Please see below for a link to the webinar recording for the Trials Methodology Research Partnership:

Use of routine data in trials in the UK

Sharon Love (UCL) and Andrew McKay (University of Liverpool)

20 August 2020

The slides are also available below.

For any queries, please contact uktmn@nottingham.ac.uk

https://www.youtube.com/watch?v=W7iYHmxodXk

TMRP Webinar:
Use of routine data in trials in the UK

Welcome, and thank you for joining us.

Please remain muted and ensure your camera stream is turned off

MRC Clinical Trials Unit

Smarter studies
Global impact
Better health

UCL

How much are we using routinely collected electronic health data in trials in the UK?

Sharon Love
Associate Professor Trial Conduct Methodology
MRC Clinical Trials Unit at UCL
20th August 2020

Routinely Collected Health Data (RCHD)

- Electronic medical records, routinely collected health data, registry data, administrative databases etc…
- “(health) data collected without specific a priori research questions” (RECORD)
- Examples mortality data (cause and date of death), hospital admissions

Routinely Collected Health Data Examples

Mortality (ONS)
Hospital episode statistics (HES)

NCRAS

RCHD for outcomes - benefits

- National data capture
  - Including patients otherwise lost to follow-up/withdrawn from study visits
- Reduce burden on patients and site staff
  - Long term follow-up
- Bias
  - Reduces recall bias
  - Ensures fairer follow-up
  - Objective outcome assessment
  - Economic analysis
- Cost effective (potentially)

Plan

- Review of trials using RCHD
- Comparison of trial and registry data
- The changes due to COVID-19
- Future

Trials

Lensen et al. Trials (2020) 21:398
https://doi.org/10.1186/s13063-020-04329-8

Open Access

RESEARCH

Access to routinely collected health data for clinical trials – review of successful data requests to UK registries

Sarah Lensen[1], Archie Macnair[1], Sharon B. Love[1*], Victoria Yorke-Edwards[1], Nurulamin M. Noor[1], Meredith Martyn[1], Alexandra Blenkinsop[1], Carlos Diaz-Montana[1], Graham Powell[2], Elizabeth Williamson[3], James Carpenter[1,4†] and Matthew R. Sydes[1†]

Abstract

Background: Clinical trials generally each collect their own data despite routinely collected health data (RCHD) increasing in quality and breadth. Our aim is to quantify UK-based randomised controlled trials (RCTs) accessing RCHD for participant data, characterise how these data are used and thereby recommend how more trials could use RCHD.

We conducted a systematic review of RCTs accessing RCHD from at least one registry in the UK between
urposes of informing or supplementing participant data. A list of all registries holding RCHD
re registries published release registers, these were searched for RCTs

Systematic review: Aims

- How many UK trials are accessing RCHD to inform participant data?
- Which RCHD sources have been accessed for trials?
- Which trial types (disease area, size etc.)?
- How is the data being used?

Systematic review: Aims

- How many UK trials are accessing RCHD to inform participant data?
- Which RCHD sources have been accessed for trials?
- Which trial types (disease area, size etc.)?
- How is the data being used?

PROSPERO International prospective register of systematic reviews

Routine electronic health records (EHR) used as trial data by randomised controlled trials in the UK

Sarah Lensen, Archie Macnair, Matt Sydes, James Carpenter, Sharon Love, Victoria Yorke-Edwards, Elizabeth Williamson, Graham Powell

