



# Central statistical monitoring of investigator-led clinical trials in oncology

Marc Buyse<sup>1,2,3</sup> · Laura Trotta<sup>3</sup> · Everardo D. Saad<sup>4</sup> · Junichi Sakamoto<sup>5,6</sup>

Received: 25 April 2020 / Accepted: 14 June 2020 / Published online: 23 June 2020  
© Japan Society of Clinical Oncology 2020

## Abstract

Investigator-led clinical trials are pragmatic trials that aim to investigate the benefits and harms of treatments in routine clinical practice. These much-needed trials represent the majority of all trials currently conducted. They are however threatened by the rising costs of clinical research, which are in part due to extensive trial monitoring processes that focus on unimportant details. Risk-based quality management focuses, instead, on “things that really matter”. We discuss the role of central statistical monitoring as part of risk-based quality management. We describe the principles of central statistical monitoring, provide examples of its use, and argue that it could help drive down the cost of randomized clinical trials, especially investigator-led trials, whilst improving their quality.

**Keywords** Investigator-led trials · Central statistical monitoring · Data quality · Trial costs

## Introduction

Medical practice largely relies on the evidence generated by clinical trials, particularly randomized controlled trials (RCTs). These are considered the gold-standard approach for evaluating therapeutic interventions due to their capacity to allow for inferences about causal links between treatment and outcomes [1]. A general property of experimental research is that internal validity (i.e., the reliability of results) and external validity (i.e., their generalizability) tend to move in opposite directions in response to attempts to control trial features such as the population, the intervention,

and the assessment of outcomes. This gives rise to different attitudes towards clinical trials in general, and RCTs in particular: one that prioritizes internal validity (the explanatory attitude), and one that places more emphasis on the generalizability of results (the pragmatic attitude) [2]. Industry-sponsored trials, here defined as trials that aim to investigate experimental drugs with largely unknown effects, are typically characterized by an explanatory approach, which is suitable for the development of these novel agents or combinations. In contrast, investigator-led clinical trials, here defined as trials that aim to investigate the benefits and harms of treatments in routine clinical practice, are typically characterized by a pragmatic attitude. Table 1 characterizes some of the contrasts between an explanatory and a pragmatic approach to clinical trials. These contrasts have direct implications on the conduct of investigator-led trials, notably with regards to ways of ensuring their quality, which is the focus of this paper. Investigator-led clinical trials belong to a research area known as comparative-effectiveness research. We note that “real-world evidence” is a broader concept, given that it is often applied to observational research, something that falls outside the scope of our paper [1, 3].

## Why are investigator-led clinical trials needed?

Industry-sponsored clinical trials are essential for the development of new treatments. These clinical trials need to fulfil

✉ Marc Buyse  
marc.buyse@iddi.com

<sup>1</sup> International Drug Development Institute (IDDI),  
San Francisco, CA, USA

<sup>2</sup> Interuniversity Institute for Biostatistics and Statistical  
Bioinformatics (I-BioStat), Hasselt University, Hasselt,  
Belgium

<sup>3</sup> CluePoints, Louvain-la-Neuve, Belgium

<sup>4</sup> International Drug Development Institute (IDDI), 30 avenue  
provinciale, 1340 Ottignies-Louvain-la-Neuve, Belgium

<sup>5</sup> Tokai Central Hospital, Kakamigahara, Japan

<sup>6</sup> Epidemiological and Clinical Research Information Network  
(ECRIN), Kyoto, Japan

**Table 1** Explanatory versus pragmatic approach to clinical trials

Approach	Explanatory	Pragmatic
Type of trial	Industry-sponsored	Investigator-led
Primary purpose of trial	Regulatory approval	Public health impact
Patient selection	Fittest patients	All patients
Effect of interest	“Ideal” treatment effect	Actual treatment effect
Endpoint ascertainment	Centrally reviewed	Per local investigator
Preferred control group	Untreated (when feasible)	Current standard of care
Experimental conditions	Strictly controlled	Clinical routine
Volume of data collected	Large, for supportive analyses	Key data only
Data quality control	Extensive and on-site	Limited and central only

