






## SYSTEMATIC REVIEW

# Why are feasibility studies accessing routinely collected health data? A systematic review [version 1; peer review: awaiting peer review]

Aziza Mirza<sup>1</sup>, Victoria Yorke-Edwards <sup>1</sup>, Sarah Lensen<sup>2</sup>, Macey L. Murray<sup>1</sup>, Carlos Diaz-Montana<sup>1</sup>, James Carpenter<sup>1,3</sup>, Matthew R. Sydes <sup>1</sup>, Sharon B. Love <sup>1</sup>

<sup>1</sup>MRC Clinical Trials Unit at UCL, UCL, London, WC1V 6LJ, UK

<sup>2</sup>Department of Obstetrics and Gynaecology, University of Melbourne, Melbourne, Victoria, 3052, Australia

<sup>3</sup>Medical Statistics, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK

---

**V1** First published: 16 Aug 2021, 10:815  
<https://doi.org/10.12688/f1000research.52486.1>  
Latest published: 16 Aug 2021, 10:815  
<https://doi.org/10.12688/f1000research.52486.1>

---

## Abstract

**Background:** Feasibility trials are often undertaken to determine whether a larger randomised controlled trial (RCT) is achievable. In a recent review, 15 feasibility trials accessed routinely collected health data (RCHD) from UK national databases and registries. This paper looks at attributes of these trials and the reasons why they accessed RCHD.

**Methods:** We extracted data from all publicly available sources for the 15 feasibility studies found in a previous review of trials successfully accessing RCHD in the UK between 2013–2018 for the purpose of informing or supplementing participant data. We extracted trial characteristics, the registry accessed, and the way the RCHD was used.

**Results:** The 15 feasibility RCTs were conducted in a variety of disease areas, and were generally small (median sample size 100, range 41–4061) and individually randomised (60%, 9/15). The primary trial outcome was predominantly administrative (non-clinical) (80%, 12/15) such as feasibility of patient recruitment. They were more likely to recruit from secondary care (67%, 10/15) settings than primary (33%, 5/15).

NHS Digital was the most commonly accessed registry (33% (5/15)) with SAIL databank (20% (3/15)), electronic Data Research and Innovation Service (eDRIS) and Paediatric Intensive Care Audit Network (PICANET) (each 13% (2/15)) also being accessed. Where the information was clear, the trials used RCHD for data collection during the trial (47%, 7/15), follow-up after the trial (27%, 4/15) and recruitment (13%, 2/15).

**Conclusions:** Between 2013 and 2018, 15 feasibility trials successfully

## Open Peer Review

**Approval Status** AWAITING PEER REVIEW

Any reports and responses or comments on the article can be found at the end of the article.

accessed UK RCHD. Feasibility trials would benefit, as with other trials, from guidance on reporting the use of RCHD in protocols and publications.

### Keywords

feasibility trial, pilot trial, systematic review, routinely-collected health data, registry, RCT, electronic health record

**Corresponding author:** Sharon B. Love ([s.love@ucl.ac.uk](mailto:s.love@ucl.ac.uk))

**Author roles:** **Mirza A:** Formal Analysis, Investigation, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Yorke-Edwards V:** Data Curation, Software, Writing – Review & Editing; **Lensen S:** Conceptualization, Data Curation, Writing – Review & Editing; **Murray ML:** Conceptualization, Writing – Review & Editing; **Diaz-Montana C:** Conceptualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Carpenter J:** Conceptualization, Writing – Review & Editing; **Sydes MR:** Conceptualization, Writing – Review & Editing; **Love SB:** Conceptualization, Project Administration, Supervision, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** This work was supported by Health Data Research UK, an initiative funded by UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities and by a grant from the Medical Research Council MC\_UU\_12023/24.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2021 Mirza A *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Mirza A, Yorke-Edwards V, Lensen S *et al.* **Why are feasibility studies accessing routinely collected health data? A systematic review [version 1; peer review: awaiting peer review]** F1000Research 2021, 10:815 <https://doi.org/10.12688/f1000research.52486.1>

**First published:** 16 Aug 2021, 10:815 <https://doi.org/10.12688/f1000research.52486.1>

## List of abbreviations

eDRIS - Data Research and Innovation Service  
 PICANET - Paediatric Intensive Care Audit Network  
 RCHD – routinely-collected health data  
 RCT – randomised controlled trial  
 UKCRC – UK Clinical Research Collaboration

