What do we know about trial data collection?

Study	Approach	Outcome	Conclusions
A good use of time? Providing evidence for how effort is invested in primary and secondary outcome data collection in trials	Compared time spent collecting primary outcome data to the time spent secondary outcome data.	Median time spent on primaries – 56.1h Median time spent on secondaries – 190.7h	Trial teams should explicitly consider how long it will take to collect the data for an outcome and decide whether that time is worth it given importance of the outcome to the trial.
Using systematic data categorisation to quantify the types of data collected in clinical trials: the DataCat project	Categorised types of data collected across a range of trials and assessed what proportion of collected data each category represents.	Primary outcome data: 5.0%/11.2% Secondary outcome data: 39.9%/42.5% Non-outcome data (identifiers, demographic data): 32.4%/36.5%	Trialists should work to make sure that the data they collect are only those essential to support the health and treatment decisions of those whom the trial is designed to inform. Additional data may be considered wasteful in context of limited public funding for clinical research.
Getting it wrong most of the time? Comparing trialists' choice of primary outcome with what patients and health professionals want	Explored how important patients and healthcare professionals consider the outcomes measured in published trials.	The primary outcome ranked as most important outcome: 28% In additional study: no PPI was included in outcome selection (might help)	Trialists must consult with patients and healthcare professionals to identify the outcomes they will need to inform their future decisions about the usefulness of the intervention being tested.
Electronic Data Capture Versus Conventional Data Collection Methods in Clinical Pain Studies: Systematic Review and Meta-Analysis	Comparing electronic vs conventional (paper-and-pencil, telephone, in-person) data collection methods	Electronic data are equivalent or superior to conventional methods in terms of score equivalence, data completeness, ease, efficiency, and acceptability.	If the appropriate psychometric evaluations are in place, electronic data collection methods are a feasible means to collect pain data in clinical and research settings.

What do we know about trial data collection?

Study	Approach	Outcome	Conclusions
Methods for the Collection of Resource Use Data within Clinical Trials: A Systematic Review of Studies Funded by the UK Health Technology Assessment Program	Systematic review of data collection method resource use and costs. Methods examined were categorised as follows: • Medical records taken from routine primary and secondary care sources • Prospective forms completed by trial researchers or health-care professional • Prospective forms completed by trial researchers or health-based professional (based on patient recall) • Patient or carer-completed diaries • Patient completed or carer-completed forms Frequency of data collection was examined as well.	A checklist of good practice relating to economic data collection. List of practices: • Perspective • Identify resources for measurement • Data collection/ analysis plan • Resource use data collection • Baseline cost data • Piloting • Validation • Non-trial estimates of resource use • Method of costing • Unit Costs • Standardised reporting format Each practice has a corresponding recommendation, available in the conclusion of the paper.	Economic data collection is variable, even among a homogeneous selection of trials designed to meet the needs of a common organization (NHS).

What do we know about trial data collection?

Study	Approach	Outcome	Conclusions
Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature	Comparison of response rates of web surveys to alternative data collection methods	Response rate of web-based data collection 12.9% lower than alternative methods	Web-based data collection present lower RRs compared to alternative methods. However, it is not recommended to interpret this as a meta-analytical evidence due to the high heterogeneity of the studies.
Engaging Transgender People in NIH-Funded HIV/AIDS Clinical Trials Research	Describes improvements in the inclusion of transgender people. Regarding that collection: Two-step data collection forms of collecting sex at birth and gender identity as 2 independent variables. An explanation of "birth sex" for those unfamiliar with the distinction was added and among the gender terms, "other" was replaced with "additional category," as individuals do not think of themselves as "other."		

What do we know about trial data collection?

Study	Approach	Outcome	Conclusions
Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups	Literature review regarding the barriers to participation for members of socioeconomically disadvantages groups. One of the phases of research examined was data collection and measurement.	Data collection improvements: • Inclusive language and methods – simplify reading age of study materials, translate. • Flexible data collection methods – tailoring to participant circumstances (e.g., if can't be reached by phone -> door knock interviews/ online surveys) • Use of technology to gather data • Pilot testing measures – involving local community partners	To tackle the challenges of research with socially disadvantaged groups, and increase their representation in health and medical research, researchers and research institutions need to acknowledge extended timeframes, plan for higher resourcing costs and operate via community partnerships.

COMET Initiative – brings together people interested in the development and application of agreed standardised sets of outcomes (core outcome sets). These represent the minimum that should be measured and reported in all clinical trials of a specific condition.