commercial interests and market expectations, which may not always address all patients’ needs [4]. Moreover, clinical trials that lead to the approval of novel drugs or devices often have shortcomings that have been recognized for decades. Such shortcomings include the strictness of the eligibility criteria, the choice of comparators, the effect size of interest, the choice of outcomes, and insufficient data on long-term toxicity [5]. Arguably, some of these shortcomings are a by-product of the general principles underlying marketing approval by regulatory agencies, such as the Japanese Pharmaceutical and Medical Devices Agency (PMDA), the European Medicines Agency (EMA), and the US Food and Drug Administration (FDA). These agencies must determine whether a new drug is sufficiently safe and effective to be made available for clinical use, which requires a careful assessment of the quality of the pivotal trial design, conduct, data and analysis whilst allowing safe and effective new drugs to enter the market quickly [6]. However, the need remains to generate additional, post-approval evidence on novel drugs or devices [6, 7]. Such evidence is required for clinical practice, as it provides a far better understanding of the effectiveness and safety of competing interventions in “real life”. Moreover, it allows the assessment of patients and settings not necessarily covered by the initial approval, thus leading to potential extensions of indications and refinement of the drug usage in patient subgroups. Even for newly approved drugs, many questions of clinical interest typically remain unanswered at the time of approval, including the duration of therapy, dose or schedule modifications that may lead to a better benefit/risk ratio, combinations of the new drug with existing regimens, and so on. Likewise, repurposing of existing drugs, whose safety and efficacy profile is well documented in other indications, is more likely to be attractive in the setting of investigator-led trials than to pharmaceutical companies for whom a given product ceases to be financially attractive towards the end of its life-cycle [8]. Finally, large, simple trials that address questions of major public health importance have been advocated for decades as one of the pillars of evidence-based medicine [9]. All in all, more and larger investigator-led trials are needed,

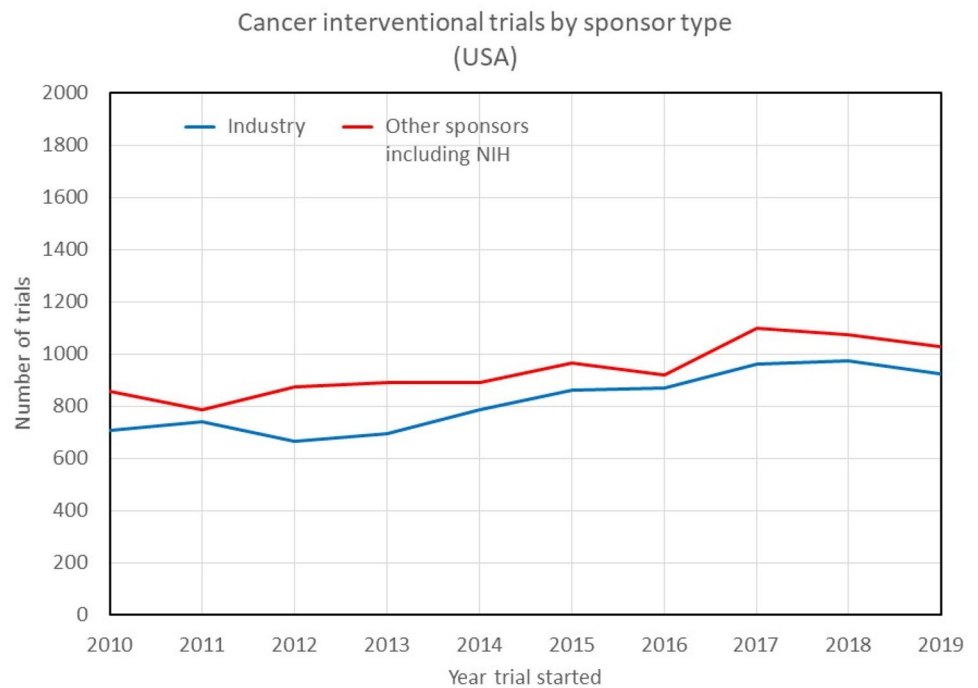
and it is crucially important to identify ways of conducting them as cost-effectively as possible [10, 11]. In particular, excessive regulation of investigator-led trials, using industry-sponsored trials as a model, is both unnecessary and counterproductive [11]. Taruno et al. (2019) have shown that, in Japan, the enactment of the Clinical Trial Act in March 2018 has resulted in a halving of the number of interventional cancer trials registered in the UMIN and jRCT clinical trial registries (Fig. X (Table 3) in [12]).

