University of Basel

Department of Clinical Research

Universitätsspital Basel

Development and assessment of evidence-based strategies towards increased feasibility and transparency of investigator-initiated clinical trials in Switzerland

Inaugural dissertation

to be awarded the degree of **Dr. sc. med**.

presented at the

Faculty of Medicine

of the University of Basel

by

Katharina Klatte

from Tokyo, Japan

Basel, 2022

Approved by the Faculty of Medicine

On application of

Prof. Dr. med. et phil. Matthias Briel (first supervisor)

Prof. Dr. med. Mirjam Christ-Crain (second supervisor)

Prof. Dr. Kerry Hood (external expert)

Prof. Dr. med. Christiane Pauli-Magnus (further advisor)

Basel, ……

……

Prof. Dr. Primo Schär, Dean

TABLE OF CONTENTS

ACKNOWLEDGMENTS

Firstly, I would like to thank my first PhD advisor, Prof Dr. Matthias Briel, for the continuous support, patience, motivation and expertise provided to me throughout my doctoral work. Thank you for always providing constructive feedback and guiding me through this time of education.

Secondly, I would like to extend my gratitude to Prof. Christiane Pauli-Magnus. Thank you for giving me the opportunity to work at the Department of Clinical Research and complete my PhD in this inspiring environment. I have always received support and guidance when needed.

I would also like to thank my second PhD advisor Prof. Mirjam Christ-Crain, for her positive attitude and encouragement throughout my thesis work. I appreciate her regular participation in my PhD annual meetings and her support throughout my PhD journey.

Besides my advisor, I would like to thank Prof. Kerry Hood, as external expert, for her positive attitude and willingness to be part of this thesis work.

My sincere thanks also go to my colleagues at the Clinical Trial Unit (CTU) and the Basel Institute for Clinical Epidemiology and Biostatistics (CEB). Benjamin Speich, Dmitry Gryaznov, and Stefan Schandelmaier, thank you for including me into your projects and teaching me valuable skills while finding my path into clinical research. I want to thank my colleagues Viktoria Gloy, Alexandra Schulz, Pascal Benkert, Andrea Wiencierz and Gilles Dutilh for supporting me with their positive attitude and experience.
I especially want to acknowledge my friend Suvitha Subramaniam, who has been by my side from the beginning of this journey, encouraging me, supporting me and motivating me to complete this work.

I want to thank my dear friend Janine Wasmuth for her emotional support throughout writing this thesis.

I would like to thank my mother for spending most of her vacations taking care of my children to allow for some extra working hours and relaxing time. And I would like to

thank my father for supporting me throughout writing this thesis and my academic life in general.

A special thanks to my four wonderful children Tom, Mio, Laura and Luise for coping with an always working mother and to the three big children for taking good care of their youngest sibling.

Last but not least, I would like to thank my husband, Max, for his incredible amount of support. Taking on a big part of our daily house and organizational work while working full-time. Thank you for living this crazy life with me.

PLAIN SUMMARY

This work addresses the obligation to minimize research waste by identifying barriers and needs for support in important processes of clinical research and by proposing efficient strategies to improve the quality of research practice. Major sources of waste in clinical research have been identified by the "Increasing Value, Reducing Waste" series in *The Lancet* in 2014. Two considerations in this series address the problem of inefficient trial management and insufficient research transparency. Collected evidence suggests that inefficient management and monitoring of the procedural conduct of trials are a major source of waste even in well-designed studies addressing important questions. The absence of a continuous oversight of established trial processes endanger completion of trials in a set timeframe or even cause premature discontinuation. Increasing feasibility of clinical trials by providing an evidence-based strategy to effectively support the conduct of clinical trials at the University Hospital of Basel that has the potential to be transferred to the whole academic network for clinical research in Switzerland was aspired in this thesis. Along with feasibility, it is important that information of a trial including results is publicly available. In Switzerland, prospective registration of a clinical trial in a primary trial registry has been made mandatory by law in 2014 (Art 56 Human Research Act). We analyzed research transparency in terms of trial registration and results publication in a local setting in Switzerland to assess the successful implementation and enforcement of national efforts and identify potential barriers.

