ORINOCO
Tidying up outcomes in clinical trials

ORINOCO: Optimising Resource-use IN Outcome Collection

Sponsor

University of Aberdeen

Chief investigator

Prof Shaun Treweek

Co-investigators

Dr Katie Gillies, Dr Heidi Gardner, Mr David Pickles, Mr Adel El Feky, Ms Annabel Dawson

Funding

Funding secured from the Chief Scientist Office (CSO).

Location

Health Services Research Unit, Health Sciences Building, Foresterhill, University of Aberdeen

Project summary

Trials are one of the best ways of testing treatments, but they can be expensive and time consuming. The amount of data collected has a big influence on both cost and time.

We aim to understand how much time trial teams spend collecting the most important trial data (called primary outcomes) compared to the other data they collect (secondary outcomes). Outcomes are things like pain, blood pressure, or weight. Small-scale work suggests that trial teams spend most of their time on the less important outcomes. Our proposed large-scale work will find out whether this is correct. We also want to understand the time taken to collect core outcome sets–an agreed minimum amount of information–compared with trials that do not use them to see if they improve efficiency or worsen it.

Once we have the above, we will speak with trial teams and others involved in trials to understand what will help them to plan and fund their work more efficiently and also to develop guidance trial teams can use in the future. We hope our results will make it more likely that time isn't given to less important outcomes at the expense of the most important

Co-ordination: Health Services Research Unit, University of Aberdeen
Project page: https://www.abdn.ac.uk/hsru/what-we-do/research/projects/orinoco-826.php

Project start date: 1st March 2019
Anticipated interview start date: 1st February 2020
End date: 31st March 2021

Introduction

Trials are becoming more expensive[1]. Increasingly complex trial protocols contribute to this[2,3]: unlike many things in trials, this complexity is modifiable by trialists themselves.

Much of this complexity is a product of the outcomes selected and the data collected to report them. In addition to the work done by participants to provide data, the trial team must create data collection forms (called CRFs) and build a data management system to store them: in Aberdeen, building just the serious adverse events item of a CRF from a stock template takes 4.5 hours. It is not surprising that data collection is estimated to consume well over 30% of all work hours spent on trials[4].

How these hours of work are spent is important because participants, trial teams, funders and other trial stakeholders are more interested in some outcomes than others. The most important outcomes are called primary outcomes and the size of the trial is defined by the few (often just one) primary outcomes selected by the trial team. Future judgements as to whether the intervention works are largely framed around the primary outcomes. All other outcomes are, by definition, of less importance and are widely known as secondary outcomes.

There are no systematic reviews specifically focused on how data collection effort is distributed across outcomes. Some facts are known though. A large US study involving 15 pharmaceutical companies and 116 protocols found that for Phase III (i.e. later stage, definitive) trials, 7% of data collection items were linked to primary outcomes, 36% to key secondaries, 32% with basic medical history etc and 25% had nothing to do with the trial research questions[3]. The latter 25% were estimated to cost $3.7 billion annually in the US. A Cochrane systematic review comparing entries in trial registries to published trial reports found that 10% - 18% of primary outcome data and 44% of secondary outcome data were not published[5]. In a review of all trials submitted to a German ethics committee between 2000 and 2002, Kirkham *et al* found that only 47% of the two and half million items of outcome data collected from participants in these 308 trials were fully published[6]. Hind *et al* report growing global concern over escalating trial costs, concluding that UK Clinical Trials Units need to explain their funding requests, and that more research on barriers to implementing evidence-based strategies that minimise costs is needed[1]. A 2017 systematic review of trial resource use and costs found little empirical data and concluded that such data are urgently needed to improve value for money[7]. Heneghan *et al* give a catalogue of problems with data collection, including lack of relevance to decision-makers and poorly specified and collected data[8]. Overall the picture is one of many trials collecting more data than are needed to answer their research questions and not all collected data are published. Collecting unnecessary data is ethically questionable, wastes resources and can threaten the collection of essential data: longer forms certainly have lower response rates[9]. Data that languish unseen in investigators' cupboards is research waste writ large; they will not improve patient care. A more streamlined approach is needed.