### Importance of investigator-led clinical trials in oncology

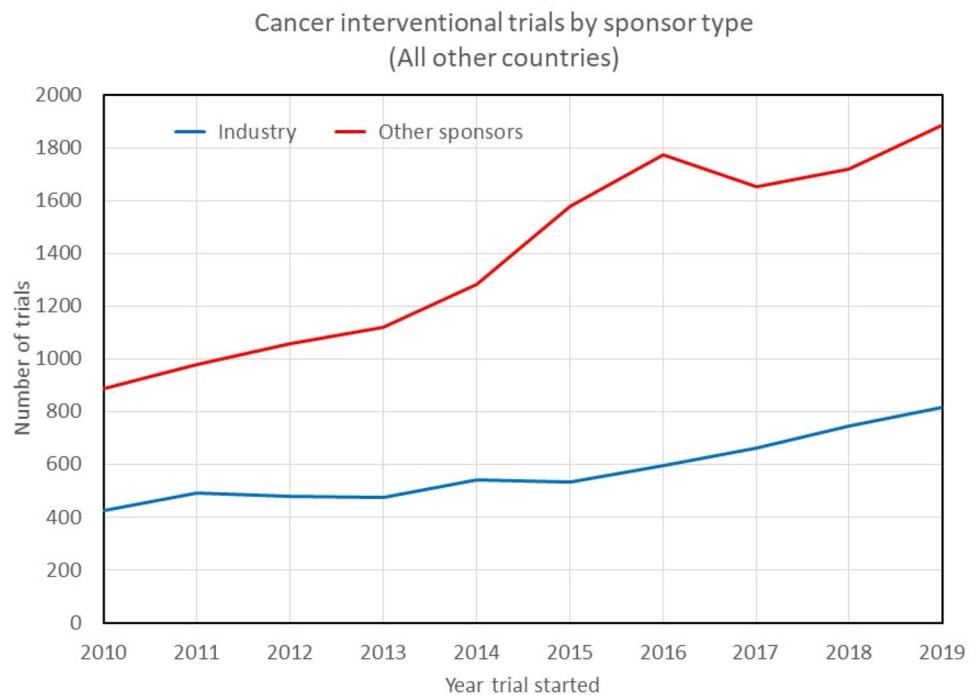
Publicly available clinical-trial registries are useful to assess the importance of investigator-led clinical trials in worldwide clinical research. The longest established and largest registry is ClinicalTrials.gov, with 333,529 trial protocols as of March 19, 2020. ClinicalTrials.gov contains trial protocols from both the US and other countries, and distinguishes between four major types of funders: (1) Industry (e.g., pharmaceutical and device companies), (2) the US National Institutes of Health, (3) other Federal agencies (e.g., FDA, Centers for Disease Control and Prevention, or Department of Veterans Affairs), and (4) all others (including individuals, universities, and community-based organizations). For the purposes of this paper, we focus on clinical trials conducted by sponsors other than the pharmaceutical and device industry, i.e., funder types (2)–(4), as opposed to funder types (1). We call these trials “investigator-led” clinical trials for simplicity.

Figures 1 and 2 show the number of registered interventional clinical trials in oncology, by funder type and year the trial started, in the US (Fig. 1) and all other countries (Fig. 2). In the US, about 2000 such trials were reported to have started in 2019, about 1000 being industry trials and about 1000 investigator-led trials (roughly half of which sponsored by NIH and other federal agencies, and half by other sponsors). In other countries, about 2600 such trials were reported in 2019, about 800 being industry trials versus about 1800 investigator-led trials (with

**Fig. 1** Number of cancer interventional trials by sponsor type and year trial started (USA)



**Fig. 2** Number of cancer interventional trials by sponsor type and year trial started (all other countries)



this latter number having increased fastest over time). There may be substantial under-reporting of clinical trials to ClinicalTrials.gov, especially for non-US trials and for investigator-led trials, so it is conservative to assume that investigator-led trials outnumber industry-sponsored trials worldwide. As such, investigator-led trials have the potential to generate much of the evidence upon which

the treatment of cancer patients is decided. Yet, as stated above, investigator-led trials may be under threat because of excessive regulation and bureaucracy, and the accompanying direct and indirect costs.

## Clinical trial costs

The rising costs of clinical trials have been a matter of major concern for some time [13]. The contribution of clinical trials to the overall costs of drug development is not known with precision, but recent estimates suggest that pivotal clinical trials leading to FDA approval have a median cost of US\$ 19 million; such costs are even higher in oncology and cardiovascular medicine, as well as in trials with a long-term clinical outcome, such as survival [14]. Interestingly, the cost of clinical trials was found to have huge variability, with more than 100-fold differences at the extremes of the cost distribution among the 138 trials surveyed [14]. The extent to which the skyrocketing costs of clinical research depend on individual components of clinical-trial conduct can vary substantially across trials, and likely when industry-sponsored studies are compared with investigator-led trials. In industry-sponsored trials, a great deal of resources are spent in making sure that the data collected in clinical trials are free from error. This is usually done through on-site monitoring (site visits) including source-data verification and other types of quality assurance procedures, alongside with centralized monitoring including data management and the statistical monitoring that is the focus of the present paper. While some on-site activities make intuitive sense, their cost has become exorbitant in the large multicenter trials that are typically required for the approval of new therapies [15]. It has been estimated that for large, global clinical trials, leaving aside site payments, the cost of on-site monitoring represents about 60% of the total trial [16].