## Introduction

Feasibility trials are considered the best way to identify the practicability and achievability of a definitive (and expensive) randomised controlled trial (RCT).<sup>1,2</sup> By asking the question ‘can this be done?’, conducting a feasibility trial before a main trial can help avoid research waste.<sup>1,3,4</sup>

Routinely collected health data (RCHD), are sought from national databases and registries to both supplement and replace data that would commonly be collected as part of traditional trial data collection.<sup>5,6</sup>

The trials community recognises considerable issues in accessing RCHD data for trials<sup>7</sup> and we may anticipate that applicants would not seek to access RCHD for feasibility studies which, by nature, are small and short-term. Our recent review<sup>5</sup> found about 1 in 10 applications for RCHD was for a feasibility study and therefore we determined to find out more about these studies.

## Methods

### Eligibility

Our systematic review<sup>5</sup> (PROSPERO CRD42019123088), published in 2020, considered successful applications for routinely collected health data (such as NHS Digital) from publicly-accessible release registers. This identified 160 successful applications for RCHD relating to RCTs between 2013 and 2018. Eligibility for that review included RCHD being used for any of the following: (i) replacing conventionally collected trial data; (ii) cross-checking against existing trial data (including participant-reported data); (iii) cross-checking RCHD from different sources; (iv) prompt the trial team to further investigate a possible outcome measure or event; (v) cost-effectiveness analysis; and (vi) solely methodological purposes. Use of RCHD to evaluate the number of potential patients without using the RCHD for the purposes of recruitment was excluded. This was achieved by initially selecting every entry containing the following terms rand\*, trial, RCT, study, placebo, phase. The resultant list was then screened independently by two people to achieve the final list of all RCTs. Any disagreements were resolved by discussion and re-checking.

The publicly available application summary for these 160 trials was hand-searched by two people and considered eligible for this review of feasibility studies if any of the following criteria were met: (i) included a description of a feasibility or pilot RCT in the study design, (ii) aimed to develop a full scale RCT following the conclusion of the feasibility trial, or (iii) included feasibility as a trial outcome. Again, at this stage, any disagreements were resolved by discussion and re-checking.

### Data collection and analysis

The underpinning data extraction has been detailed previously.<sup>5</sup> In summary, information was extracted from the publicly-accessible data release registers and any publicly available sources (trial websites, publications, trial registrations and protocols, last searched April 2020). Two authors independently extracted data onto a piloted data extraction form and any disagreements were resolved by discussion and re-checking. Trial teams were not contacted for information or clarification. Data collection included information about the RCT, the source of RCHD, and the way in which the data were used (e.g. linkage identifiers used, category of data use).

For each included feasibility trial, any previous access of RCHD prior to 2013 was also included. We extracted data about the trial, the RCHD source and the planned use of the RCHD.

For these feasibility trials we also collected the wording used to describe the use of RCHD in publicly available trial protocols, websites, publications and registrations (Last checked April 2020).

The dataset was stored and analysed descriptively in Microsoft Excel (2016).

### Patient and public involvement

There was no patient and public involvement in this review.