There are two main barriers to streamlined data collection:

1. Knowing what to measure.
2. Poor awareness at the planning stage of how data collection impacts resource use.

Core Outcome Sets[10] – an agreed set of outcomes defined as the minimum to measure and report in a given type of trial – help to address the first problem and are recommended[11], though far from all conditions have them. The second barrier has received less attention and core outcomes could conceivably make it worse by contributing further time to data collection if trialists continue to collect their own outcomes of choice in addition to the core outcomes. We focus on the 2nd barrier but will look at the use of core outcome sets on workload.

Aims

Our overall aims are to:

1. Increase quantified awareness among trialists of how data collection effort is distributed.
2. Reduce research waste by increasing the chance that trial teams focus effort on important outcome data that can be collected, analysed and reported within the resources available.

Research questions

1. Across trials completed in the last five years, how much time is spent collecting primary outcome data compared to time spent collecting secondary outcome data?
2. What is the impact on overall data collection time of trialists using a core outcome set in their trial and how does this differ from trials not using core outcome sets?
3. What are trial stakeholder views on the best way to use the information and implement recommendations coming from 1 and 2 to improve trial data collection planning and the match between data collection work and resourcing?

Study design

There will be three phases:

- Phase 1: Identifying trials and outcomes
- Phase 2: Obtaining timings for each outcome
- Phase 3: Stakeholder consultation

Phase 1. Identifying trials and outcomes

We will randomly select 120 trials from Pubmed that meet the criteria below:

- Phase III trials only, feasibility studies for Phase III trials are excluded
- Trials can be in any disease area
- Conducted in any country
- Non-commercial (by non-commercial we mean trials that are not funded and run by commercial organisations such as pharmaceutical companies. A trial that is commercially-funded but run by an academic team would be eligible for inclusion.)
- Indexed by PubMed in the last 5 years

- Interventions can target any part of the trial (staff or patients)
- Trial interventions must aim to impact on a health-related outcome, trials looking at methodological outcomes are excluded.

There is one further consideration:

1. To ensure a spread of trials across different disease areas (e.g. oncology, neurology, infectious disease), we will add in another rule; a maximum of 15% of the trials selected can be from one broad disease category. Once 18 trials have been selected in any broad disease area, further trials in that disease area will be excluded.

To these 120 trials we will make two additions.

Firstly, we will add 20 trials (12.5% of our total) from the same period that have used a core outcome set. As mentioned earlier, we will do this because core outcome sets are recommended[12] but the impact on workload of mandating a set of outcomes for a trial is unclear. We will select a random sample of 2012-2016 trials from the 189 completed trials using the rheumatoid arthritis core outcome set (comprising 8 outcomes) in Kirkham *et al*[13]. This was the first such set so is more mature than those in other clinical areas; Kirkham has given us the list of trials he identified. If any of the other 120 selected trials used core outcomes, we will add them to our core outcome subset.

Secondly, we will add 20 public health intervention trials. We do this because there may be specific issues relevant to public health intervention trials, particularly around collection of multiple behavioural measures over extended periods.

We will create a search that specifically looks for public health interventions from the last 5 years and then randomly select 20 from that pool that adhere to the eligibility criteria below. We also expect to pick up some public health trials in our other searches (though not many) meaning our sample of public health intervention trials will be a little over 20.

Eligibility criteria for the public health trials:

- Trials must target a public health issue, i.e. a health issue that has the potential to impact the health of a significant proportion of the population
- Conducted in any country
- Non-commercial (definition as above)
- Indexed by PubMed in the last 5 years

There is one further consideration:

1. To ensure a spread of trials across different public health issues (e.g. vaccinations, health screening, obesity), we will add in another rule; a maximum of 25% of the trials selected can be from one broad category. Once 4 trials have been selected in any broad category, further trials in that category will be excluded.

With samples representing Phase III trials, public health trials and trials using a core outcome set, this gives a total of 170 trials, a balance between enough trials to provide compelling data and a sample that is unmanageably large. For each we will identify primary and secondary outcome measures and the measurement instruments used to collect them. We will obtain the protocol or trial registration entry for as many as we can to overcome (and report on) the potential problem of selective outcome reporting in trial reports.

Phase 2: Obtain timings for each outcome

Phase 2 will assign a 'Time to collect' to each outcome measure so that we can calculate a ratio of time distribution between primary and secondary outcome collection for each trial.

Data collection methods

To collect timing data, we will use a rubric of methods based on DP's learnings from the pilot study[14].