## The effectiveness of central versus on-site monitoring

The high costs of monitoring could be justified if monitoring activities were likely to have an impact on patient safety or on the trial results [17]. Yet, there is no evidence showing that extensive data monitoring has any major impact on the quality of clinical-trial data, and none of the randomized studies assessing more intensive versus less intensive monitoring has shown any difference in terms of clinically relevant treatment outcomes [18–22]. Besides, there may also be a lack of effectiveness of sending large numbers of data queries to the centers as part of the data management process. In one limited study, only six queries were found (0.4% of 1395 queries) that might have influenced the results of three phase 1 cancer clinical trials, had the discrepancy not been revealed [23]. But without question, the most time-consuming and least efficient activity is source-data verification, which can take up to 50% of the time spent for on-site visits, hence it is especially important to make sure that such time is well spent. A large retrospective study of 1168 industry-sponsored clinical trials has shown that only 1.1% of all data

were changed as a result of source-data verification [24]. Moreover, it has been shown via simulations that random errors, which comprise most of the errors detected during source-data verification, have a negligible impact on the trial results [25]. In contrast, systematic errors (those that create a bias in the comparison between the treatment groups of a randomized trial) can have a huge impact on the trial results, but these types of errors can either be prevented or detected and corrected centrally [25, 26]. All in all, the monitoring of clinical trials needs to be re-engineered, not just for investigator-led trials, but also for industry-sponsored trials. To instigate and support this much-needed transition, regulatory agencies worldwide have advocated the use of risk-based quality management, including risk-based monitoring and central statistical monitoring (CSM) [27, 28].

## Risk-based quality management

The central principle of risk-based quality management is to “focus on things that matter”. What matters for a randomized clinical trial is to provide a reliable estimate of the difference in efficacy and tolerance between the treatments being compared. It is important to stress that the criteria to assess efficacy and tolerance may differ between industry-sponsored trials and investigator-led trials. For instance, in terms of efficacy, industry-sponsored trials often use the centrally reviewed progression-free survival (PFS), which may provide the most sensitive indicator of the antitumor effect of a treatment, while investigator-led trials use the locally assessed PFS, which may provide the most relevant indicator of disease progression for clinical decision-making (for instance to change therapy). Neither of these two assessments of PFS is better than the other; they serve different purposes and have their own advantages and limitations. Centrally reviewed PFS is arguably a “cleaner” endpoint, but it is quite expensive to measure and does not reflect clinical routine; as such it is neither feasible nor desirable in investigator-led trials. In terms of safety, investigator-led trials can collect much simpler data than industry-sponsored trials of drugs for which safety has not yet been demonstrated. Typically, in investigator-led trials, the occurrence of Common Terminology Criteria for Adverse Events grade 3 or 4 toxicities will suffice, plus any unexpected toxicity not known to be associated with the drug being investigated. Finally, medical history and concomitant medications, which may be important to document drug interactions with an experimental treatment, serve no useful purpose in investigator-led trials. All in all, investigator-led trials should collect radically simpler data than industry-sponsored trials.

Similarly, data quality needs to be evaluated in a “fit for purpose” manner: while it may be required to attempt to reach 100% accuracy in all the data collected for a pivotal trial of an experimental treatment, such a high bar is by no

means required for investigator-led trials, as long as no systematic bias is at play to create data differences between the randomized treatment groups (for instance, a higher proportion of missing data in one group than in the other) [25]. Both types of trials may benefit from central statistical monitoring of the data; industry-sponsored trials to target centers that are detected as having potential data quality issues, which may require an on-site audit, and investigator-led trials as the primary method for checking data quality.

### Principles of central statistical monitoring

Central statistical monitoring (CSM) is part of risk-based quality management [29]. As shown in Fig. 3, the process starts with a Risk Assessment and Categorization Tool (RACT) [30]. CSM helps quality management by providing statistical indicators of quality based on data collected in the trial from all sources.

A “Data Quality Assessment” of multicenter trials can be based on the simple statistical idea that data should be broadly comparable across all centers [31]. Note that this idea is premised on the fact that data consistency is an acceptable surrogate for data quality. Note also that other tools of central monitoring can be used in addition, to uncover situations in which data issues occur in most (or sometimes all) centers; these other tools, which include “Key Risk Indicators” and “Quality Tolerance Limits”, are beyond the scope of this article. Taken together, all these tools produce statistical signals that may reveal issues in specific centers. Actions must then be taken to address these issues, such as contacting the center for clarification, or in some cases performing an on-site audit to understand the cause of the data issue (Fig. 3).