**Results**

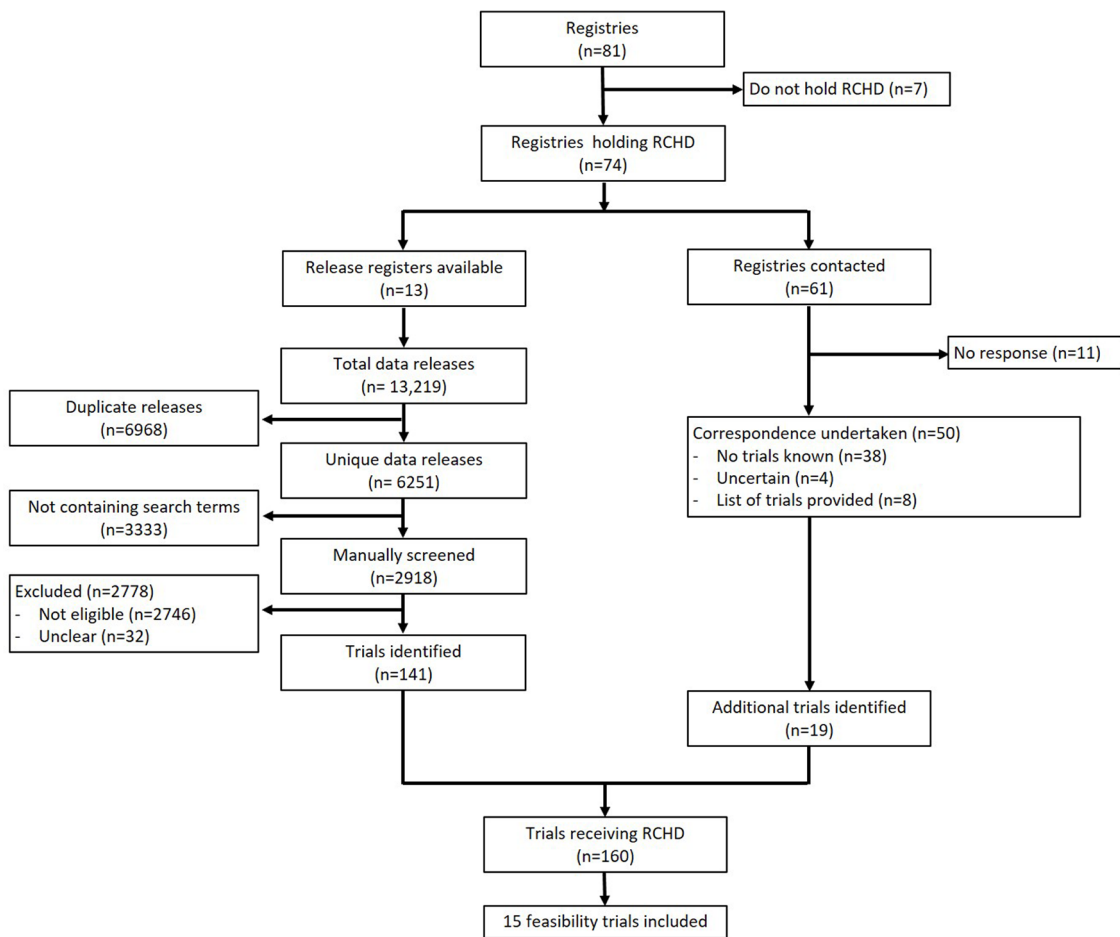
The PRISMA flowchart is given in **Figure 1**.

**RCT characteristics**

The search and extraction of data for feasibility trials was conducted between November 2019 and April 2020 and the data set is available from the UCL Research Data repository.<sup>8</sup> The characteristics of the 15/160 (9%) feasibility trials that accessed RCHD are shown in **Table 1**. Feasibility studies were generally small (median sample size 100, range 41-4061 participants). All 15 (100%) had, by the data freeze, completed recruitment and follow-up. The majority (9/15, 60%) were individually randomised trials; the remainder (6/15, 40%) were cluster randomised. The main disease category was mental health (4/15, 27%) followed by cardiovascular, infection and stroke (each 3/15, 20%). RCT intervention was often compared to standard-of-care (13/15, 87%), with only 2/15 (13%) to be assessed against a placebo-based control. The majority of interventions in these trials were a change to patient management (10/15, 67%). For example, direct transport to specialist centres rather than Accident & Emergency services or providing paramedics with additional training. 12/15 (80%) trials had an administrative, non-clinical primary outcome measure (e.g. can the trial team recruit 60 patients in one year); the other 3/15 (20%) had a clinical primary outcome measure. All trials recruited only in the UK. Most were coordinated through a UK Clinical Research Collaboration (UKCRC) Registered Clinical Trials Unit (9/15, 60%) and were predominantly open label (9/15, 60%). More trials recruited in secondary care (10/15, 67%) than primary care (5/15, 33%).

**RCHD access and use**

The most commonly accessed data sources were NHS Digital (5/15, 33%), SAIL databank (3/15, 20%), electronic Data Research and Innovation Service (eDRIS, formerly called ISD) (2/15, 13%) and Paediatric Intensive Care Audit Network (PICANET) (2/15, 13%) (**Table 2**), with each trial only accessing one source. Data collected from RCHD application logs



**Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart.**

**Table 1. Characteristics of feasibility RCTs receiving RCHD during 2013-2018.**

	Feasibility trials successfully accessing UK RCHD 2013-2018		Other trials successfully accessing UK RCHD 2013-2018	
	N	%	N	%
N	15		145	
UK based	15	100	111	77
Trial completed	15	100*	90	62
Randomisation				
Individually randomised	9	60	126	87
Cluster randomised	6	40	19	13
Intervention				
Treatment	12	80	104	72
Drug	3	20	73	50
Other	9	60	31	22
Primary prevention	2	13	26	18
Screening	1	7	15	10
Masking				
Open	9	60	104	72
Research staff only	3	20	9	6
Participant only	2	13	2	1
Research staff and participant	1	7	30	21
Coordination				
Registered CTU	9	60	94	65
Non-registered CTU	5	33	17	12
Unclear	1	7	34	23
Recruitment setting				
Primary care	5	33	36	25
Secondary care	10	67	109	75
Disease category				
Mental health	4	27	8	6
Cardiovascular and stroke	3	20	43	30
Infection	3	20	5	3
Cancer	1	7	46	32
Other	4	27	43	30
Comparison				
Standard of care	13	87	125	86
Placebo	2	13	20	14
Primary outcome				
Administrative	12	80	0	0
Feasibility	11	73	145	100
Non-feasibility	1	7		
Clinical	3	20		
Trial size Median (range)	100 (41-4061)		2141 (60-6,000,000)	
Number of years recruiting (median)	1		4	

\*10/15 completed at the time of the original review.

indicated that the most common reasons for requesting RCHD were for data collection during the feasibility trial period (7/15, 47%), follow-up after the feasibility trial period (4/15, 27%), and recruitment (2/15, 13%) with 2/15 (13%) not clearly stating the reason (Table 2).

In some instances, the planned use of RCHD in the trial release register was inconsistent with information gathered from publicly available sources, such as protocols, publications and trial websites. Of the 15 trials, 40% (6/15) had the same RCHD use stated in publicly-available literature and the RCHD release register, for 3/15 (20%) the uses in both sources had some overlap (i.e. the use was similar but not the same) and for 6/15 (40%) the uses for the RCHD were disparate (Table 3). For 2/6 trials with consistent stated use, RCHD use was unclear in both the application and the publicly available literature, with the latter not mentioning the use of RCHD. In those with overlap, the trial documentation was more explicit than the release register. One trial for example had stated in the release register that RCHD would be used for

**Table 2. Registries used by feasibility RCTs and reason for RCHD use.**

	Recruitment	Follow-up	Data collection*	Unclear	n	%
ISD/eDRIS	1			1	2	13%
NHFD		1			1	7%
NHS-D/HSCIC			4	1	5	33%
NNRD at NDAU			1		1	7%
PICANET			2		2	13%
SAIL		3			3	20%
TARN	1				1	7%
<b>n</b>	<b>2</b>	<b>4</b>	<b>7</b>	<b>2</b>	<b>15</b>	
<b>%</b>	<b>13%</b>	<b>27%</b>	<b>47%</b>	<b>13%</b>		<b>100%</b>

\*Three specifically adverse events (3/15 = 20%).

**Table 3. Reasons for RCHD access from publicly available documentation and registry access log.**

Consistency of reason for RCHD access	Reason for RCHD access, collected from publicly available documentation	Reason for RCHD access on registry access log
Consistent	Unclear	Unclear
	Unclear	Unclear
	Recruitment	Recruitment
	Recruitment	Recruitment
	Data collection	Data collection
	Data collection	Data collection
Overlap	Follow-up, hospital admissions and mortality	Follow up
	Primary outcome data	Data collection
	Hospital admission and mortality	Follow up
Disparate	Cancer and death	Adverse Events
	Hospital attendances, cancer diagnoses and deaths	Adverse Events
	Service use data e.g. hospital attendance	Adverse Events
	Full trial data	Follow up
	Hospital care data	Follow up
	Unclear	Data collection

'data collection'. In the trial protocol however, RCHD use was described as: "Primary outcome measure: (...) this will be measured using routinely collected NHS data. The Hospital Episode Statistics (HES) system provides a central record of an individual's use of all emergency departments in England and data will be extracted from this system (using participants' NHS numbers) to provide information on individual participants' use of emergency departments at baseline and over the 12 months of follow-up". This is more specific than claimed in the release register. Out of the six trials for which the stated uses in the RCHD release register and publicly available literature were disparate, one trial had stated in the release register that RCHD would be accessed for adverse events whereas RCHD was accessed specifically for "Hospital attendances, cancer diagnoses and deaths".

## Discussion

Of the 160 trials previously identified as successfully accessing RCHD from UK registries between 2013-2018, 9% (15/160) were feasibility trials. Of these, 73% (11/15) had a primary aim of feasibility. As far as we are aware, this is the first report to describe the use of RCHD in feasibility trials.