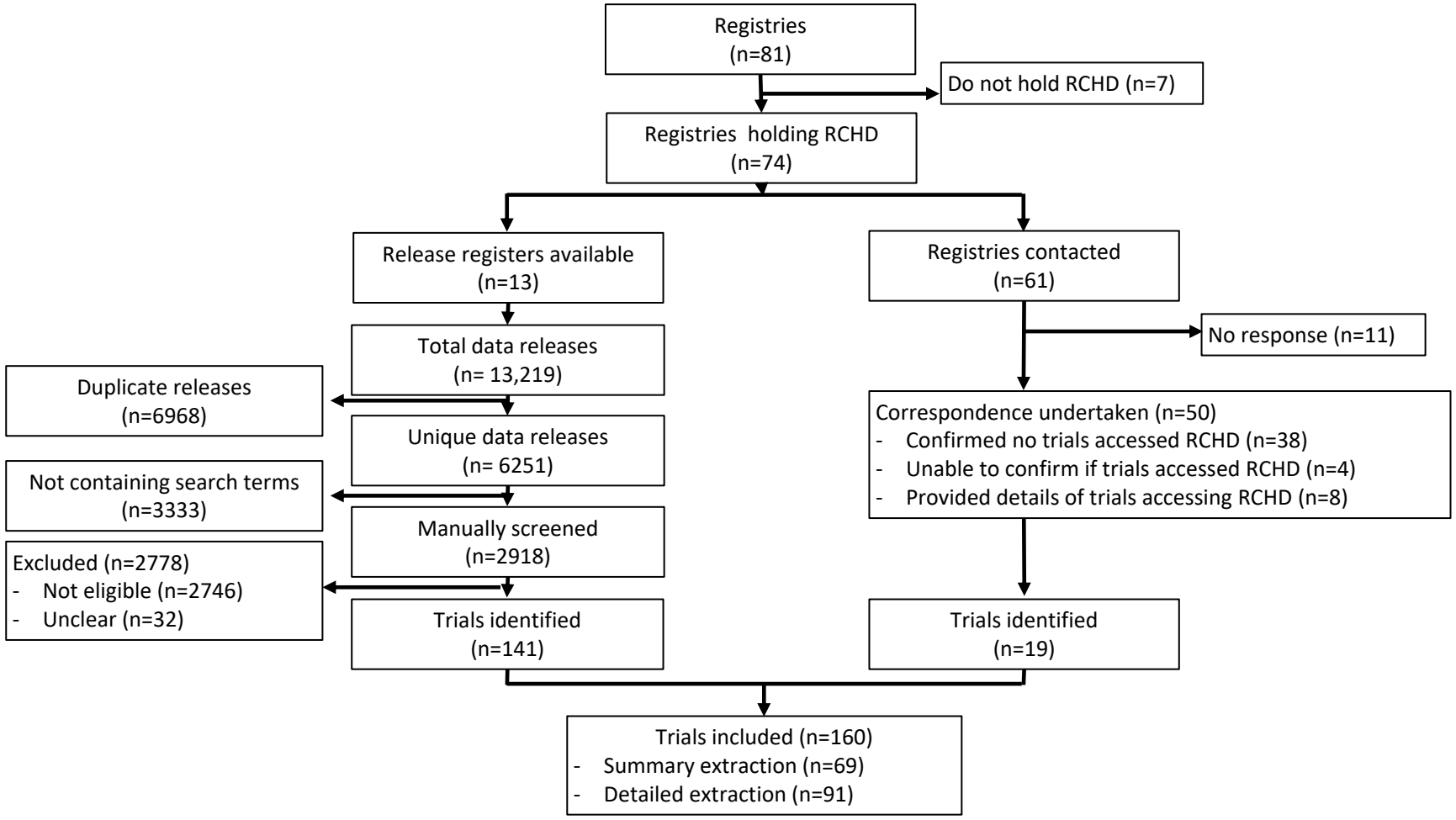
Systematic review: Methods 1

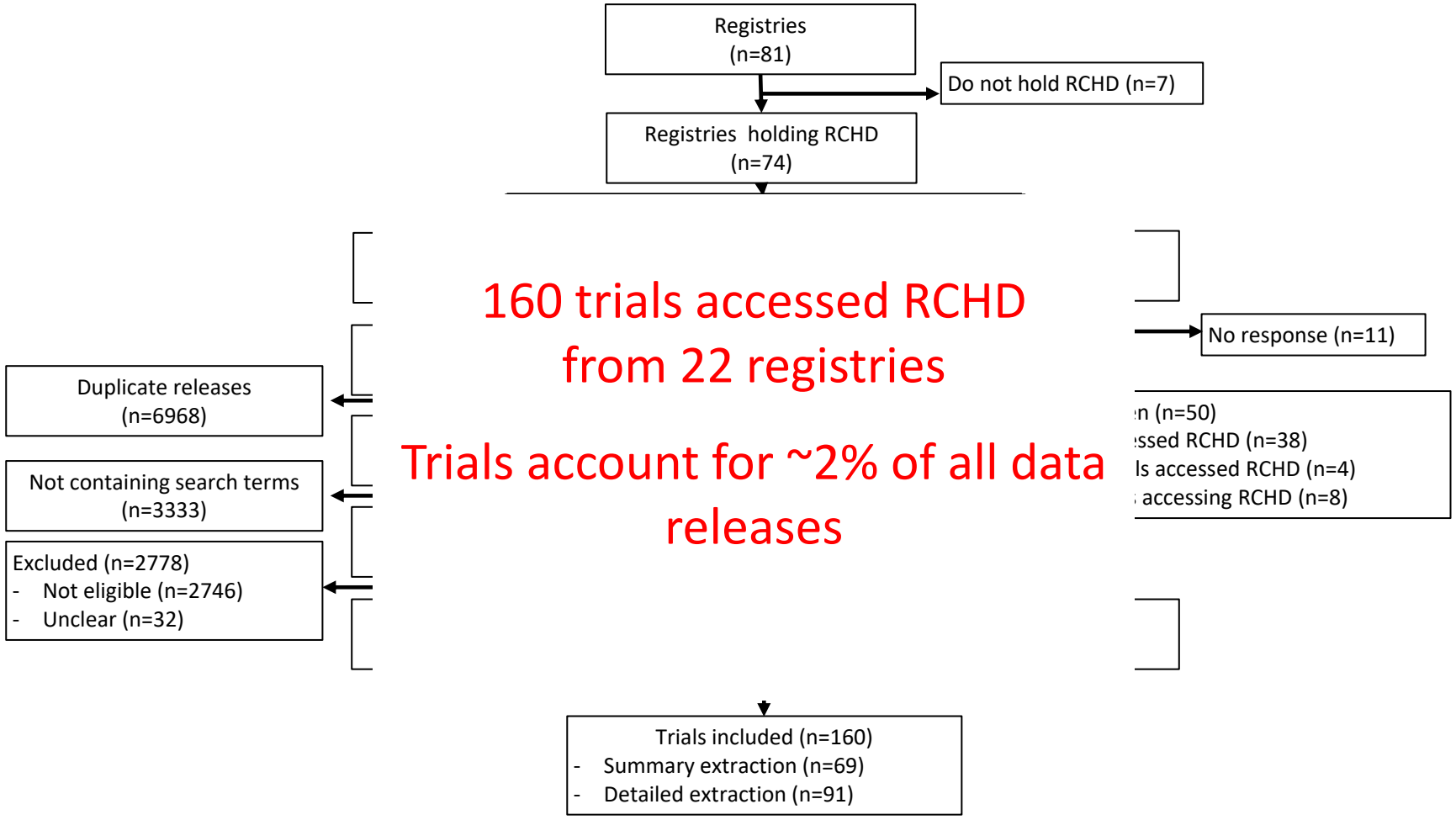
- Develop list of UK RCHD sources (registries)
  - Excluding: cohorts, biobanks, records only held at the point of care (e.g. GP practice)
- Search for trials accessing these sources 2013-2018
- Trial eligibility
  - RCT (individual or cluster)
  - Accessing RCHD for participant data (baseline or outcome)

Systematic review: Methods 2

- Data collection
  - Detailed extraction for 2017-2018 releases
  - Duplicate extraction onto CRF and single data entry into Macro
- Identification of trial related material
  - Trial website (protocol, PIS/PICF, SAP etc)
  - Trial registration page
  - Trial results publications (incl supplementary material)

Results


Registries
(n=81)
Do not hold RCHD (n=7)
Registries holding RCHD
(n=74)
Release registers available
(n=13)
Registries contacted
(n=61)
Total data releases
(n= 13,219)
No response (n=11)
Duplicate releases
(n=6968)
Unique data releases
(n= 6251)
Correspondence undertaken (n=50)
- Confirmed no trials accessed RCHD (n=38)
- Unable to confirm if trials accessed RCHD (n=4)
- Provided details of trials accessing RCHD (n=8)
Not containing search terms
(n=3333)
Manually screened
(n=2918)
Excluded (n=2778)
- Not eligible (n=2746)
- Unclear (n=32)
Trials identified
(n=141)
Trials identified
(n=19)
Trials included (n=160)
- Summary extraction (n=69)
- Detailed extraction (n=91)