Although it is a simple idea to perform a central data quality assessment based on the consistency of data across all centers, the statistical models required to implement the idea are necessarily complex to properly account for the natural variability in the data [32, 33]. Essentially, a central data quality assessment is efficient if:

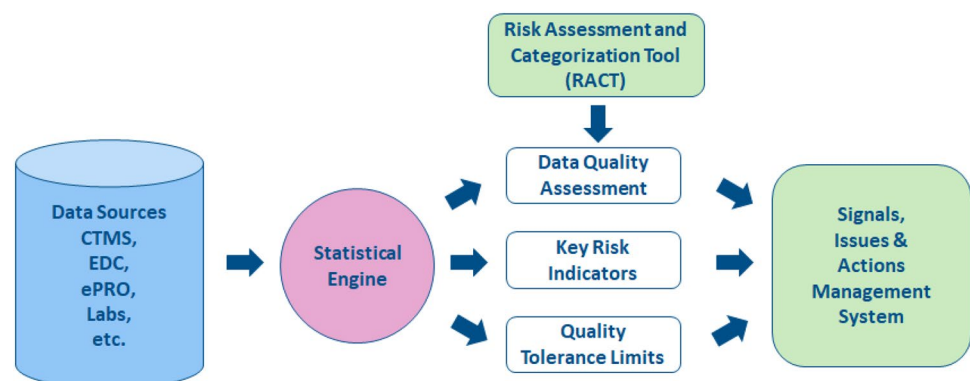
1. data have undergone basic data management checks, whether automated or manual, to eliminate obvious errors (such as out-of-range or impossible values) that can be detected and corrected without a statistical approach;
2. data quality issues are limited to a few centers, while the other centers have data of good quality;
3. all data are used, rather than a few key data items such as those for the primary endpoint or major safety variables;
4. many statistical tests are performed, rather than just a few obvious ones such as a shift in mean or a difference in variability.

It is worth emphasizing the last two points, namely that it is statistically preferable to run many tests on all data collected than on a few data items carefully selected for their relevance or importance. Hence, what matters for a reliable statistical assessment of data quality is volume rather than clinical relevance. The reason is that the power of statistical detection comes from an accumulation of evidence, which would not be available if only important items and standard tests were considered [34]. In addition, investigators pay more attention to key data (such as the primary efficacy endpoint or important safety variables), which, therefore, do not constitute reliable indicators of overall data quality. This being said, careful checks of key data are also essential, but such checks, for the most part, are not statistical in nature.

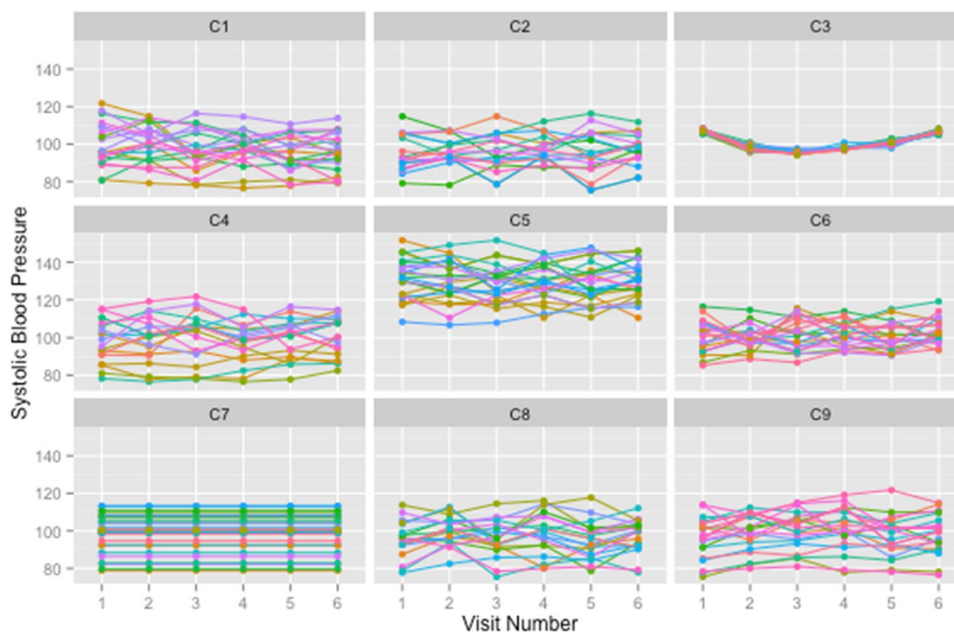
### Central statistical monitoring in practice

Figure 4 shows a made-up example of systolic blood pressure, measured during six successive visits, in nine centers (numbered C1–C9) of a fictitious multicentre trial. Each colored line represents one patient. It is easy, even visually, to spot centers that deviate from the norm: a lack of variability is apparent in center C3, an upward shift in mean in center C5, and data propagation in center C7. While these inconsistencies are too extreme to be commonly seen in practice, others may escape visual scrutiny and yet be revealing of issues worth investigating further. For instance, the

**Fig. 3** The Risk-Based Quality Management process



**Fig. 4** A made-up example of systolic blood pressure, measured during six successive visits, in 9 centers (numbered C1–C9) of a multicentre trial. Each colored line represents the systolic blood pressure of one patient over time



data of center C6 may well be inconsistent with the data of other centers, as it seems to have smaller variability, but it is impossible to tell from Fig. 4 if this inconsistency falls beyond the play of chance.