In a first step, we systematically reviewed existing evidence on effective monitoring strategies both in the medical literature and across international clinical research stakeholder groups. Monitoring strategies varied in their methodological approach but the effectiveness of risk-based and triggered approaches could be shown with moderate certainty. However, we did not find evidence on the effect of these methods on the overall trial conduct. Based on these findings, we then engaged local, national and international stakeholder representatives in the creation of a comprehensive risk-tailored approach integrating monitoring in the broader context of trial management. We systematically reviewed information on risk indicators commonly used to guide monitoring in the academic setting and in industry and identified risk elements extended to the overall management of a clinical trial. In order to continuously visualize the status of identified risk elements throughout the study conduct, we initiated the user-centered development of a supporting study dashboard. The final risk-tailored approach consisted of the following components: A study-specific risk assessment prior to study start, selection and development of data

based pathways addressing the identified risks, and the continuous visualization of the status of risk elements in a study dashboard. The generic content of the dashboard provides continuous information and support for risk indicators applicable to almost all clinical trials (Data quality, Recruitment, Retention, and Safety management) and the optional content is based on further study-specific items identified during the risk assessment (e.g. Follow-up visits, Re-consent process, Sampling management, Imaging quality). User-testing of the risk assessment and study dashboards revealed that the continuous oversight of most critical elements efficiently supports the trial-related work procedures of principle investigators, trial managers and trial monitors.

In a second project of this thesis, we assessed current trial registration and publication for clinical intervention studies approved by the Ethics Committee North and Central Switzerland (EKNZ) in the last five years. Registration of all clinical trials would provide an overview of what research is being conducted at present and registries constitute an ideal platform for the publication and dissemination of research results. Identifying factors influencing registration and potential barriers provides a basis for further initiatives to increase trial registration. Prospective trial registration has increased over the last five years and trials with higher risk category, multicenter trials and trials taking advantage of Clinical Trials Unit services were associated with higher registration rates. Although prospective trial registration prevalence has improved within the last five years within the EKNZ approved studies, a strong need for support in the registration process was identified in our qualitative evaluation.

The impact of this work and whether it eventually increases feasibility and transparency in clinical research critically depends on its implementation, evaluation, and refinement. Sharing current knowledge on effective monitoring strategies with trialists and monitors to choose evidence-based strategies for their trials constitutes a major support for investigator-initiated trials in the academic environment. The advancement of a risk-based trial monitoring approach into a comprehensive risk-tailored approach supporting the overall conduct of a trial and considering trial monitoring as an integrative part of trial management has the potential to efficiently optimize study processes. While an uptake of the study specific risk assessment and the use of a study dashboard as a standard process would be aspired for all RCTs in the future, improving the timeline and resources needed for the development of a study specific dashboard will be important to advance the generation of affordable and efficient dashboards for investigator-initiated trials. Sharing evidence on the registration behavior and perceived barriers by researchers in the local setting of the EKNZ helps to understand underlying processes and test measures for improvement. Supporting researchers in the process of trial registration and

educating research institutes and investigators about the need and advantages of trial registration, has the potential to facilitate the implementation of automated processes and Standard Operating Procedures (SOPs) ensuring the registration of all clinical trials. Establishing trial registries as a primary platform for sharing research results should be aspired in the future.

INTRODUCTION

1.1 The need for increased feasibility and reduced waste in investigator-initiated clinical trials

Randomized clinical trials (RCTs) are the gold standard approach to investigate the benefits and harms of a health care intervention. Randomization addresses bias resulting from confounding and allows us to draw causal conclusions on the measured intervention effect by balancing the prognosis of patients among the different treatment arms at the outset of a trial. However, RCTs are often complex as well as resource and cost intensive to plan and perform, which is a major barrier in performing RCTs.[1-3]

One out of four RCTs is prematurely discontinued.[4 5] The main reason identified for RCT discontinuation was poor recruitment.[5] Most RCTs in the academic setting have limited resources for the support of the trial conduct including clinical trial management and monitoring. Premature discontinuation of a significant proportion of government-funded clinical trials and recruitment delays inflating costs have recently been identified as a substantial problem of clinical research in Switzerland.[6] The problem of discontinued clinical trials became also apparent during the Covid-19 pandemic as one third of trials started were discontinued 6 months later and many delayed. [7-9] If studies are discontinued due to poor recruitment or administrative/ management issues, trial participants are exposed to potential health risks and other burdens without the research question being answered. In addition, this constitutes a major source of research waste.

The “Increasing Value, Reducing Waste” series in *The Lancet* [10-15] in 2014 provided convincing evidence for various sources of waste in biomedical research, including clinical trials. Two important aspects in this discussion were the need for more efficient trial management and the need for increased transparency of clinical research. In order to improve the conduct of clinical research, the key authors of the series initiated the Reduce Research Waste and Reward Diligence (REWARD) campaign (https://www.thelancet.com/ campaigns/efficiency/statement). In terms of feasibility and trial conduct, this campaign underlines the need for an increase in efficiency of participant recruitment, participant retention, data monitoring, and data

sharing and formulates the need for a larger evidence base on efficiency improving strategies.

In this PhD thesis, we aimed to address the problem of inefficient trial management considering the whole process of clinical trial conduct including trial monitoring and address insufficient transparency focusing on trial registration and results publication.