In order for us to make comparisons between outcome collection timing from both the developers' and the users' perspective, we will collect data from the users' perspective using the following methods:

- Primary contact with the trial team involved in measuring the outcome – using a standard template, we will email the corresponding author listed on the published paper, as well as the trial's Chief Investigator and the Trial Manager where these details are available. A reminder email will be sent two weeks after the initial contact if the team have not responded.
- Web-based resources such as the Shirly Ryan Ability Lab Rehabilitation Measures Database (https://www.sralab.org/rehabilitation-measures).
- A crowd-sourcing approach – we aim to exhibit at conferences that have a good mix of trialists, healthcare professionals engaged with conducting trials, and patients (EBMLive 2019 and ICTMC 2019), to collect outcome timing data from a range of healthcare professionals, researchers and trial managers.

For every trial we will also aim to collect data from the original paper detailing development of the outcome measure to provide us with timing from the developers' perspective.

Where different trials have used the same outcome measure (e.g. two trials use SF-36 to measure quality of life) we will aim to get timings from each trial team. This will give a 'Time to collect' range for that outcome, which acknowledges that there is a degree of uncertainty regarding timings. Moreover, trials may use the same outcome measure but administer it differently. SF-36 for example could be completed face-to-face with a member of trial staff, or as self-report. These modes may lead to different timings.

Collecting data from trials using core outcome sets

Although trials using a core outcome set might be expected to use the same outcome measures, this is not the case: only the outcome is specified (i.e. the 'what'), not how to measure it (i.e. the 'how'). Trials using the same core outcomes may have the same outcomes (e.g. fatigue) but use different

measurement instruments (e.g. the Multidimensional Assessment of Fatigue measure and a 10cm visual analogue scale) to measure outcomes. It may be possible to use a core outcome set but reduce workload by choosing a particular way of measuring the outcome. The 'Time to collect' data for trials using core outcome sets will be obtained in the same way as for other trials.

Collecting data for composite outcomes

It is likely that composite outcomes (outcomes in which a number of individual outcomes (e.g. a number of different serious morbidities) are combined to produce a single outcome (e.g. overall serious morbidity)) will have been collected in more than a few of the trials in our sample. We will include these outcomes and collect data for them using the same rubric as detailed above.

Adjusted outcomes

It is likely that adjusted outcomes (outcomes that have validated measurement tools that have been modified or adjusted in some way or used in a 'non-standard' way) will have been collected in more than a few of the trials in our sample. We will include these outcomes and collect data for them using the same rubric as detailed above, though the fact that an outcome measure has been adjusted may limit the sources that we can use to collect timing data. If we are able to collect data on a substantial number of adjusted outcomes, we will aim to make comparisons between time spent collecting adjusted and non-adjusted outcomes.

Repeated outcome measurements

We will also record the number and timing of measurement points. For example, a trial may have a baseline measurement point and a primary measurement point of 12-months but also have two intermediate measurement points at three and six months. These extra measurement points add work but it is not always clear how they contribute to answering the trial research question, though they can substantially increase data collection workload. During outcome timing data collection, we will be mindful the participants in different arms of the trial may have different experiences so that the data collected accurately reflects the experiences of trial participants (e.g. certain outcomes may be collected in one group and not the other).

Minimum outcome collection times

We will not assign a minimum time spent on any data collection item as outcomes are likely to have been measured numerous times; assigning an arbitrary minimum time would therefore not only introduce potential inaccuracy once, but repeatedly.

Timing ranges

In instances where timing details are provided as a range (e.g. a blood draw usually takes between 5 and 10 minutes), we will use the mid-point of the two timings provided (in this case, 7.5 minutes).

Combined timings

In instances where trial teams are unable to provide individual timings for outcomes, where data for multiple outcomes were collected within a single clinic appointment for example, we will divide the time by the number of outcomes collected. For example, if a trial team reports that a clinic appointment takes 45 minutes and involves the collection of data for one primary outcome and four secondary outcomes, we will assign each outcome with a data collection time of 9 minutes.

Where possible we will use timing data collected from other trial teams to improve the accuracy of these figures, e.g. if timing data for the primary outcome in the example above (let's call it Trial A) has already been collected in another trial in our portfolio (Trial B), and Trial B's team has estimated a time to collect of 15 minutes, then Trial A's timing data would be assessed as:

- 45 minutes for one primary outcome and 4 secondary outcomes
- 45 minutes minus 15 minutes (primary outcome) = 30 minutes (for 4 secondary outcomes)
- 30 minutes divided by 4 (secondary outcomes) = 7.5 minutes per secondary outcome.