Figure 4 depicts only one variable, but the power of the statistical approach is to perform many tests on all variables. This can lead to a large number of tests: in a trial of 100 centers, if data are available on 400 variables, and if five tests on average are performed on each variable, the system generates  $100 \times 500 \times 5 = 200,000$  tests. There is obviously a need to summarize the statistical information produced by all these tests in an overall inconsistency index. Essentially, if  $P_{ij}$  represents the  $P$  value of the  $j^{\text{th}}$  statistical test in center  $i$ , the data inconsistency score for center  $i$  is equal to  $DIS_i = \exp\left(\frac{1}{\sum_{j=1}^N w_j} \sum_{j=1}^N w_j \log P_{ij}\right)$ , where  $w_j$  is a weight that accounts for the correlation between the tests. Put simply, the  $DIS$  is a weighted geometric mean of the  $P$  values of all tests performed to compare center  $i$  with all other centers. In fact, the calculation of the  $DIS$  is more complex than this formula suggests, but the technical details are unimportant here [35]. Venet et al. discusses other ways of combining many statistical tests to identify data issues in multicenter trials [31]. It is visually useful to display the  $DIS$  as a function of center size, as shown in Fig. 5 [36]. When the trial includes many centers, it may be useful to limit the number of centers found to have statistical inconsistencies by setting the false discovery rate to a low probability, such as 5% [37].

Timmermans et al. provide a detailed example of CSM applied to a completed trial, the Stomach Cancer Adjuvant Multi-Institutional Trial (SAMIT) Group Trial, involving 1495 patients across 232 centers in Japan, which was

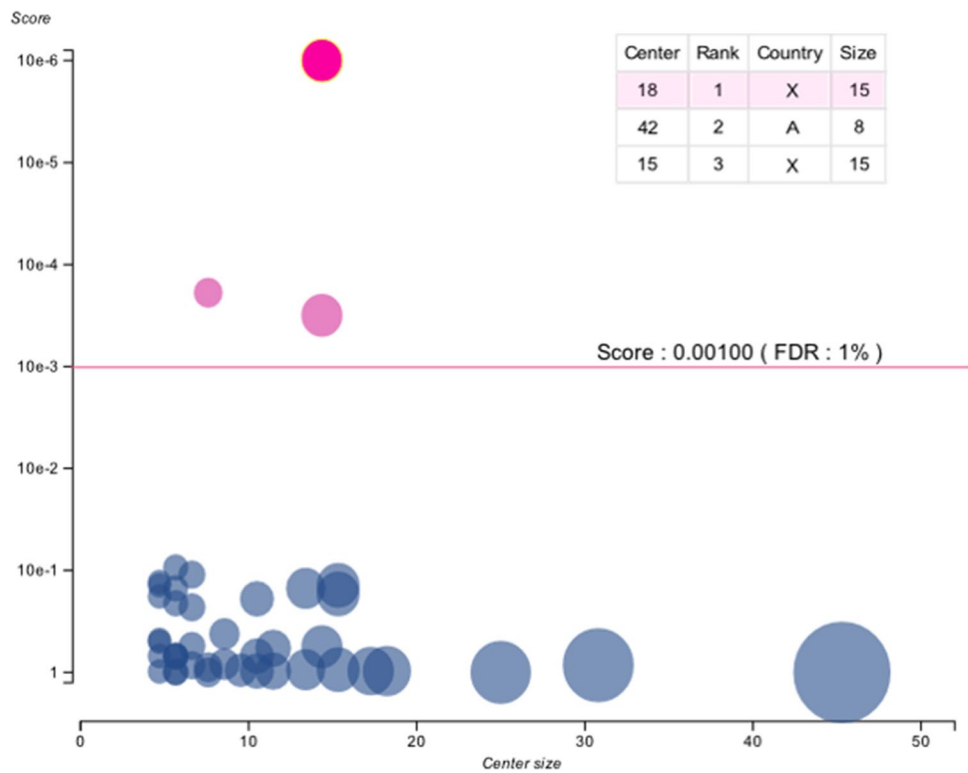
subsequently published [36, 38]. This trial, like many trials in oncology, included many centers with only a couple of patients [36]. Table 2 shows the main findings of CSM in this trial, which led to further checks and data corrections prior to final analysis [38]. This example shows the power of CSM to identify data issues even in small centers, providing a large enough number of patient-related variables are included in the analysis [34]. Table 2 also shows the actions taken, when required, to correct the few data issues that remained in this final dataset. It is noteworthy that some of the statistical findings led to no action if an explanation was found for them (e.g., visits on unusual days of the week), or if, upon further investigation, the findings seemed likely to be due to the play of chance.

