses reported staggered opening times to reflect variation in time required at each site to obtain approvals.

Eleven (65%) UK respondents searched a trial registry for competing trials compared with 100% of ECRIN respondents (Table 1, Question 6). Thirty-five percent of UK and fifty percent of ECRIN respondents were aware of other trials competing to recruit the same patient population. Co-enrolment was considered for only half of UK and one-third of ECRIN respondents. One-third of UK and half of ECRIN respondents reported a need to be optimistic about the predicted recruitment rate for the trial to be attractive to the funder. Free text comments highlighted the difficulties this practice would lead to during trial conduct. Additional comments stated that the estimates were reflective of recruitments rates if things went well, accepting that this may not be the case with an inability to accurately predict researcher performance and stability of local clinical services.

3.2. Statisticians' survey

The survey was conducted between 18 September 2018 and 6 November 2018 among the 51 UKCRC-registered CTUs of whom 50 (98%) responded (46 responses completed online, four responses completed at the network meeting). The ECRIN EuCos circulated the survey between 13 November 2018 and 29 January 2019 to 53 participants of whom 19 (36%) responded.

Table 2 summarizes the survey results.

Table 1. Survey results of the Chief Investigators' survey (UK & ECRIN)

Question	Answers	UK *N* = 17 *n* (%)	ECRIN *N* = 6 *n* (%)	Overall *N* = 23 *n* (%)
1) In determining the disease or condition prevalence, what sources of data were available to you to use? *Please select all that apply.*	Population-based data on geographical areas covered by sites	9 (53)	5 (83)	14 (61)
	Disease/condition incidence data	8 (47)	6 (100)	14 (61)
	Audit data from a single site	5 (29)	3 (50)	8 (35)
	Audit data from multiple sites	14 (82)	4 (67)	18 (78)
	Estimates obtained from sites based on their experience/perceptions rather than available data	5 (29)	3 (50)	8 (35)
	Feasibility or pilot study	5 (29)	3 (50)	8 (35)
	Previous RCTs in similar populations	7 (41)	4 (67)	11 (48)
	Other: Please specify[a]	4 (24)	1 (17)	5 (22)
2) In considering the translation of these data sources to your trial population which of the following adjustments did you make within your grant application to predict recruitment in to your study? *Please select all that apply.*	Estimated impact of specific eligibility criteria	15 (88)	6 (100)	21 (91)
	Ethnic minorities	0 (0)	0 (0)	0 (0)
	Seasonal effects	4 (24)	2 (33)	6 (26)
	Consent rate	13 (76)	3 (50)	16 (70)
	Other: Please specify[b]	6 (35)	1 (17)	7 (30)
	None	1 (6)	0 (0)	1 (4)
3) Within your trial's recruitment period, did you assume that all sites would be open for the same length of time?	Yes	2 (12)	5 (83)	7 (30)
	No: Please specify	15 (88)	1 (17)	16 (70)
4) Within your trial's recruitment period, did you assume that all sites would have the same average recruitment rate?	Yes	4 (24)	0 (0)	4 (17)
	No: Please specify	13 (76)	6 (100)	19 (83)
5) In considering recruitment to your trial, were you aware of any trials recruiting at the same time that would compete for the same patient population?	Yes: Please specify any strategy used to allow for the impact on your recruitment	6 (35)	3 (50)	9 (39)
	No	11 (65)	3 (50)	14 (61)
6) Did you search a trial registry for competing trials?	Yes	11 (65)	6 (100)	17 (74)
	No	6 (35)	0 (0)	6 (26)
7) Is your trial open to coenrollment (e.g., patient enrollment to more than one trial)?	Yes: If yes, what restrictions apply?	9 (53)	2 (33)	11 (48)
	No	8 (47)	4 (67)	12 (52)
8) In estimating your recruitment rate, there may be a need to be optimistic about your recruitment rate for the trial to be attractive to the funder. Do you feel that this issue impacted the recruitment rate used?	Yes	6 (35)	3 (50)	9 (39)
	No	11 (65)	3 (50)	14 (61)

[a] Local patient survey; data compiled by a specific NIHR biomedical research unit; national data on disease activity; multiple sources.