Registries
(n=81)
Do not hold RCHD (n=7)
Registries holding RCHD
(n=74)
160 trials accessed RCHD from 22 registries
Trials account for ~2% of all data releases
No response (n=11)
n (n=50)
ssed RCHD (n=38)
ls accessed RCHD (n=4)
accessing RCHD (n=8)
Duplicate releases
(n=6968)
Not containing search terms
(n=3333)
Excluded (n=2778)
- Not eligible (n=2746)
- Unclear (n=32)
Trials included (n=160)
- Summary extraction (n=69)
- Detailed extraction (n=91)

Registry	Total Trials n=160
NHS-Digital	108 (68%)
ISD-Scotland	35 (22%)
Public Health England (PHE)	15 (9%)
SAIL	9 (6%)
Intensive Care national Audit and Research centre (ICNARC)	7 (4%)
NHS Wales	7 (4%)
Paediatric Intensive care Audit Network (PICANet)	6 (4%)
Clinical Practice Research Database (CPRD)	4 (3%)
NHS Blood and Transplant (NHSBT)	3 (2%)
Trauma audit and Research Network (TARN)	3 (2%)
National Emergency Laparotomy audit (NELA)	2 (1%)
Neonatal Research Database (NNRD)	2 (1%)
Public Health Wales (PHW)	2 (1%)
UK Renal Registry (UKRR)	2 (1%)
ResearchOne	2 (1%)
Other	7 (7%)

Datasets accessed

	Death	Cancer registration	Hospital visits	Other datasets
Trials (2017-2018) N=91	69 (76%)	29, (32%)	50 (55%)	26 (29%)

Trial characteristic	Total (N=160) (n, %)
Randomisation	
Individual	136 (85%)
Cluster	24 (15%)
Design	
Screening	16 (10%)
Treatment	116 (73%)
Primary Prevention	28 (18%)
Setting	
Primary Care	41 (26%)
Secondary Care	119 (74%)
Intervention	
CTIMP	76 (48%)
Surgical	13 (8%)
Other	71 (44%)
Trial size	median 1590 range 41-6,000,000
1-500	41 (26%)
500-5000	74 (46%)
>5000	42 (26%)
Disease Category	
Cancer	47 (29%)
Cardiovascular/stroke	46 (29%)
Pregnancy/childbirth	9 (6%)
Mental health	12 (8%)
Infection	8 (5%)
Endocrine and diabetes	4 (3%)
Inflammatory disorder	5 (3%)
Other	29 (18%)
Coordinated by Registered CTU	
Yes	92 (57%)
No	29 (18%)
Unclear	39 (24%)
International	
Yes	32 (20%)
No	125 (78%)
Highest impact factor journal (n=91)	
BMJ	2 (2%)
JAMA	6 (7%)
Lancet	16 (18%)
NEJM	3 (3%)
Other	8 (9%)
Not published	56 (62%)

Comparison to HRA (2015 approvals)

Recruitment start date range: 1979-2018

	RCTs accessing RCHD (n=160)	HRA in 2015 (n=963)
Primary care	41 (26%)	48 (5%)
Secondary care	119 (74%)	846 (95%)
Therapeutic area		
Cancer	47 (29%)	168 (17%)
Cardiovascular and stroke	46 (29%)	121 (13%)
Pregnancy and childbirth	9 (6%)	30 (3%)
Infection	8 (5%)	55 (6%)
Drug trial	76 (48%)	515 (53%)
Cluster trial	24 (15%)	29 (3%)
Feasibility/pilot	17 (11%)	177 (18%)
Sample size (median, range)	1590 (41 - 6,000,000)	275 (6 - 30,000)
UK only	125 (78%)	450 (50%)
International trials	32 (20%)	443 (50%)

	Total N=91
RCHD only	52 (58%)
Cross-checking self-reported data	29 (32%)
Cross-checking trial data	28 (31%)
Cost-effectiveness	25 (28%)
Trigger case-review	22 (24%)
Methodology	11 (12%)
RCHD cross-checking	9 (10%)
Unclear	13 (14%)

Summary of findings

- 22 registries have provided data to 160 trials (2013-2018)
  - Small proportion of data releases (2%)
  - Small proportion of UK trials (3%)
- Commonly: large cancer or cardiovascular trials
- 2/3 of trials accessed data from NHS Digital
- Data use to inform outcomes varied substantially