We first evaluated the available evidence on the effectiveness and efficiency of monitoring strategies and their impact on trial conduct in terms of recruitment and retention in a Cochrane Methods Review (first thesis manuscript). [16]

Trial monitoring is important to maintain oversight of the trial conduct, ensure conform data collection and documentation processes, and to consider patient rights and safety. However, focusing on the most critical items of trial conduct, ensuring patient safety and results validity, in monitoring is key. Extensive on-site monitoring with source data verification creates financial and logistical barriers to the conduct of clinical trials, in particular investigator-sponsored trials, without evidence of benefit for trial participants or increased trial validity.[17-21] Efficient and continuous monitoring of a clinical trial is important to initiate support in case of problems related to the conduct, documentation or safety of a clinical trial. However, as trial monitoring is only one element of the responsibility to maintain oversight of the study progress, accuracy of procedures, documentation and data collection, it is sensible to consider monitoring as an integrative part of efficient trial management. Timely participant recruitment, comprehensive participant follow-up, and high quality of collected data are important goals for trial management and monitoring. Given the limited study budget in investigator-initiated trials, prioritization of resources to the most critical elements of a trial is essential. In comparison to investigator-sponsored trials, industry-sponsored RCTs tend to perform better in avoiding discontinuation due to poor recruitment.[5] Along with a profound feasibility assessment, continuous monitoring of recruitment throughout the study would allow for early corrective measures in one or more participating centers. In addition, adjustments to the sample size estimations within the planning phase have to be considered throughout the study based on continuous data evaluation. Along with recruitment and retention problems, organizational problems in other areas of a clinical trial management, including training of staff, follow-up visits, Investigational Medicinal Product (IMP) management, have been identified [22] [18] [23] For example, organizing and integrating follow-up visits into clinical routine within a predetermined timeframe constitutes a complex management procedure critical for the standardized collection of outcome

data. Thus, a comprehensive approach supporting trial management and providing a continuous overview on critical study elements is needed. Early awareness and interference in case of issues endangering the progress of trial is important as discontinuation has a large impact in terms of research waste and ethical obligations towards patients. In the second thesis manuscript we report the development and functionality of a risk-tailored approach towards efficient management and monitoring of investigator-initiated trials".

Another important source of waste in the field of RCTs is an ongoing intransparency in terms of non-registration of trials and non-publication of results. The International Clinical Trial Registration Platform (ICTRP) of the World Health Organization (WHO) primary registries constitutes an environment for sharing information on planned, ongoing, and completed clinical trials. The mission of this platform is to ensure that a complete overview of trials is available to all stakeholders in health care.[24] Measures to improve trial registration practice are needed, especially in the academic setting. Prospectively registration of trials is more prevalent in industry-sponsored trials (91.6 %) than investigator-sponsored trials (74.8 %).[5] Similarly, non-publication of trial results in trial registries or peer-reviewed journals was the case for only 3.9 % of industry-sponsored trials compared to 23.8 % of investigator-sponsored trials. Since trial discontinuation is often associated with non-publication, useful data gets lost and can constitute substantial research waste.[12] Even if data from discontinued trials commonly do not allow answering the primary research question, the data can still be used in meta-analyses. In addition, lessons learnt should be shared in order to prevent future RCTs from repeating potential mistakes in planning and conduct. [4] Making all planned, on-going and completed research publicly available will improve research transparency and will ultimately strengthen the validity and value of the scientific evidence base and ensure public trust.

1.2 The need for evidence on effective monitoring strategies

Trial monitoring should be effective in terms of protecting patient rights and safety and ensuring valid data generation, analysis and interpretation as requested by the International Conference of Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use – Good Clinical Practice Guideline (ICH GCP).[25] However, no standard implementation of an evidence-based monitoring approach has been developed so far and inefficient trial monitoring has been identified as a major cost driver for RCTs. [11] [23] [26] A broad consensus among

stakeholders for the adoption of risk proportionate approaches to clinical trial monitoring exists [27-31]; and various new risk-adapted monitoring approaches have been developed [32] and are currently in clinical research practice. However, evidence for a successful incorporation of new risk-adapted and central monitoring approaches need to be gathered through empirical research.

Risk-adapted monitoring can be based on an initial assessment of the risk associated with an individual trial protocol and may incorporate further trial or site-specific classifications into different risk categories. [33 34] The proposed monitoring plan for a health care trial is then dependent on the respective risk classification of the assessment and is complemented by central data monitoring. Central or centralized monitoring is a concept that is based on a remote evaluation of electronically available study data. It involves processing of data tables, data entry information and statistical tests performed on variables collected in the database with the aim to identify centers with poor data quality or with other problems in trial conduct. [35]. Monitoring strategies incorporating central monitoring as a tool to trigger early corrective actions have been introduced and their effectiveness and efficiency needs to be further evaluated.[36] [37] The first Manuscript of this thesis entitled "Monitoring in clinical intervention trials" describes our systematic review of the available evidence on the effectiveness of monitoring strategies.