[b] Difficulties in recruiting; logistics of recruitment; availability of research nurses; data on rates of recruitment from previous studies; a general rule: 50% of what the Principal Investigator estimates; eligible Vs consent rate, e.g., we expect the recruitment to be something between 30 and 50% of the eligible population depending on the trial question; impact of recruiters.

Table 2. Survey results to closed-ended questions of the Statisticians' Survey (UK and ECRIN)

Question	Answer	UK, $N = 50$ *n* (%)	ECRIN, $N = 19$ *n* (%)	Overall, $N = 69$ *n* (%)
Introductory questions				
1) Who usually leads recruitment prediction for a clinical trial within your unit? *Please select all that apply.*	Chief investigator	33 (66)	13 (68)	46 (67)
	Trial coordinator	28 (56)	5 (26)	33 (48)
	Statistician	29 (58)	10 (53)	39 (57)
	Other (e.g., IT team, senior staff)	6 (12)	1 (5)	7 (10)
2) Do you believe a statistician should be involved in the recruitment prediction process?	Yes	43 (86)	13 (68)	56 (81)
	No	3 (6)	6 (32)	9 (13)
3) When predicting the recruitment rate at the pretrial planning stage, where do you find the information about the prevalence of the condition being studied, the eligibility of patients, the consent rate of participants etc.? Please provide information[a]	Published literature	28 (56)	10 (53)	38 (55)
	Research team experience	28 (56)	10 (53)	38 (55)
	Previous studies	22 (44)	4 (21)	26 (38)
	Registry data/audit data/patient databases/hospital data	22 (44)	5 (26)	27 (39)
	Feasibility surveys/pilot studies/sites' questionnaire	24 (48)	5 (26)	29 (42)
	Conservative interpretation of previous experience or consent rate	4 (8)	0 (0)	4 (6)
	PPI[b] engagement group	2 (4)	0 (0)	2 (3)
	Projections were not particularly evidence based	1 (2)	0	1 (1)
	NA	0 (0)	1 (5)	1 (1)
4) In considering recruitment to trials in your CTU, are you usually confident that you are aware of other trials recruiting at the same time that would compete for the same patient population?	Not confident at all	0 (0)	1 (5)	1 (1)
	Not very confident	6 (12)	11 (58)	17 (25)
	Neither	10 (20)	1 (5)	11 (16)
	Fairly confident	28 (56)	5 (26)	33 (48)
	Very confident	6 (12)	1 (5)	7 (10)
Recruitment prediction				
5) In addition to the number of patients and the number and size of sites, what factors would you routinely consider when predicting rates of recruitment? *Please select all that apply.*	Staggered site openings	48 (96)	13 (68)	61 (88)
	Seasonal variation	24 (48)	9 (47)	33 (48)
	Holiday periods	21 (42)	9 (47)	30 (43)
	Other	9 (18)	6 (32)	15 (22)
6) Do you use any statistical model for recruitment prediction?	Yes	3 (6)	4 (21)	7 (10)
	No	47 (94)	15 (79)	62 (90)
7) Are you aware of any of the statistical approaches listed below for use in recruitment prediction? *Please select all that apply.*	Poisson model—assumes a constant average rate of recruitment	23 (46)	13 (68)	36 (52)
	Poisson gamma model—which models variability in	13 (26)	8 (42)	21 (30)

(*Continued*)