On application, the reasons for requesting RCHD were data collection (47%, 7/15), follow-up (27%, 4/15) and recruitment (13%, 2/15), with the remaining two (13%, 2/15) not giving a clear reason. These are varied reasons.

Data collected from the registry applications was inconsistent with the wording found in publicly available documentation. Apart from guidance on the definition of feasibility studies, there needs to be clearer guidance on reporting the use of RCHD in trial protocols and publications. As per previous publications,<sup>9</sup> trial management details are often not included in papers. *SPIRIT guidelines*, section 18a, “Plans for assessment and collection of outcome, baseline, and other trial data” indicates this information should be in protocols; *CONSORT guidelines* 2010, section 4b “Settings and locations where the data were collected” indicates it should be in publications.

This study has several limitations. The scope of this review is restricted to registries in the UK, and captured trials accessing RCHD during a relatively short window (2013–2018). Local datasets could have been requested which were not part of the underpinning review. There may be future value in interviewing the trial teams about their experiences accessing and using RCHD, since the information available from publicly accessible data release requests was limited. The data used to inform this review was collected from a variety of sources in order to provide a comprehensive picture of the use of RCHD by included trials, however it remains possible we missed relevant sources. We aimed to include feasibility trials in this review and included trials that indicated that they were feasibility trials (and for which no information to contradict this was identified). However, it is possible that researchers conducting small trials refer to these as feasibility studies, despite not aiming to generate feasibility information per se; by definition, feasibility trials should only be started with the aim of getting information for a follow-on larger trial.<sup>1</sup> For example, one trial expressed the sentiment that it would be valuable to conduct a larger trial, suggesting feasibility may not have been the original plan.

## Conclusion

Between 2013 and 2018, 15 feasibility trials in the UK successfully accessed RCHD for reasons other than solely participant recruitment. As with other trials, feasibility trials would benefit from guidance on specifying the use of RCHD in protocols and publications.

## Data availability

### Underlying data

UCL Research Data: Underlying data for ‘Why are feasibility studies accessing routinely collected health data? A systematic review’, <https://doi.org/10.5522/04/14743836.v1>.<sup>8</sup>

This project contains the following underlying data:

- Data file 1: Feasibility trials using electronic health data from registries.

### Reporting guidelines

UCL Research data: PRISMA checklist and flow diagram for ‘Why are feasibility studies accessing routinely collected health data? A systematic review’, <https://doi.org/10.5522/04/14743836.v1>.<sup>8</sup>

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) (CC-BY 4.0).

## Acknowledgements

Thank you to Archie Macnair, Meredith Martyn, Alexandra Blenkinsop and Nurulamin M Noor for double-checking the original data extraction.

## References

1. Eldridge SM, Lancaster GA, Campbell MJ, *et al.*: **Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework.** *Plos One.* 2016; **11**.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
2. Thabane L, Ma J, Chu R, *et al.*: **A tutorial on pilot studies: the what, why and how.** *BMC Med Res Methodol.* 2010; **10**.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Morgan B, Heijdenberg J, Hinrichs-Krapels S, *et al.*: **Do feasibility studies contribute to, or avoid, waste in research?** *Plos One.* 2018; **13**.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. Myers BA, Pillay Y, Hornsby WG, *et al.*: **Recruitment effort and costs from a multi-center randomized controlled trial for treating depression in type 2 diabetes.** *Trials.* 2019; **20**.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

5. Lensen S, Macnair A, Love SB, *et al.*: **Access to routinely collected health data for clinical trials - review of successful data requests to UK registries.** *Trials.* 2020; **21**.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. McKay AJ, Jones AP, Gamble CL, *et al.*: **Use of routinely collected data in a UK cohort of publicly funded randomised clinical trials.** *F1000Res.* 2020; **9**: 323. 2020/11/12.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Lugg-Widger FV, Robling M: **Routinely collected data for trialists: The need for continued conversations and solution sharing.** *Clin Trials.* 2019; **16**: 217–218. 2018/11/18.  
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Mirza A, Yorke-Edwards V, Sydes M, *et al.*: **Feasibility trials using electronic health data from registries.** *UCL Research Data. Dataset.* 2021.  
[Publisher Full Text](#)
9. Beaumont D, Arribas M, Frimley L, *et al.*: **Trial management: we need a cadre of high-class triallists to deliver the answers that patients need.** *Trials.* 2019; **20**.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)



The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**