Assessing data missingness

The CSO reviewers of the 1st stage of this proposal asked if an assessment of data missingness could be considered. This is an interesting suggestion because the worst possible outcome choice would be one that was both time-consuming to collect and 'lossy', with data often missing in whole or part. Missingness is poorly reported, which means a comprehensive look at this issue would add a substantial amount of work to the proposal. Instead, we will be more exploratory, which will inform future work. We will:

- Rank *secondary* outcomes by frequency of use and 'Time to collect' for our 170 trials.
- Select the top 10. These are frequently used, time-consuming, secondary outcomes.
- Attempt to assess missingness for these outcomes (proportion wholly missing, as well as proportion 100% complete) in trials using them. This will involve looking at the protocol, trial report(s) and contacting trial teams.

In addition to feasibility information about our method, this gives missingness data for outcomes that are a) used often b) time-consuming c) generally underpowered anyway.

Phase 3: Stakeholder consultation

The final phase will involve using the data from Phases 1 and 2 to develop two outputs:

1) Stakeholder views on the best way to use these data to improve the trial efficiency
2) Guidance that funders and others can give to applicants to improve trial planning.

In total we will talk to around 50 trial stakeholders and we will do this in two ways. Firstly, we will have one-to-one interviews with around 35 stakeholders (e.g. chief investigators, sponsor representatives, funders, PPI partners) to explore their views on how 'Time to collect' data can be used to improve trial planning. Secondly, we will talk to 3 trial teams with low primary: secondary outcome ratios and 3 with high ratios (over 1:6, the median from our pilot) to explore how they selected their outcomes and how 'Time to collect' data might have helped them to plan. For all of this work we will purposively look for a diverse range of funders/trial speciality/types of intervention. With both groups we will ask about how they make decisions regarding the number and timing of data collection points, which will help to tease out the reasons behind why trialists collect data outside baseline and the primary measurement point.

The approximately 35 individuals needed for the one-to one interviews will be identified through existing contacts (e.g. the CTU network, UK Trial Managers' Network, R&D contacts, funding panels, Trial Forge, PPI partner organisations). Some interviews will be done by telephone or Skype, depending on location and we will develop a topic guide for the interviews to ensure we ask the same key questions of all interviewees, tailored to stakeholder group. The trial teams we speak to regarding the low and high primary: secondary outcome collection ratios identified from Phase 2 will be purposefully selected with regard to intervention, speciality, funder etc but will by necessity be somewhat constrained by geography. We expect most to be UK or Ireland-based because this will make organising the discussion with a team of individuals easier (e.g. we could visit). It will, however, be informative to involve some trials from outside the UK and Ireland to explore system-level differences that may affect outcome choice. In the UK, NIHR often specifies outcomes to be collected and they recommend core outcome sets. This may be less prevalent elsewhere so discussions outside the UK and Ireland are worth exploring.

We will take advantage of stakeholder engagement opportunities such the International Clinical Trials Methodology Conference in November 2019 and the UK Trial Managers' Network Annual meetings, both of which we regularly speak at. The UK CTU Directors' meeting and MRC Methodology Hub meetings also provide potential opportunities for some of the interviews mentioned above, as well as more opportunistic engagement that will serve to increase our sample size for Phase 3. Our PPI partner Annabel Dawson, our trial retention prioritisation work with the James Lind Alliance, KGs' work with PPI Partners and HG's public engagement work will all provide help and opportunities for engaging with members of the public and patients, again potentially increase the number of views and perspectives contributing to Phase 3.