### CSM findings

Experience from actual trials [29, 31, 32, 36, 39] as well as extensive simulation studies [35] have shown that a statistical data quality assessment based on the principles outlined above is quite effective at detecting data errors. Experience from actual trials suggests that data errors can be broadly classified as:

1. fraud, such as fabricating patient records or even fabricating entire patients [34, 39, 40]
2. data tampering, such as filling in missing data, or propagating data from one visit to the next [31]
3. sloppiness, such as not reporting some adverse events, making transcription errors, etc. [33]
4. miscalibration or other problems with automated equipment [32]

**Fig. 5** Bubble plot showing the data inconsistency scores of all centers in a multicentre trial as a function of center size (number of patients). Centers shown in magenta have data that are highly inconsistent with data from other centers. *FDR* false discovery rate



**Table 2** Summary of findings in SAMIT trial [36, 38]

Center size	Main finding	Action taken
N=2	Unusually low variability in drug dose	Likely due to chance – no action needed
N=2	Too many missing dates of relapse	Dates of relapse retrieved and added
N=2	Unusual similarities in laboratory values	Likely due to chance – no action needed
N=6	Too many visits on a Saturday	Center open on Saturday – no action needed
N=17	Atypical values for blood tests	Set incorrect zero values to missing
N=24	Too many visits on a Saturday	Center open on Saturday—no action needed

Whilst some of these data errors are worse than others, in so far as they may have a more profound impact on the results of the trial, all of them can potentially be detected using CSM, at a far lower cost and with much higher efficiency than through labor-intensive methods such as source data verification and other on-site data reviews. Investigator-led trials generate more than half of all randomized evidence on new treatments, and it seems essential that this evidence be submitted to statistical quality checks before going to print and influencing clinical practice.

**Acknowledgment** This paper is based on an invited presentation by the first author at the 57th Annual Meeting of the Japan Society of Clinical Oncology (JSCO2019), Fukuoka, October 24–26, 2019. The manuscript was finalized in April 2020, after the world had been hit by the COVID-19 pandemic. The pandemic resulted in a major disruption of on-site monitoring activities of clinical trials, and a corresponding shift towards centralized monitoring [41]. It is likely that this shift will become permanent and will, albeit in the most unfortunate

of circumstances, provide empirical evidence of the usefulness of the approach discussed in the present paper.

**Compliance with ethical standards**

**Conflicts of interest** MB is stockholder of CluePoints and IDDI, LT is employed by CluePoints, and EDS is employed by IDDI.

**References**

- Collins R, Bowman L, Landray M, et al (2020) The magic of randomization versus the myth of real-world evidence. *N Engl J Med* 382:674–678
- Schwartz D, Lellouch J (1967) Explanatory and pragmatic attitudes in therapeutical trials. *J Chron Diseases* 20:637–648
- Sherman RE, Anderson SA, Dal Pan GJ et al (2016) Real-world evidence - what is it and what can it tell us? *N Engl J Med* 375:2293–2297