Table 2. Continued

Question	Answer	UK, $N = 50$ *n* (%)	ECRIN, $N = 19$ *n* (%)	Overall, $N = 69$ *n* (%)
	center recruitment rates using a gamma distribution			
	Bayesian approaches requiring a prior for recruitment to be specified	12 (24)	9 (47)	21 (30)
	Other	2 (4)	2 (11)	4 (6)
	None	24 (48)	4 (21)	28 (41)
8) Have you ever simulated recruitment data to support your pretrial planning?	Yes, routinely	0 (0)	1 (5)	1 (1)
	Sometimes	16 (32)	6 (32)	22 (32)
	Never	34 (68)	12 (63)	46 (67)
9) If you do not use any of the approaches mentioned above for recruitment prediction, what is the reason for this? *Please select all that apply.*	I prefer using a simple approach (e.g., using Excel) rather than assuming statistical distributions for recruitment prediction	22 (44)	8 (42)	30 (43)
	I am not familiar with these models for recruitment prediction	17 (34)	2 (11)	19 (28)
	I am familiar with some/all of these models, but I don't know how to implement them for recruitment prediction	6 (12)	2 (11)	8 (12)
	I am not convinced of the value of implementing these models	27 (54)	2 (11)	29 (42)
	Other	8 (16)	8 (42)	16 (23)
Recruitment monitoring and implementation of statistical models via Web application				
10) How do you routinely monitor recruitment during the course of a trial? *Please select all that apply.*	Tables showing the expected and actual recruitment rates	43 (86)	14 (74)	57 (83)
	Recruitment graphs showing the expected and actual recruitment rates	49 (98)	11 (58)	60 (87)
	Individual recruitment targets for each site	41 (82)	10 (53)	51 (74)
	Common recruitment target for all sites	24 (48)	9 (47)	33 (48)
	Comparison of overall recruitment rates for each site with recruitment rate over recent months	31 (62)	8 (42)	39 (57)
	Other	8 (16)	0 (0)	8(12)
11) Are you aware of any software/Web platforms for planning and monitoring patient recruitment?	Yes	3 (6)	2 (11)	5 (7)
	No	47 (94)	17 (89)	64 (93)
12) If a user-friendly Web application implementing some of the aforementioned models became freely available, would you be interested in using it for predicting and/or monitoring of the trial recruitment? *Please select all that apply.*	No, I don't believe it is a statistical issue, and it is best handled by the trial team	4 (8)	3 (16)	7 (10)
	Not for prediction, but I would be interested in using it for monitoring	3 (6)	1 (5)	4 (6)
	Yes, I want to improve prediction of recruitment	14 (28)	5 (26)	19 (28)
	Yes, I want to use it for both initial prediction and monitoring of recruitment	27 (54)	11 (38)	38 (55)
	Other	18 (38)	3 (16)	21 (30)

[a] Free-text responses have been categorized into common themes.
[b] Patient and Public Involvement.

The majority believe that statisticians should be involved in predicting recruitment (86% UK, 68% ECRIN); however, statisticians were reported to have been involved in leading the process in only just over half of the studies. Respondents from ECRIN reported were less confident than UK respondents in their awareness of other trials competing to recruit from the same patient population (12% UK, 63% ECRIN) and were less likely to adjust for staggered site openings (96% UK, 68% ECRIN).

Use of statistical models to predict recruitment was low overall (10%) but higher within ECRIN respondents (6% UK, 21% ECRIN) who also had greater awareness of the individual statistical approaches with 48% of UK respondents not aware of any method compared with 21% within ECRIN.

At the pre-trial planning, only 32% of respondents sometimes simulated data to support recruitment prediction, whereas 67% of them never did (Table 2, Question 8). The time investigators would need to dedicate to perform simulations is an additional challenge, especially if they are not convinced of their value.

The majority of respondents who sometimes simulated data are at least aware of the Poisson model (73%, 16/22), whereas six of them are not aware of any statistical model for recruitment prediction (27%, 6/22). On the other hand, almost half of respondents who never simulated data, are not aware of any model (48%, 22/46), whereas 43% of them are at least aware of the Poisson model (20/46). However, because of the small sample size, we cannot conclude a definitive correlation between knowledge of statistical models and use of recruitment simulations.

A sizable proportion of respondents (44% UK, 42% ECRIN) preferred to use a simple approach rather than statistical distributions to predict recruitment. Slightly over half of UK respondents were unconvinced of the value of implementing these methods in comparison to only 11% of ECRIN respondents (Table 2, Question 9).

A minority of respondents were aware of existing software or Web platforms to support planning and monitoring of patient recruitment with over a quarter of respondents being interested in such a resource for predicting recruitment and over half interested for both prediction and monitoring (Table 2, Question 12). Free text responses indicated that time to learn how to use such an application and funder willingness to support any associated costs were a concern. Other participants expressed an interest in comparing any methods alongside those already used in practice to determine whether any time or resource investment was worthwhile. One participant expressed concerns on whether the requirements of more complex clinical trials could be met by such a resource.

Response to the free-text question about further practices, tools or resources that could potentially improve prediction accuracy are provided in Box 1, with suggestions for how funders/trial teams could monitor recruitment progress/milestones summarized in Box 2.

Box 1 Further practices or tools to improve recruitment prediction

Question 13: Please give details of any further practices or tools/resources that you think could influence your future practice, in terms of prediction accuracy in patient recruitment.