- Review of trials using RCHD
- Comparison of trial and registry data
- The changes due to COVID
- Future

Trial and RCHD comparison for death data

- Barry et al 2013
- Herrington et al 2015
- Submitted a comparison of BOSS trial and RCHD
- SWAT 125: Comparison of trial-collected and routinely-collected death data [Available from: https://www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/FileStore/Filetoupload,976743,en.pdf]

- Review of trials using RCHD
- Comparison of trial and registry data
- The changes due to COVID
- Future

Acceleration of changes in 2020

- RCHD for outcomes
- RCHD available in weeks
- RCHD retention

- Review of trials using RCHD
- Comparison of trial and registry data
- The changes due to COVID
- Future

Future

- RCHD will become useable for outcomes
- Improvement in the application process
- RCHD will be clarified as source data
- RCHD will be kept with trial data

Use RCHD efficiently for all trials

How much are we using routinely collected electronic h in t K?

Thank You

Shar
Associ essor Trial Conduct Methodology
MRC Clinical Trials Unit at UCL
2020

References

Lensen S, Macnair A, Love SB, Yorke-Edwards V, Noor NM, Martyn M, Blenkinsop A, Diaz-Montana C, Powell G, Elizabeth Williamson E, Carpenter J and Sydes MR. Access to routinely collected health data for clinical trials – review of successful data requests to UK registries. Trials 2020, 21:398

Barry SJE, Dinnett E, Kean S, Gaw A, Ford I, Are Routinely Collected NHS Administrative Records Suitable for Endpoint Identification in Clinical Trials? Evidence from the West of Scotland Coronary Prevention Study. PLoS ONE, 2013. 8(9)

Herrington W, Wallendszus K and Bowman L, Can vascular mortality be reliably ascertained from the underlying cause of death recorded on a medical death certificate? Evidence from the 2800 adjudicated heart protection study deaths. Trials 2015, 16(Suppl 2):P61

SWAT 125: Comparison of trial-collected and routinely-collected death data [Available from: https://www.qub.ac.uk/sites/TheNorthernIrelandNetworkforTrialsMethodologyResearch/FileStore/Filetoupload,976743,en.pdf]

Questions?

Please use the message box to type your questions

TMRP Webinar Series:

"*Use of routine data in trials in the UK*"

20/08/2020

Use of routinely collected outcome data in a UK cohort of publicly funded randomised clinical trials

Andrew McKay

Liverpool Clinical Trials Centre, University of Liverpool

a.mckay@liverpool.ac.uk

Acknowledgements to collaborators

- **Prof. Paula Williamson**
  - MRC North West Hub for Trials Methodology Research, Department of Health Data Science, University of Liverpool, Liverpool, UK
- **Prof. Andrew Farmer**
  - Chair of the NIHR Health Technology Assessment Programme General Board
  - Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
- **Prof. Carrol Gamble**
  - Liverpool Clinical Trials Centre, University of Liverpool, a member of Liverpool Health Partners, Liverpool, UK
  - MRC North West Hub for Trials Methodology Research, Department of Health Data Science, University of Liverpool, Liverpool, UK
- **Dr. Ashley Jones**
  - Liverpool Clinical Trials Centre, University of Liverpool, a member of Liverpool Health Partners, Liverpool, UK

Introduction

- Later phase clinical trials are expensive – increased focus on methodology to support innovative and efficient delivery.
  - Collection of consistent and reliable data is still required.
- Many sources of routinely collected health data (RCHD).
  - E.g. medical records, registries and hospital activity data.
- Progress in achieving connectivity, data linkage and security.
- Extent of RCHD sources being used to deliver efficient clinical trials is unclear.

Example of recent evidence of clinical trials using RCHD for research purposes (Fitzpatrick *et al.*, 2018) (I)

- A scoping review:
  - RCTs extended by record linkage to enable long-term follow-up.
  - Explore additional insights into the long-term treatment effects and harms of treatment.

- 113 trials identified:
  - 1945-2016 with 1-50 years additional follow-up.
  - Nordic countries (43%), USA (23%), UK (22%).

- Outcomes: Mortality (78%), cancer (36%), cardiovascular events (33%).