4. Negrouk A, Lacombe D, Cardoso F et al (2017) Safeguarding the future of independent, academic clinical cancer research in Europe for the benefit of patients. *ESMO Open* 2:e000187
5. Naci H, Davis C, Savovic J et al (2019) Design characteristics, risk of bias, and reporting of randomised controlled trials supporting approvals of cancer drugs by European Medicines Agency, 2014–16: cross sectional analysis. *BMJ* 366:5221
6. Pease AM, Krumholz HM, Downing NS, et al (2017) Postapproval studies of drugs initially approved by the FDA on the basis of limited evidence: systematic review. *BMJ* 357:j1680
7. Cipriani A, Ioannidis JPA, Rothwell PM et al (2020) Generating comparative evidence on new drugs and devices after approval. *Lancet* 395:998–1010
8. Bertolini F, Sukhatme VP, Bouche G (2015) Drug repurposing in oncology—patient and health systems opportunities. *Nat Rev Clin Oncol* 12:732–742
9. Yusuf S, Collins R, Peto R (1984) Why do we need some large, simple randomized trials? *Stat Med* 3:409–422
10. More Trials - to do more trials better. Available at <https://moretrials.net/> (Accessed 14 Apr 2020).
11. Landray MJ, Bax JJ, Alliot L et al (2017) Improving public health by improving clinical trial guidelines and their application. *Eur Heart J* 38:1632–1637
12. Taruno H, Oba M, Matsui K, et al (2019) Analysis of the status of specified clinical trials using JRCT (Japan Registry of Clinical Trials). *Jpn Pharmacol Ther* 47:s148–s158
13. Erdmann J (2005) Researchers facing increasing costs for clinical research, with few solutions. *J Natl Cancer Inst* 97:1492–1494
14. Moore TJ, Zhang H, Anderson G, et al (2018) Estimated costs of pivotal trials for novel therapeutic agents approved by the US Food and Drug Administration, 2015–2016. *JAMA Intern Med* 178:1451–1457
15. Califf RM, Karnash SL, Woodlief LH (1997) Developing systems for cost-effective auditing of clinical trials. *Control Clin Trials* 18:651–660
16. Institute of Medicine (US) (2010) Forum on drug discovery development and translation. Transforming Clinical Research in the United States. National Academies Press, Washington DC
17. Reith C, Landray M, Devereaux PJ et al (2013) Randomized clinical trials—removing unnecessary obstacles. *N Engl J Med* 369:1061–1065
18. Tudur Smith C, Stocken DD, Dunn J et al (2012) The value of source data verification in a cancer clinical trial. *PLoS ONE* 7:e51623
19. Brosteanu O, Schwarz G, Houben P et al (2017) Risk-adapted monitoring is not inferior to extensive on-site monitoring: results of the ADAMON cluster-randomised study. *Clin Trials* 14:584–596
20. Stenning SP, Cragg WJ, Joffe N et al (2018) Triggered or routine site monitoring visits for randomised controlled trials: results of TEMPER, a prospective, matched-pair study. *Clin Trials* 15:600–609
21. Wyman Engen N, Huppler Hullsiek K, Belloso WH et al (2020) A randomized evaluation of on-site monitoring nested in a multinational randomized trial. *Clin Trials* 17:3–14
22. Journot V, Pignon JP, Gaultier C et al (2011) Validation of a risk-assessment scale and a risk-adapted monitoring plan for academic clinical research studies—the Pre-Optimon study. *Contemp Clin Trials* 32:16–24
23. Pronker E, Geerts BF, Cohen A, et al (2011) Improving the quality of drug research or simply increasing its cost? An evidence-based study of the cost for data monitoring in clinical trials. *Br J Clin Pharmacol* 71:467–470
24. Sheetz N, Wilson B, Benedict J et al (2014) Evaluating source data verification as a quality control measure in clinical trials. *Ther Innov Regul Sci* 48:671–680
25. Buyse M, Squifflet P, Coart E, et al (2017) The impact of data errors on the outcome of randomized clinical trials. *Clin Trials* 14:499–506
26. Baigent C, Harrell FE, Buyse M, et al (2008) Ensuring trial validity by data quality assurance and diversification of monitoring methods. *Clin Trials* 5:49–55
27. European Medicines Agency (2011) Reflection paper on risk based quality management in clinical trials. *Eur Med* 1:94–103
28. U.S. Department of Health and Human Services (2013) Food and drug administration guidance for industry. Oversight of Clinical Investigations. A Risk-Based Approach to Monitoring. National Academies Press, DC
29. Timmermans C, Venet D, Burzykowski T (2016) Data-driven risk identification in phase III clinical trials using central statistical monitoring. *Int J Clin Oncol* 21:38–45
30. Transcelerate. RBM Assets. Available at <https://transceleratebiopharmainc.com/assets/rbm-assets/> (Accessed 14 Apr 2020).
31. Venet D, Doffagne E, Burzykowski T et al (2012) A statistical approach to central monitoring of data quality in clinical trials. *Clin Trials* 9:705–713
32. Desmet L, Venet D, Doffagne E et al (2014) Linear mixed-effects models for central statistical monitoring of multicenter clinical trials. *Stat Med* 33:5265–5279
33. Desmet L, Venet D, Doffagne E et al (2017) Use of the beta-binomial model for central statistical monitoring of multicenter clinical trials. *Stat Biopharma Res* 9:1–11
34. Sakamoto J (2016) A hercule poiro of clinical research. *Gastric Cancer* 19:21–23
35. Trotta L, Kabeya Y, Buyse M et al (2019) Detection of atypical data in multicenter clinical trials using unsupervised statistical monitoring. *Clin Trials* 16:512–522
36. Timmermans C, Doffagne E, Venet D et al (2016) Statistical monitoring of data quality and consistency in the Stomach Cancer Adjuvant Multi-institutional Trial Group Trial. *Gastric Cancer* 19:24–30
37. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188
38. Tsuburaya A, Yoshida K, Kobayashi M et al (2014) Sequential paclitaxel followed by tegafur and uracil (UFT) or S-1 vs. UFT or S-1 monotherapy as adjuvant chemotherapy for T4a/b gastric cancer (SAMIT): a phase 3 factorial randomised controlled trial. *Lancet Oncol* 15:886–893
39. George S, Buyse M, Piantadosi S (2020) Fraud in clinical trials. In: Piantadosi S, Meinert C (eds) Principles and Practice of Clinical Trials. Springer, New York
40. Buyse M, George SL, Evans S et al (1999) The role of biostatistics in the prevention, detection and treatment of fraud in clinical trials. *Stat Med* 18:3435–3451
41. U.S. Department of Health and Human Services (2020) Food and drug administration. Guidance for industry, investigators, and institutional review boards. FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic. Available at <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/fda-guidance-conduct-clinical-trials-medical-products-during-covid-19-pandemic> (Accessed 15 Apr 2020).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.