- Training (e.g., work with CIs to show the value of involving statistics for recruitment purposes before and during the trial/workshops for trial statisticians/challenge the clinicians, etc.) (8/69, 12%)
- Better engagement (e.g., with potential sites, using standard questionnaires to elicit proposed recruitment target/easily accessible & timely input from clinical communities & access to relevant patient groups, etc.) (5/69, 7%)
- Raising awareness of the available approaches (2/69, 3%)
- Current challenges and conflict with NIHR CRN targets (e.g., sites come back and ask to change their local recruitment target to ensure they are not challenged/providing the funding for realistic timelines/building in flexibility with timelines for project management and funding/allowing for reallocation of research resources to new studies toward the end of the study, etc.) (3/69, 4%)
- Recommendations to improve prediction (e.g., building up a database of actual recruitment in our studies that could be referred to in future/use anonymized registry of patients with relevant disease along with demographic information/valid international data on disease incidence and prevalence/comparing predictions and targets to what actually happened so future predictions can be improved/comparing recruitment in pilot studies with that in full trials/getting funders to request more rigorous methods to estimate recruitment, etc.) (10/69, 14%)
- Tool/model (a tool that automatically integrates recruitment predictions for individual sites in a multicentre trial into an overall prediction for the trial/simple, robust methods yielding accurate results/a smoothed time-autocorrelated prediction might be helpful/any tool that helps to maintain engagement/enthusiasm) (8/69, 12%)
- Demonstrating evidence that these models actually work in practice (e.g., it is important to show that prediction ability of a model/tool is better than the simpler ways/it can be cumbersome to gather all information to feed into prediction tools and it would require further input from a statistician, etc.) (2/69, 3%)
- No response (35/69, 51%)
- Response not clear (3/69, 4%)

Box 2 Suggestions for monitoring recruitment progress

Question 14: Do you have any comments or suggestions on how funders/trial teams monitor recruitment progress/milestones?

- Educate funders (e.g., Funders' expectations of trials set up and recruitment rate should be more realistic/they should be less rigid in setting targets and monitoring against those targets/funders want more for less/funding panels to define the feasibility of recruitment rates and convince the administrative funders that studies need more time to be successfully delivered, etc.) (3/69, 4%)
- Educate Chief Investigators (CIs tend to be overoptimistic/Methodologists/statisticians usually try to be conservative but this is challenged by CIs, etc.) (2/69, 3%)
- Build on skills and experience gained from previous trials (e.g., pass on the skills and experience of trials teams that manage to recruit to time and budget/ a sort of rule of thumb is that 20-40% of potentially eligible people approached to take part will consent to inclusion/in cancer trials we tend to expect about 50% of eligible patients to consent etc.) (2/69, 3%)
- Take into account factors related to trial/outcomes/intervention/condition being studied (e.g., seasonal factors/clinic frequency/TTE[1] considerations/staggered entry assumptions etc.) (4/69, 6%)
- Allow for delays outside of the control of the trial management team during the course of the trial (e.g., due to budgeting/staffing/resources/new trials opening/delays in agreeing contracts/whole centers dropping out etc.) (2/69, 3%)
- Central database with recruitment information from previous trials to accompany the online tool (1/69, 1%)
- Uncertainty to be considered (e.g., any recruitment estimates at the onset of a trial will be based on assumptions, e.g., average recruitment rate per site or something similar/there are so many variables involved and I am not sure there are any decent ways of getting around that/initial recruitment predictions tend to be very inaccurate/any prediction or monitoring of recruitment using sophisticated modeling may not be any better than using simple projections, etc.) (5/69, 7%)
- Generic programming (programming something generic is important, because statisticians are already under a lot of time pressure and deadlines, so they want to avoid an overload of duties) (1/69, 1%)
- No response (52/69, 75%)
- Response not clear (3/69, 4%)

[1] TTE: time to event.

4. Discussion

This survey is the first to identify current practice on methods to predict recruitment in clinical trials and raises hypotheses about different practices in the United Kingdom compared with Europe and the perceived value of more complex statistical approaches.