1.3 The need for a risk-tailored approach to support trial management and monitoring

Almost half (48.7 %) of clinical trials approved between 2016 and 2020 in Switzerland were multicenter trials.[38] Multicenter trials traditionally require a large amount of resources for monitoring and a complex trial management infrastructure. The involvement of many different sites constitutes a challenge for the oversight on the study conduct for monitors, trial managers, and principal investigators. In addition, resources are limited, especially in the academic setting. [3 6] Hence, new efficient approaches to optimize the process of oversight of investigator-initiated trials, encompassing monitoring and management aspects, are urgently needed.

Trial sites may have different challenges when integrating a trial in their daily clinical routine and therefore need support in different aspects of the study conduct.

A risk assessment extended to the whole management process of a trial including elements of trial monitoring addresses the need for prioritizing study specific elements critical for the conduct of a study at hand. The integration of central data application, like for example the continuous overview of follow-up visits, into the trial

management process beyond the surveillance of monitoring elements enables the continuous evaluation of trial progress. An up-to-date overview for the trial manager or principal investigator of the overall study progress in terms of recruitment, retention, completeness and timeliness of data entry as well as on the completeness of follow-up visits is crucial for maintaining oversight and leading a clinical trial. A real time evaluation and visualization of critical study risks enables early interference in case of need for support and the timely initiation of preventive measures.

Based on findings from our systematic review on monitoring strategies and a contextual analysis reflecting stakeholder perceptions, we addressed the need for a comprehensive risk-tailored approach for optimizing the management of trial conduct integrating trial monitoring in a broader context of risk management. The development and user testing of this risk-tailored approach and study dashboard are described in the Manuscript entitled “Development of a Risk-tailored Approach and Dashboard for Efficient Management and Monitoring of Investigator-Initiated Trials”.

1.4 The need for increased transparency in clinical trials

Transparency of clinical research is important to ensure that decisions in health care are based on all relevant research findings in the health field of interest. Hence, the publication of all results available including the results of discontinued studies is essential. Registration of all clinical studies would provide an overview of what research is being conducted at present and would emphasize the ethical obligations of researchers, sponsors and publishers with regards to the publication and dissemination of their research results. If evidence-based decisions in healthcare are informed by results available in a complete trial registry, more attention would be given to studies not publishing their results in an appropriate timeframe and could promote the obligation to publish results in the health field community. Hence, publicly accessible information on trials including research methods and protocols, results and interpretation (reports and publications) is essential to ensure a contribution of all conducted clinical trials to the progress of clinical knowledge.

For many years, the scientific community and others have worried about reporting biases such that negative results from clinical trials may be less likely to be published than positive results.[39] One of the proposals to address this potential bias and to enhance transparency was a comprehensive clinical trials register that would inform the scientific community and the public which trials had been started [40]. The today’s largest public trial registry, ClinicalTrials.gov, was originated as a result of patient

activist lobbying in 1997, claiming registration of prospective clinical trials in a public database from inception to reporting of final results.[41] Meanwhile, several clinical trials registers have been implemented (ISRCTN, EudraCT, and ANZCTR[42-44]). The proportion of registered trials increased substantially after 2004 when the International Committee of Medical Journal Editors (ICMJE) recommended publishing trial reports only if the trial was registered.[45] More recently, the World Medical Association included a statement in the Declaration of Helsinki that "every research study involving human subjects must be registered" which will likely further increase the proportion of registered trials.[46] In Switzerland, prospective registration of a clinical trial in a primary trial registry has been made mandatory by law in 2014 (Art 56 Human Research Act).[47]

In view of these positive developments, trial registries and specifically ClinicalTrials.gov would theoretically represent an ideal source of information for researchers, patient, care givers and policy makers about research questions, trial designs and endpoints, trials status and trial results. However, there is still no incentive to ensure completeness and accuracy provided in trial registries. Various studies have examined trial registration and, in particular, prospective trial registration based on published RCTs [48-52], with prospective registration rates of RCTs ranging from 61 % in 2007 to 83% in 2012.[53] While this suggests a continuing international improvement in prospective registration in the decade following the statement of the ICMJE in 2004, [45]compliance with result reporting remains still low. Of all trials reported to be terminated or completed between January 2008 and August 2012, only 13% reported summary results within 12 months after trial completion, whereas 38% reported results at any time up to September 2013.[52] Furthermore, an analysis of registered trials from the "Trials tracker" initiative revealed in 2021 that only 75.3 % of trials from all sponsors have been published two years after trial completion. A similar analysis comprising trials in the EU on the basis of the EudraCT registry revealed that 81.1 % accessed on May 5th 2022 were registered. However, the sensitivity of such continuous automated search processes for study results has not been examined yet. [54-57]

A recent meta-research study in Switzerland, Germany, UK and Canada revealed that non-publication of RCTs has declined, but remains common. The study outlined that 21% of unpublished trials could not be identified in registries impeding any uptake of knowledge and only 16% of investigator-sponsored trials had results reported in a trial registry.[5] In view of these findings, further efforts are needed to increase compliance with registration and publication requirements and quality of recorded items as a basis for clinical research transparency.