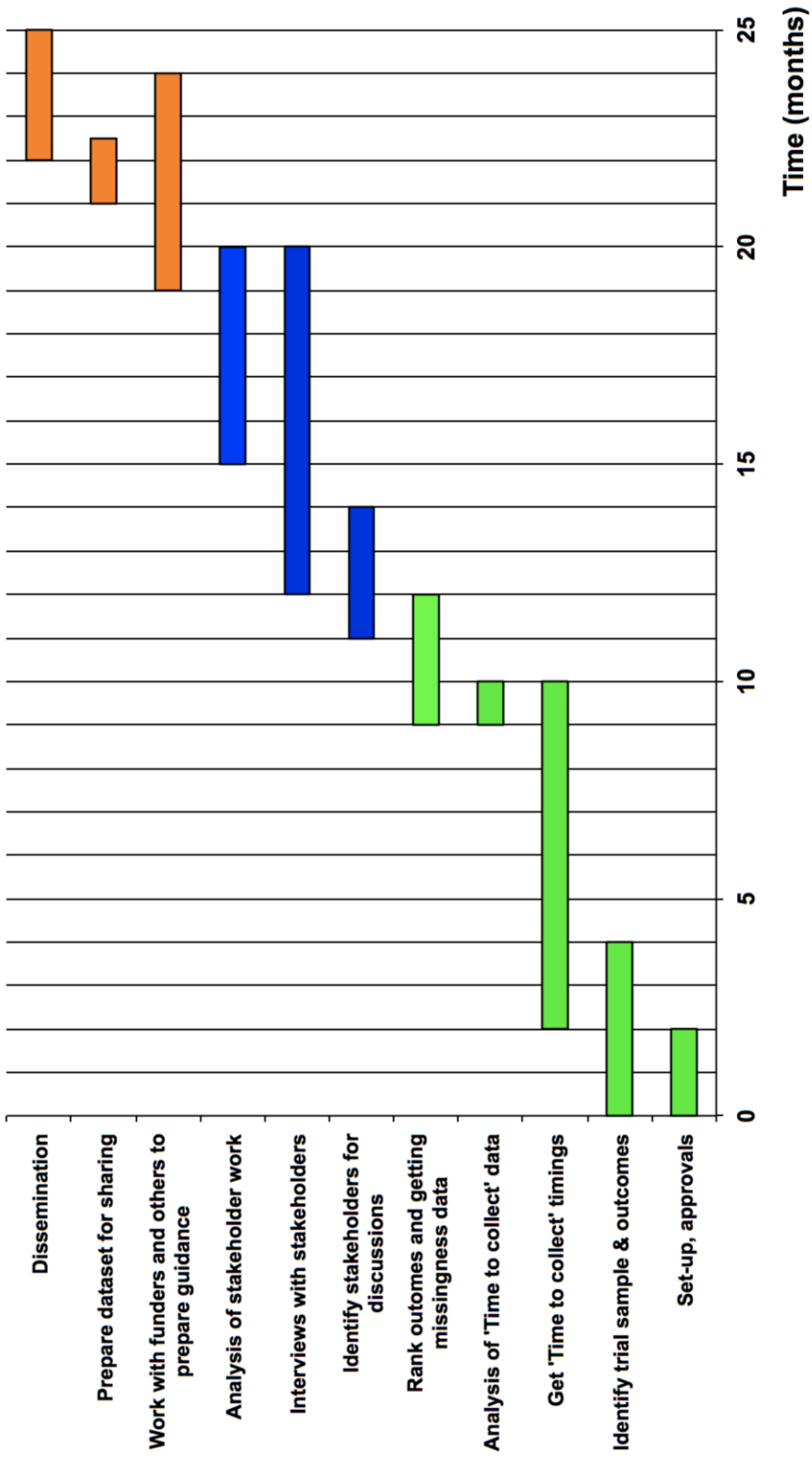
Moving to our guidance work, this will involve funders in the UK and Ireland, particularly the NIHR Evaluation, Trials and Studies Coordinating Centre (NETSCC) and the Health Research Board in Ireland. The Director of NETSCC, Matt Westmore, is on the Trial Forge Steering Committee and has agreed to work with us on this guidance. Oonagh Ward, Programme Manager–Infrastructure, Networks & Interventions at the Irish Health Research Board has also agreed to help. We would also like to involve CSO stakeholders.

To develop the guidance we will bring together a panel of around 15 - 20 people representing key

stakeholders. The guidance would use the data from Phases 2 and 3 and we will also summarise the data collection problem and current evidence. The guidance will be explicitly practical, meaning the intention is to produce something that is credible but easily acted upon. As an electronic supplement to the guidance, all the 'Time to collect' data we collect in Phase 2 will be made available publicly for sharing.

The primary purpose of the guidance will be to reduce the chances that a trial team a) collects unnecessary data b) chooses a way of collecting data that is less efficient than it could be and c) does not allocate enough resources to data collection in its budget. We anticipate that our funder partners will endorse and signpost applicants to this guidance.