Example of recent evidence of clinical trials using RCHD for research purposes (Fitzpatrick *et al.*, 2018) (II)

- 48% with statistically significant treatment effects in trial extension phase.
  - 28% of these showed treatment effects significant only in this period.
- 11% with statistically significant harms in trial extension phase.
  - 88% of these showed harms significant only in this period.
- Key finding: Some treatment benefits extend beyond the trial and some treatment harms only become apparent after the trial is complete.
- Shows value of long-term follow-up facilitated by RCHD.
- The authors "*recommend that researchers routinely request permission from trial participants to study long-term treatment effects using linkage to RCHD*".

Study aim

- Study aim: To ascertain current practice amongst a United Kingdom (UK) cohort of recently funded and ongoing randomised controlled trials (RCTs) in relation to sources and use of routinely collected outcome data.

- We define RCHD to be data collected without specific a priori research questions developed prior to using the data for research.

- UK cohort: National Institute for Health Research Health Technology Assessment (NIHR HTA).

Methods – Inclusion criteria

- The following inclusion criteria were used:
  1. Ongoing RCT of any type including feasibility or pilot work, funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme;
  2. availability of a protocol; and
  3. use of RCHD for at least one study outcome.

Methods – Searching

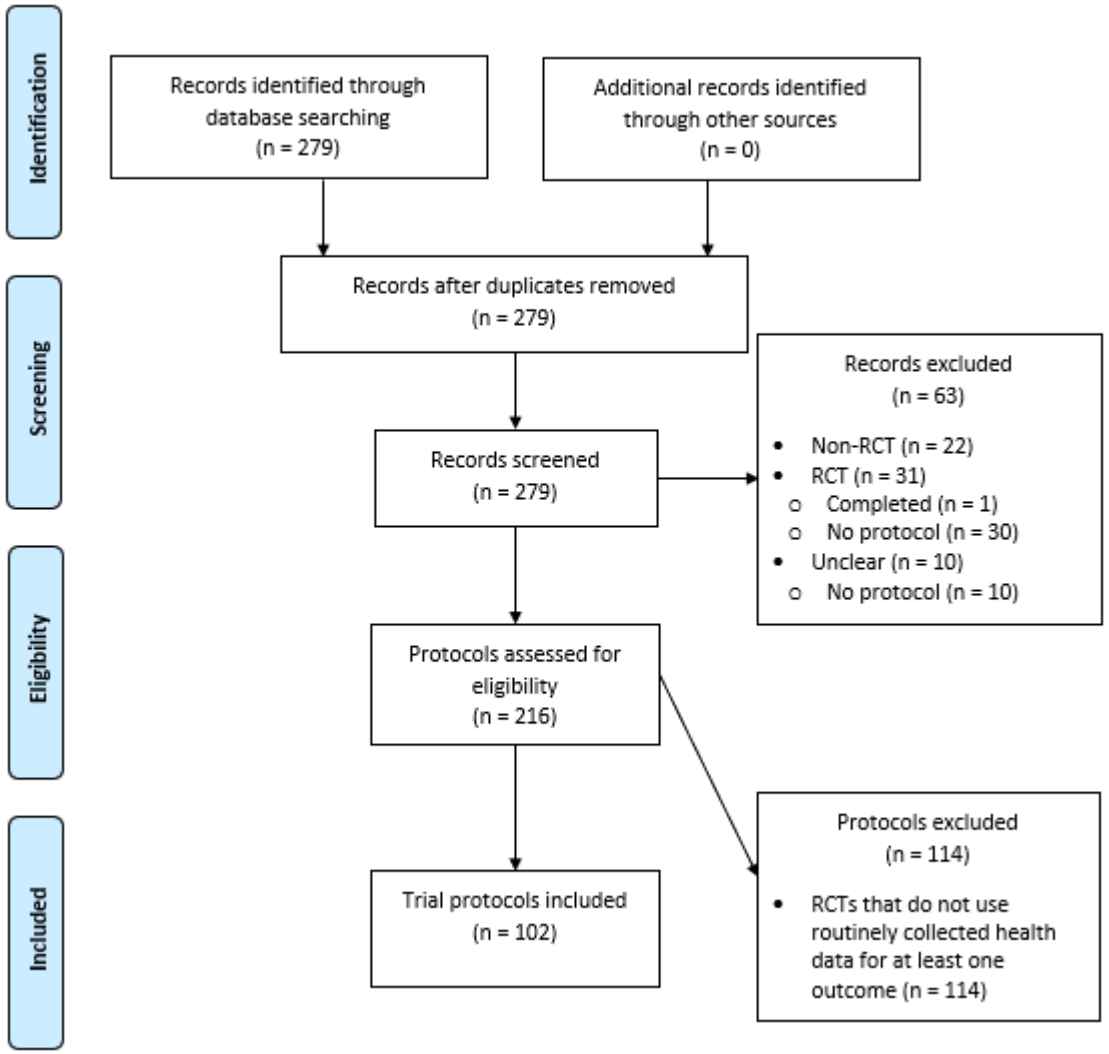
- A search of the NIHR Journals Library* was undertaken to find protocols registered as of 25/10/2019. The search fields and terms used were:
  - Search term: ‘Random’.
  - Research type: ‘Primary research’.
  - Programme: ‘HTA’.
  - Status: ‘Research in progress’.
- In the absence of a protocol, the study was excluded.
- For studies with multiple protocol versions, the most recently available version was used.