In this PhD work, we started such efforts to increase transparency of clinical trials on a local level by collaborating with the Ethics Committee of Northwestern and Central Switzerland (EKNZ) in a mixed methods study on the non-registration and non-publication of approved clinical trials. In the third thesis manuscript titled “Towards full clinical trial registration and results publication: longitudinal meta-research study in Northwestern and Central Switzerland“, we report a mixed-methods study on the non-registration and non-publication of clinical trials in the local context of Northwestern Switzerland and propose measures based on our findings to ensure full registration and results publication.

1.5 Main Objectives of this PhD

1. To systematically summarize the existing evidence on the effectiveness of risk-based monitoring strategies for intervention trials in health care.

2. To design and user test an effective risk-tailored trial management approach integrating elements of trial monitoring based on empirical evidence and stakeholder input.

3. To evaluate the prevalence of prospective registration and results publication in clinical trials approved by the Ethics Committee Northwestern and Central Switzerland (EKNZ) between 2016 and end of 2020 and to investigate factors associated with trial registration and reasons for non-registration.

1.6 Contributions by the PhD student

I had the great opportunity to be a PhD student at the Clinical Trial Unit (CTU) of the Department of Clinical Research (DKF) at the University Hospital in Basel. Working in such a multidisciplinary environment enabled me to acquire knowledge on different processes involved in clinical research. Since two of my projects required close collaboration with several different teams within the DKF and within the local research community including principal investigators, ethics committees and trial managers, organizing these projects combined the challenging task of integrating different needs and ideas and the opportunity to learn from experts of various fields of clinical research. Throughout my PhD, I was able to perform many of the activities independently, while learning the fundamentals of evidence synthesis, comparative-effectiveness research, project management, data analysis and clinical trial setup.

The first step of my PhD work was to gather current evidence on the effectiveness of trial monitoring approaches in a systematic literature review. This work required the development of a search strategy, which I conducted with the guidance of two experienced information specialists. I independently performed the deduplication of identified literature and organized and performed the abstract, full text screening and data extraction in duplicate. After conceptualization with my supervisors I developed a data extraction form and programmed the content into the Cochrane Epi 4 review tool.[58] Together with my supervisors I developed the review protocol, including endpoint definitions, quantitative or qualitative analysis plans for main and secondary outcomes, which was published in the Cochrane library.[16] I analyzed all data available from included studies and performed a meta-analysis where possible.

The next step in my PhD journey was to coordinate a project group in the development of a new monitoring strategy based on an extensive literature search, a contextual analysis and the involvement of local, national and international stakeholders. This process required the organization of stakeholder meetings, interviews, and discussions with various team members, literature reviews, and summaries of current practices, needs and decisions. This iterative process took several months and resulted in the proposal of a new concept including a study-specific risk assessment guide and an accompanying manual. After the development of the new concept, we set up the structure for a study dashboard based on the risk-tailored monitoring approach for two ongoing RCTs at the University Hospital of Basel. My colleague Suvitha Subramaniam from the data science team introduced me to the concept and structure of R shiny programming. After the iterative process of dashboard refinement, which included many meetings with the study teams, we started the user-testing phase and I conducted semi-structured interviews with trial monitors, trial coordinators/trial managers and principal investigators.

Within the project on trial registration, the close collaboration with the EKNZ enabled me to work with a large data set of approved trials. I set up a data extraction form using internal data software of the University Hospital Basel. I coordinated the search and data extraction for all trials in duplicate. In the analysis phase, I analyzed both qualitative and descriptive quantitative data.

I critically interpreted all data together with my supervisors and co-authors and developed first drafts for all manuscripts, coordinated the critical revision by co-authors, submitted and revised manuscripts as first and co-author, and presented and discussed our work at local and national meetings.

In addition to my PhD work, I also had the opportunity to contribute to a workshop at the German Evidence based Medicine Network in 2020, and the CAS Study Nurse Course of the DKF. I was fortunate to be exposed to a collaborative environment and contributed to several projects not directly related to my PhD. Joining different projects provided further insights into other topics such as subgroup analysis in oncological trials, reporting quality of trial protocols, or the reliability of information across trial registries. Several of these projects have led to publications that I co-authored (section 3).