Survey responses clearly indicate that the statistical models available are not being implemented. The absence of a robust demonstration of their benefits in comparison with simple approaches is a key barrier to their uptake. The statistical literature is restricted to the evaluation of these models in simulations or in retrospective trials [7,13,14]. It lacks a real-time prospective evaluation using the same limited information sources to support parameter estimation across models at the design stage, which are then used to monitor actual accrual. Furthermore, the survey suggested that this evidence is required before trial statisticians being able to justify the time required to understand and implement the methodology, suggesting that software availability on its own is insufficient to change practices.

There are many factors to be taken into account when predicting recruitment and in turn defining a trial's duration. Recruitment targets cannot be realized if based on overoptimistic expectations and unrealistic timelines. The overoptimistic expectations of the research team have been reported previously [15]; however, this survey highlights the tension felt by a sizable proportion of investigators to be optimistic about their recruitment rates, for the trial to be attractive to funders. Despite this tension, there was a clear appreciation of the difficulties this would cause at later time points, with calls for funders' expectations of trial set up times and recruitment rates to be more realistic, less rigid and to allow for unforeseen delays outside the control of the trial management team.

The importance of adhering to the site initiation schedule is key and our anecdotal experience is that deviations from this often explain a substantial proportion of under recruitment. The time required to complete the administrative arrangements which need to be made to open participating sites can vary and be impacted by site engagement and capacity or by regulatory changes [16–20]. The survey indicated that staggered site openings are more commonly allowed for within the United Kingdom than across the ECRIN network. However, although the rate of site initiations may be informed by past experience, there is an inherent assumption regarding the stability of site resources to deliver the research, remaining stable throughout the trial. Although a potential solution is to improve site feasibility and capability assessments, the variables that need to be included and how they are used within resulting predictions, needs further scrutiny.

The survey demonstrates that investigators and statisticians are using a wide variety of information sources to predict their recruitment rates. However, in practice the extrapolation of these data to a specific multicenter trial often requires adjustments to be used. The size of the adjustments may be considered to be arbitrary or based on guess work, and this may in part reduce the number of factors investigators feel able to include. In addition, as the availability of routinely collected data increases to support clinical trial planning, the assessment of such sources and how they are used will be of increasing importance.

The main focus of our survey was the prediction of recruitment; however, we also aimed to ascertain how this was monitored against observed accrual. The responses demonstrated that this information is considered in multiple ways per trial with comparisons of observed recruitment rates against those predicted in graphical or tabular form, at individual site level and across all sites, and covering the entire recruitment period or restricting to recent months. Although eight respondents indicated ''other'' methods were used, the free text provided, demonstrated that the approach was consistent with the closed response categories. The limitation with these approaches is that they do not allow understanding of whether the observed variation is within reasonable limits of the prediction. This may lead to delays in remedial actions. A potential benefit of using a statistical model is the prespecification of a quantile to act as a trigger when the observed recruitment rate is inconsistent with that prespecified. One respondent commented that the use of statistical models would ''simply give a distribution of recruitment rates from which we would need to pick a final number which would be the mean, so simple multiplication would seem as appropriate given the uncertainty about the assumptions.'' This suggests that even if the uncertainty they elude to is not welcome within prediction, there is potential for their use within monitoring.

In a survey of the UKCRC-registered CTUs, the top inefficiency from recruitment of the first participant to the publication of results, was identified as the failure to meet recruitment targets due to overoptimistic or inaccurate recruitment estimates [21]. Some statisticians reported being under pressure to project optimistic recruitment rates. This is likely due to the perception that realistic rates are associated with increased budgets, beyond what funders are willing to provide. This issue was raised by our survey respondents, with participants' suggestions that training should be provided for both CIs and funders. Increasing funders' flexibility in setting timelines would be helpful and reflects additional calls on requirements with adaptive designs [22].

The majority of our respondents believe that a statistician should be involved in the recruitment prediction process but do not model recruitment as a stochastic process. This may be in part explained by time pressures, as recruitment prediction is undertaken during the unfunded preparation time of a grant application. This will be compounded given investigators and statisticians are unconvinced that the models described in the literature are worthy of the additional time required to support their use. However, the survey demonstrates the majority would be interested if the benefits were found to justify the additional time and statistical expertise required.

4.1. Limitations of the studies

This survey aimed to elicit current recruitment practice across the United Kingdom and Europe. The high-response rates from the United Kingdom are a strength of the survey; however, this means that the findings predominantly represent current practice within the United Kingdom.