Timetable


Time (months)
0
5
10
15
20
25
Dissemination
Prepare dataset for sharing
Work with funders and others to prepare guidance
Analysis of stakeholder work
Interviews with stakeholders
Identify stakeholders for discussions
Rank outomes and getting missingness data
Analysis of 'Time to collect' data
Get 'Time to collect' timings
Identify trial sample & outcomes
Set-up, approvals

Research governance

The study will be submitted to the University of Aberdeen College of Life Sciences and Medicine College Ethics Review Board for ethical review. The University has agreed to be Sponsor for the study, with Patricia Burns as our named contact in Research Governance.

Finance

This study has been peer reviewed and funding granted by the Chief Scientist Office (CSO) to the value of £224,213.00; grant agreement reference HIPS/18/04.

Data handling

Personal data

Any study data stored on laptops will be stored in encrypted files or folders and no personal identifiable data will be stored on these. Data will be transferred to a University server and deleted from the laptops as soon as the researcher is able to do so. Note that immediately transferring data to an encrypted laptop is more secure than leaving the audio recording on an un-encrypted audio device.

Direct quotes may be used in the publication of findings but these will not be attributed to named individuals and any identifiable information will be removed – we will also give participants the opportunity to withdraw their data from use in direct quotations.

All data will be stored securely on password protected University of Aberdeen computers with only the direct research team having access to this data.

Storage of data

No personal data will be explicitly collected for the research although audio recording of the interview may include personal data.

Electronic data (note there will be no identifiable data stored) will be stored on University networked drives in locked server rooms, or on an encrypted drive (e.g. encrypted drives of the laptop of the researchers, with encryption being set to at least AES-128). Any paper documentation linked to the study will be stored in locked cabinets in the Health Services Research Unit at the University of Aberdeen, or scanned and stored as per electronic data. The paper version will then be destroyed.

Study management

The study will be managed by the investigators listed above. Participants are able to contact any of the investigators for information about the study, or if they have queries. The study documentation will be available to the investigators on a shared file-space, and data stored securely for access by the research team only.

Author publication

This work will be disseminated at conferences such as Society of Clinical Trials or the UK Clinical Trials Methodology conferences. We anticipate that the results would also be published in a peer reviewed journal. In any publication resulting from this work, it will not be possible to identify individual participants or their place of work from quotations.

References

1. Hind D et al. Comparative costs and activity from a sample of UK clinical trials units. Trials 2017; 18:203.
2. Getz KA et al. Assessing the impact of protocol design changes on clinical trial performance. Am J Ther. 2008; 15: 450–7.
3. Getz KA et al. Quantifying the magnitude and cost of collecting extraneous protocol data. Am J Ther. 2015; 22: 117–24.
4. O'Leary E et al. Data collection in cancer clinical trials: Too much of a good thing? Clinical Trials 2013; 10:624–32.
5. Dwan K et al. Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database of Systematic Reviews 2011, Issue 1.
6. Kirkham JJ et al. How Much Participant Outcome Data Is Missing from Sight: Findings from a Cohort of Trials Submitted to a German Research Ethics Committee. PLoS ONE 2016; 11(6): e0157883.
7. Speich B et al. Systematic review on costs and resource use of randomised clinical trials shows a lack of transparent and comprehensive data. JCE 2017 (ePub ahead of print).
8. Heneghan C et al. Why clinical trial outcomes fail to translate into benefits for patients. Trials 2017; 18:122.
9. Edwards P et al: Follow-up by mail in clinical trials: does questionnaire length matter? Contr Clin Trials 2004, 25:31-52.
10. Williamson PR et al. The COMET Handbook: v1.0. Trials 2017; 18(S3):280.
11. NIHR 2017: http://bit.ly/2pfz4zA.
12. Fernie G et al. A good use of time and money? A study of how trial data collection effort is distributed across different categories of data. Trials 2017, 18 (Suppl 1): P239.
13. Kirkham JJ et al. A methodological approach for assessing the uptake of core outcome sets using ClinicalTrials.gov: findings from a review of randomised controlled trials of rheumatoid arthritis. BMJ Open 2017; 357: j2262.
14. Pickles D et al. Time well spent? A comparison of the work associated with collecting primary and secondary outcomes. Trials 2017, 18 (Suppl 1): P409.