FIRST AUTHOR PUBLICATIONS

2.1 Manuscript I: Monitoring strategies for clinical intervention studies

Katharina Klatte[1], Christiane Pauli-Magnus[1], Sharon B Love[2], Matthew R Sydes[2], Pascal Benkert[1], Nicole Bruni[1], Hannah Ewald[3], Patricia Arnaiz Jimenez[1], Marie Mi Bonde[1], Matthias Briel[1]

[1]Department of Clinical Research, University Hospital Basel and University of Basel, Basel, Switzerland
[2]MRC Clinical Trials Unit at UCL, University College London, London, UK
[3]University Medical Library, University of Basel, Basel, Switzerland

Status: Published Cochrane Library. 08 December 2021

Abstract

Background

Trial monitoring is an important component of good clinical practice to ensure the safety and rights of study participants, confidentiality of personal information, and quality of data. However, the effectiveness of various existing monitoring approaches is unclear. Information to guide the choice of monitoring methods in clinical intervention studies may help trialists, support units, and monitors to effectively adjust their approaches to current knowledge and evidence.

Objectives

To evaluate the advantages and disadvantages of different monitoring strategies (including risk-based strategies and others) for clinical intervention studies examined in prospective comparative studies of monitoring interventions.

Search methods

We systematically searched CENTRAL, PubMed, and Embase via Elsevier for relevant published literature up to March 2021. We searched the online 'Studies within A Trial' (SWAT) repository, grey literature, and trial registries for ongoing or unpublished studies.

Selection criteria

We included randomized or non-randomized prospective, empirical evaluation studies of different monitoring strategies in one or more clinical intervention studies. We applied no restrictions for language or date of publication.

Data collection and analysis

We extracted data on the evaluated monitoring methods, countries involved, study population, study setting, randomization method, and numbers and proportions in each intervention group. Our primary outcome was critical and major monitoring findings in prospective intervention studies. Monitoring findings were classified according to different error domains (e.g. major eligibility violations) and the primary outcome measure was a composite of these domains. Secondary outcomes were individual error domains, participant recruitment and follow-up, and resource use. If we identified more than one study for a comparison and outcome definitions were similar across identified studies, we quantitatively summarized effects in a meta-analysis using a random-effects model. Otherwise, we qualitatively summarized the results of eligible studies stratified by different comparisons of monitoring strategies. We used the GRADE approach to assess the certainty of the evidence for different groups of comparisons.

Main results

We identified eight eligible studies, which we grouped into five comparisons.

1. Risk-based versus extensive on-site monitoring: based on two large studies, we found moderate certainty of evidence for the combined primary outcome of major or critical findings that risk-based monitoring is not inferior to extensive on-site monitoring. Although the risk ratio was close to 'no difference' (1.03 with a 95% confidence interval [CI] of 0.81 to 1.33, below 1.0 in favor of the risk-based strategy), the high imprecision in one study and the small number of eligible studies resulted in a wide CI of the summary estimate. Low certainty of evidence suggested that monitoring strategies with extensive on-site monitoring were associated with considerably higher resource use and costs (up to a factor of 3.4). Data on recruitment or retention of trial participants were not available.

2. Central monitoring with triggered on-site visits versus regular on-site visits: combining the results of two eligible studies yielded low certainty of evidence with a

risk ratio of 1.83 (95% CI 0.51 to 6.55) in favor of triggered monitoring intervention. Data on recruitment, retention, and resource use were not available.

3. Central statistical monitoring and local monitoring performed by site staff with annual on-site visits versus central statistical monitoring and local monitoring only: based on one study, there was moderate certainty of evidence that a small number of major and critical findings were missed with the central monitoring approach without on-site visits: 3.8% of participants in the group without on-site visits and 6.4% in the group with on-site visits had a major or critical monitoring finding (odds ratio 1.7, 95% CI 1.1 to 2.7; P = 0.03). The absolute number of monitoring findings was very low, probably because defined major and critical findings were very study specific and central monitoring was present in both intervention groups. Very low certainty of evidence did not suggest a relevant effect on participant retention, and very low-quality evidence indicated an extra cost for on-site visits of USD 2,035,392. There were no data on recruitment.

4. Traditional 100% source data verification (SDV) versus targeted or remote SDV: the two studies assessing targeted and remote SDV reported findings only related to source documents. Compared to the final database obtained using the full SDV monitoring process, only a small proportion of remaining errors on overall data were identified using the targeted SDV process in the MONITORING study (absolute difference 1.47%, 95% CI 1.41% to 1.53%). Targeted SDV was effective in the verification of source documents, but increased the workload on data management. The other included study was a pilot study, which compared traditional on-site SDV versus remote SDV and found little difference in monitoring findings and the ability to locate data values despite marked differences in remote access in two clinical trial networks. There were no data on recruitment or retention.