The network structure within the United Kingdom facilitated survey distribution in a controlled approach using the network secretariat ensuring a targeted delivery and response, whereas the ECRIN approach used a more fluid hub and spoke model where the CIs and the statisticians were contacted by the EuCos in each country. Other surveys targeting statisticians across the UKCRC CTU network have achieved similarly high-response rates [23−25], and we have been unable to identify similar surveys across ECRIN. The lower response rates from ECRIN may be a result of these different network infrastructures; however, they may also impacted by the survey being restricted to the English language.

The comparison of practices between the United Kingdom and Europe therefore needs to be interpreted with caution as this could reflect response bias within ECRIN with those with particular interests in recruitment prediction taking part. This may be an explanation for the greater awareness of the statistical methods in the ECRIN respondents.

Furthermore, the investigators in the UK survey were identified from the website of the largest public funder of clinical trials in the United Kingdom. The sampling framework for investigators across ECRIN was by identification of the EuCos and therefore not restricted to a funding source. However, knowledge of both networks suggests that the portfolio of studies represented is restricted to noncommercial research. It would be of interest to understand differences in methods used in comparison with industry sponsored studies, with research suggesting recruitment for industry-sponsored studies being less problematic [26]. Yet, it is likely that the resources allocated and incentives provided are dissimilar and this complexity of factors warrants further detailed exploration in future research.

5. Conclusions

Approaches used to predict and monitor recruitment remain frequently unreported, and this survey provides insight from both statisticians and investigators on methods and data sources used. This study indicates that the majority of respondents did not recognize recruitment as a

stochastic process in the approaches used and stated a preference for using simple approaches. However, they consider the involvement of statisticians in the recruitment prediction process to be essential. Simple approaches will continue to be used despite the advancement of more complex statistical models until their value in improving prediction can be more robustly demonstrated. Until then, their complexity, time, and training required to aid their implementation will remain a barrier despite the potential for their added benefits in monitoring of recruitment.

CRediT authorship contribution statement

Efstathia Gkioni: Conceptualization, Methodology, Formal analysis, Data curation, Writing - original draft, Writing - review & editing. **Susanna Dodd:** Formal analysis, Data curation, Writing - review & editing. **Roser Rius:** Formal analysis, Data curation, Writing - review & editing. **Carrol Gamble:** Conceptualization, Methodology, Formal analysis, Data curation, Writing - original draft, Writing - review & editing.

Acknowledgments

We would like to express our thanks to the following people who contributed to the completion of this work:

- All the participants who took the time to complete both surveys.
- The gatekeeper who circulated the statisticians' survey within the UK CRC network (UKCRC secretariat).
- The European Correspondents from ECRIN who circulated the statisticians' survey and chief investigators' surveys.
- The Operations Director of ECRIN Dr Christine Kubiak for facilitating the communication with the European Correspondents within the ECRIN.
- Dr Ashley Jones for his feedback in piloting the survey questions and answers. Mrs. Elizabeth J. Conroy for her help in reviewing this manuscript before submission.

Funding: This project is part of the MiRoR (Methods in Research on Research) Joint Doctoral Training Programme. The European Union's Horizon 2020 Marie Sklodowska Curie Innovative Training Networks−European Joint Doctorate (ITN-EJD) under grant agreement No. 676207, fund MiRoR.

Availability of data and materials: Not applicable.

Authors' information: Not applicable.

Ethics approval and consent to participate: Ethical approval was granted from Health and Life Sciences Committee on Research Ethics (Human participants, tissues and databases) at The University of Liverpool on 5 September 2018 (reference 2282). All survey participants were provided with full written information before survey commencement with the act of completing the survey, a demonstration of consent.

Consent for publication: Not applicable.

Competing interests: The authors declare that they have no competing interests.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.03.012.

References

[1] Treweek S, Lockhart P, Pitkethly M, Cook JA, Kjeldstrøm M, Johansen M, et al. Methods to improve recruitment to randomised controlled trials: cochrane systematic review and meta-analysis. BMJ Open 2013;3:e002360.

[2] Ross SGA, Counsell C, Gillespie W, Russell I, Prescott R. Barriers to participation in randomised controlled trials: a systematic review. J Clin Epidemiol 1999;52:1143−56.