* https://www.journalslibrary.nihr.ac.uk/advancedsearch/

Methods – Data extraction - RCHD and outcomes

- Any details of data quality assessment of RCHD source prior to use.
- RCHD source name.
- Reasons for wanting outcome data from RCHD source.
- Specific outcomes and outcome type from named RCHD sources.

Results – PRISMA flow diagram [Figure 1]



Results – PRISMA flow diagram key result

- Of 216/279 (77%) NIHR HTA trials with a protocol available for further study:
  - 102/216 (47%) planned to use RCHD for at least one outcome.

Results – Reasons for sourcing outcome data from RCHD sources in 102 studies [Table 1 (I)]

	Categories (Multiple categories can apply to a single study)	Total
(1)	Supplementing data collection for withdrawn and/or lost-to-follow-up patients.	18
(2)	Supplementing data collection for unobtainable/missing data.	3
(3)	As the sole source of all outcome data.	0
(4)	As the sole source of some outcome data.	43
(5a)	As a source of some outcome data, alongside other sources for the same outcome data (e.g. CRF).	51
(5b)	As a source of some outcome data, but collected by CRF if unable to access data.	3

Results – Reasons for sourcing outcome data from RCHD sources in 102 studies [Table 1 (II)]

	Categories (Multiple categories can apply to a single study)	Total
(6a)	Registry trial*: As the sole source of outcome data with purpose-built Module to collect remaining outcome data.	1
(6b)	Registry trial*: All outcome data collected through multiple RCHD sources except for questionnaire data.	1
(6c)	Registry trial*: All outcome data collected through multiple RCHD sources except for some baseline data, questionnaire data and other patient-reported data.	1

** A registry trial is a RCT conducted using clinical observational registries as the main source of outcome data collection.*

Results – Reasons for sourcing outcome data from RCHD sources in 102 studies [Table 1 (III)]

	Categories (Multiple categories can apply to a single study)	Total
(7a)	RCHD compared to trial collected data as part of feasibility assessment criteria or as a secondary outcome.	14
(7b)	Representativeness of randomised patients compared with all eligible patients using RCHD as part of feasibility assessment criteria.	1
(8)	Participants flagged with NHS Digital/other: Check health status/notification of any deaths, causes or check health status of patient prior to contacting in case patient has died.	14
(9)	Set up mechanisms for long-term follow-up.	4
(10)	Patients asked to provide written consent for continuation in the study once have regained capacity. Those who prefer not to be actively involved in the study follow-up, then asked to provide consent to using their routinely collected NHS data.	1
	Total (across 3 slides)	155

Results – Reasons for sourcing outcome data from RCHD sources in 102 studies – Summary

- RCHD sole source of outcome data for at least one outcome in 46/102 (45%).
- Reference to prior feasibility work confirming aspects of RCHD source data quality in 5/46 (11%). [See next slide for further details]
- 14/102 (14%) will assess feasibility to use RCHD sources during trial, although specific details were often lacking.