5. Systematic on-site initiation visit versus on-site initiation visit upon request: very low certainty of evidence suggested no difference in retention and recruitment between the two approaches. There were no data on critical and major findings or on resource use.

Authors' conclusions

The evidence base is limited in terms of quantity and quality. Ideally, for each of the five identified comparisons, more prospective, comparative-monitoring studies nested in clinical trials and measuring effects on all outcomes specified in this review

are necessary to draw more reliable conclusions. However, the results suggesting risk-based, targeted, and mainly central monitoring as an efficient strategy are promising. The development of reliable triggers for on-site visits is ongoing; different triggers might be used in different settings. More evidence on risk indicators that identify sites with problems or the prognostic value of triggers is needed to further optimize central monitoring strategies. In particular, approaches with an initial assessment of trial-specific risks that need to be closely monitored centrally during trial conduct with triggered on-site visits should be evaluated in future research.

Plain language summary

New monitoring strategies for clinical trials

Our question

We reviewed the evidence on the effects of new monitoring strategies on monitoring findings, participant recruitment, participant follow-up, and resource use in clinical trials. We also summarized the different components of tested strategies and qualitative evidence from process evaluations.

Background

Monitoring a clinical trial is important to ensure the safety of participants and the reliability of results. New methods have been developed for monitoring practices but further assessments of these new methods are needed to see if they do improve effectiveness without being inferior to established methods in terms of patient rights and safety, and quality assurance of trial results. We reviewed studies that examined this question within clinical trials, i.e. studies comparing different monitoring strategies used in clinical trials.

Study characteristics

We included eight studies, which covered a variety of monitoring strategies in a wide range of clinical trials, including national and large international trials. They included primary (general), secondary (specialized), and tertiary (highly specialized) health care. The size of the studies ranged from 32 to 4371 participants at one to 196 sites.

Key results

We identified five comparisons. The first comparison of risk-based monitoring versus extensive on-site monitoring found no evidence that the risk-based approach is inferior to extensive on-site monitoring in terms of the proportion of participants with a critical or major monitoring finding not identified by the corresponding method, while resource use was three- to five-fold higher with extensive on-site monitoring. For the second comparison of central statistical monitoring with triggered on-site visits versus regular (untriggered) on-site visits, we found some evidence that central statistical monitoring can identify sites in need of support by an on-site monitoring intervention. In the third comparison, the evaluation of adding an on-site visit to local and central monitoring revealed a high percentage of participants with major or critical monitoring findings in the on-site visit group, but low numbers of absolute monitoring findings in both groups. This means that without on-site visits, some monitoring findings will be missed, but none of the missed findings had any serious impact on patient safety or the validity of the trial's results. In the fourth comparison, two studies assessed new source data verification processes, which are used to check that data recorded within the trial Case Report Form (CRF) match the primary source data (e.g. medical records), and reported little difference to full source data verification processes for the targeted as well as for the remote approach. In the fifth comparison, one study showed no difference in participant recruitment and participant follow-up between a monitoring approach with systematic initiation visits versus an approach with initiation visits upon request by study sites.

Certainty of evidence

We are moderately certain that risk-based monitoring is not inferior to extensive on-site monitoring with respect to critical and major monitoring findings in clinical trials. For the remaining body of evidence, there is low or very low certainty in results due to imprecision, small number of studies, or high risk of bias. Ideally, for each of the five identified comparisons, more high-quality monitoring studies that measure effects on all outcomes specified in this review are necessary to draw more reliable conclusions.

Introduction

Description of the problem or issue

Trial monitoring is important for the integrity of clinical trials, the validity of their results, and the protection of participant safety and rights. The International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) for Good Clinical Practice (GCP) formulated several requirements for trial monitoring (ICH 1996). However, the effectiveness of various existing monitoring approaches was unclear. Source data verification (SDV) during monitoring visits was estimated to use up to 25% of the sponsor's entire clinical trial budget, even though the association between data quality or participant safety and the extent of monitoring and SDV has not been clearly demonstrated (Funning 2009). Consistent application of intensive on-site monitoring creates financial and logistical barriers to the design and conduct of clinical trials, with no evidence of participant benefit or increase in the quality of clinical research (Baigent 2008; Duley 2008; Embleton-Thirsk 2019; Hearn 2007; Tudur Smith 2012a; Tudur Smith 2014).