[3] Sully BGO, Julious SA, Nicholl J. A reinvestigation of recruitment to randomised, controlled, multicenter trials: a review of trials funded by two UK funding agencies. Trials 2013;14:166.

[4] Healy P, Galvin S, Williamson PR, Treweek S, Whiting C, Maeso B, et al. Identifying trial recruitment uncertainties using a James lind alliance priority setting partnership−the PRioRiTy (prioritising recruitment in randomised trials) study. Trials 2018;19:147.

[5] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010;152:726−32.

[6] Barnard KD, Dent L, Cook A. A systematic review of models to predict recruitment to multicentre clinical trials. A systematic review of models to predict recruitment to multicentre clinical trials. BMC Med Res Methodol 2010;10:63.

[7] Gkioni E, Rius R, Dodd S, Gamble C. A systematic review describes models for recruitment prediction at the design stage of a clinical trial. J Clin Epidemiol 2019;115:141−9.

[8] Zweben A, Barrett D, Berger L, Murray KT. Recruiting and retaining participants in a combined behavioral and pharmacological clinical trial. J Stud Alcohol 200572−81.

[9] Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, et al. The Women's Health Initiative recruitment methods and results. Ann Epidemiol 2003;13:S18−77.

[10] Toddenroth D, Sivagnanasundaram J, Prokosch H-U, Ganslandt T. Concept and implementation of a study dashboard module for a continuous monitoring of trial recruitment and documentation. J Biomed Inform 2016;64:222−31.

[11] White D, Hind D. Projection of participant recruitment to primary care research: a qualitative study. Trials 2015;16:473.

[12] RStudio Team. RStudio: integrated development environment for R. http://www.rstudio.com/. Accessed September 10, 2019.

[13] Baldi I, Gregori D, Desideri A, Berchialla P. Accrual monitoring in cardiovascular trials. Open Heart 2017;4:e000720.

[14] Lan Y, Tang G, Heitjan DF. Statistical modeling and prediction of clinical trial recruitment. Stat Med 2019;38:945−55.

[15] Briel M, Olu KK, von Elm E, Kasenda B, Alturki R, Agarwal A, et al. A systematic review of discontinued trials suggested that most reasons for recruitment failure were preventable. J Clin Epidemiol 2016;80:8−15.

[16] Kearney A, McKay A, Hickey H, Balabanova S, Marson AG, Gamble C, et al. Opening research sites in multicentre clinical trials within the UK: a detailed analysis of delays. BMJ Open 2014;4:e005874.

[17] Hemminki A, Kellokumpu-Lehtinen P-L. Harmful impact of EU clinical trials directive. BMJ 2006;332:501−2.

[18] Hartmann M, Hartmann-Vareilles F. The clinical trials directive: how is it affecting Europe's noncommercial research. PLoS Clin Trials 2006;1:e13.
[19] Shakur H, Roberts I, Barnetson L, Coats T. Clinical trials in emergency situations. BMJ (Clin Res) 2007;334:165−6.
[20] Woolfall K, Frith L, Gamble C, Young B. How experience makes a difference: practitioners' views on the use of deferred consent in paediatric and neonatal emergency care trials. BMC Med Ethics 2013;14:45.
[21] Duley L, Gillman A, Duggan M, Belson S, Knox J, McDonald A, et al. What are the main inefficiencies in trial conduct: a survey of UKCRC registered clinical trials units in the UK. Trials 2018;19:15.
[22] Kairalla JA, Coffey CS, Thomann MA, Muller KE. Adaptive trial designs: a review of barriers and opportunities. Trials 2012;13:145.
[23] Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA 2017;318:2337−43.
[24] Hopkins C, Sydes M, Murray G, Woolfall K, Clarke M, Williamson P, et al. UK publicly funded Clinical Trials Units supported a controlled access approach to share individual participant data but highlighted concerns. J Clin Epidemiol 2016;70:17−25.
[25] Conroy EJ, Harman NL, Lane JA, Lewis SC, Murray G, Norrie J, et al. Trial Steering Committees in randomised controlled trials: a survey of registered clinical trials units to establish current practice and experiences. Clin Trials 2015;12:664−76.
[26] Kasenda B, Von Elm E, You J, Blümle A, Tomonaga Y, Saccilotto R, et al. Prevalence, characteristics, and publication of discontinued randomized trials. JAMA 2014;311:1045−52.