Results – Prior feasibility work to assess RCHD source prior to use for RCT

Reference	RCHD source data quality assessment
Goldberg (2013)	A1 minimal data set submitted routinely for all Total Ankle Replacements (TAR) to National Joint Registry (NJR).
Blackwood (2017)	Paediatric Intensive Care Audit Network (PICANet) data: • "*validated on entry and centrally*".
Mouncey (2017)	Intensive Care National Audit & Research Centre (ICNARC): • "*source of high quality, robust and representative data*".
Benger (2014)	Undertook a separate feasibility study prior to trial: • Compared collected data to Hospital Episode Statistics (HES) data. • Obtained complete data sets from routinely collected data for >95% of patients. • Recommended HES data for use in the main trial.
Griffin (2014)	Undertook a separate feasibility study prior to trial: • Tested two potential primary outcome measures (NAHS and iHOT-33). • Found both easy to use and acceptable to patients. • Chose iHOT-33 because it is the principal outcome measure for the UK Non-Arthritic Hip Registry.

Results – Categories of RCHD sources of outcome data in 46 studies where this was the sole source for at least one outcome [Table 2]

Source (Study level)	Number (%)
(i) Primary care data (all regional equivalents)	8 (17%)
(ii) HES (and/or regional equivalents)	27 (59%)
(iii) ONS (and/or regional equivalents)	27 (59%)
(iv) Data collected specifically for patient group or healthcare intervention (to include patient registries, ICNARC, ambulance service data, etc)	26 (57%)
(v) Other	5 (11%)

Discussion

- 45% of UK publicly funded trials plan to collect outcome data from RCHD sources.
  - Another cohort of 189 RCTs published since 2000 mainly in USA found to this figure to be 8% (McCord *et al.*, 2019).
- Very few trial teams described any assessments of data quality from RCHDs in the protocol.
  - Work ongoing on a CONSORT extension to determine if this should be reported in a trial publication (Kwakkenbos *et al.*, 2018) – soon to be published.
- Work ongoing: SPIRIT guidelines extension for trials using RCHD is being discussed.
  - As a minimum, we recommended trialists provide evidence of RCHD source data quality in a funding application.
- Future work: Follow cohort up to see if able to collect outcome data as planned.

Data availability

- Figshare: Use of routinely collected data in a UK cohort of publicly funded randomised clinical trials. https://doi.org/10.6084/m9.figshare.12185193 (McKay *et al.*, 2020).
- This project contains the following underlying data:
  - Data Set 1 (Study identifiers and raw data used for Figure 1 – PRISMA flow diagram)
  - Data Set 2 (Raw data used for Table 1)
  - Data set 3 (Raw data used for Supplementary Table 1)
  - Data set 4 (Raw data used for Table 2)
  - Data set 5 (Raw data showing details of outcomes using data from RCHD sources)
- This project contains the following extended data:
  - Supplementary Table 1 - EHR sources of outcome data v1.0.pdf.
- Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

References

- McKay AJ, Jones AP, Gamble CL, Farmer AJ, Williamson PR. Use of routinely collected data in a UK cohort of publicly funded randomised clinical trials [version 2; peer review: 1 approved]. *F1000Research* 2020, 9:323 (https://doi.org/10.12688/f1000research.23316.2)

- McKay A, Jones A, Gamble CL, Farmer AJ, Williamson PR: Data sets used and Supplementary Table 1. figshare. Dataset.2020. http://www.doi.org/10.6084/m9.figshare.12185193.v1

- Fitzpatrick T, Perrier L, Shakik S, et al.: Assessment of Long-term Follow-up of Randomized Trial Participants by Linkage to Routinely Collected Data: A Scoping Review and Analysis. JAMA Netw Open. 2018;1(8):e186019. 30646311 10.1001/jamanetworkopen.2018.6019 6324362

- Kwakkenbos L, Juszczak E, Hemkens LG, et al.: Protocol for the development of a CONSORT extension for RCTs using cohorts and routinely collected health data. Res Integr Peer Rev. 2018;3:9. 30397513 10.1186/s41073-018-0053-3 6205772

- McCord KA, Ewald H, Ladanie A, et al.: Current use and costs of electronic health records for clinical trial research: a descriptive study. CMAJ open. 2019;7(1):E23–E32. 30718353 10.9778/cmajo.20180096 6375253

Questions?

Please use the message box to type your questions