Recent developments at international bodies and regulatory agencies such as the European Medicines Agency (EMA), the Organisation for Economic Co-operation and Development (OECD), the European Commission (EC) and the Food and Drug Administration (FDA), as well as the 2016 addendum to ICH E6 GCP have supported the need for risk-proportionate approaches to clinical trial monitoring and overall trial management (EC 2014; EMA 2013; FDA 2013; ICH 2016; OECD 2013). This has encouraged study sponsors to implement risk assessments in their monitoring plans and to use alternative monitoring approaches. There are several publications reporting on the experience of using a risk-based monitoring approach, often including central monitoring, in specific clinical trials (Edwards 2014; Heels-Ansdell 2010; Valdés-Márquez 2011). The main idea is to focus monitoring on trial-specific risks to the integrity of the research and to essential GCP objectives, that is, risks that threaten the safety, rights, and integrity of trial participants; the safety and confidentiality of their data; or the reliable report of the trial results (Brosteanu 2017a). The conduct of 'lower risk' trials (lower risk for study participants) — which optimize the use of already authorized medicinal products, validated devices, implemented interventions, and interventions formally outside of the clinical trials regulations — may particularly benefit from a risk-based approach to clinical trial monitoring in terms of timely completion and cost efficiency. Such 'lower risk' trials

are often investigator-initiated or academic- sponsored clinical trials conducted in the academic setting (OECD 2013). Different risk assessment strategies for clinical trials have been developed, with the objective of defining risk-proportionate monitoring plans (Hurley 2016). There is no standardized approach for examining the baseline risk of a trial. However, risk assessment approaches evaluate risks associated with the safety profile of the investigational medicinal product (IMP), the phase of the clinical trial, and the data collection process. Based on a prior risk assessment, a study-specific combination of central/centralized and on-site monitoring might be effective. Centralized monitoring, also referred to as central monitoring, is defined as any monitoring processes that are not performed at the study site (FDA 2013), and includes remote monitoring processes. Central data monitoring is based on the evaluation of electronically available study data in order to identify study sites with poor data quality or problems in trial conduct (SCTO 2020; Venet 2012), whereas on-site monitoring comprises site inspection, investigator/staff contact, SDV, observation of study procedures, and the review of regulatory elements of a trial. Central statistical monitoring (including plausibility checks of values for different variables, for instance) is an integral part of central data monitoring (SCTO 2020), but this term is sometimes used interchangeably with central data monitoring. The OECD classifies risk assessment strategies into stratified approaches and trial-specific approaches, and proposes a harmonized two-pronged strategy based on internationally validated tools for risk assessment and risk mitigation (OECD 2013). The effectiveness of these new risk-based approaches in terms of quality assurance, patient rights and safety, and reduction of cost, needs to be empirically assessed. We examined the risk-based monitoring approach followed at our own institution (the Clinical Trial Unit and Department of Clinical Research, University Hospital Basel, Switzerland) using mixed methods (von Niederhausern 2017). In addition, several prospective studies evaluating different monitoring strategies have been conducted. These include ADAMON (ADApted MONitoring study; Brosteanu 2017a), OPTIMON (Optimisation of Monitoring for Clinical Research Studies; Journot 2015), TEMPER (TargetEd Monitoring: Prospective Evaluation and Refinement; Stenning 2018a), START Monitoring Substudy (Strategic Timing of AntiRetroviral Treatment; Hullsiek 2015; Wyman Engen 2020), and MONITORING (Fougerou-Leurent 2019).

Description of the methods being investigated

Traditional trial monitoring consists of intensive on-site monitoring strategies comprising frequent on-site visits and up to 100% SDV. Risk-based monitoring is a

new strategy that recognizes that not all clinical trials require the same approach to quality control and assurance (Stenning 2018a), and allows for stratification based on risk indicators assessed during the trial or before it starts. Risk-based strategies differ in their risk assessment approaches as well as in their implementation and extent of on-site and central monitoring components. They are also referred to as risk-adapted or risk-proportionate monitoring strategies. In this review, which is based on our published protocol (Klatte 2019), we investigated the effects of monitoring methods on ensuring patient rights and safety, and the validity of trial data. These key elements of clinical trial conduct are assessed by monitoring for critical or major violation of GCP objectives, according to the classification of GCP findings described in EMA 2017.

Monitoring strategies empirically evaluated in studies

All the monitoring strategies eligible for this review introduced new methods that might be effective in directing monitoring components and resources guided by a risk evaluation or prioritization.

1. Risk-based monitoring strategies

The risk-based strategy proposed by Brosteanu and colleagues is based on an initial assessment of the risk associated with an individual trial protocol (ADAMON: Brosteanu 2009). The implementation of this three-level risk assessment focuses on critical data and procedures describing the risk associated with a therapeutic intervention and incorporates an assessment of indicators for patient-related risks, indicators of robustness, and indicators for site-related risks. Trial-specific risk analysis then informs a monitoring plan that contains on-site elements as well as central and statistical monitoring methods to a different extent corresponding to the judged risk level. The consensus risk-assessment scale (RAS) and risk-adapted monitoring plan (RAMP) developed by Journot and colleagues in 2010 consists of a four-level initial risk assessment, leading to monitoring plans of four levels of intensity (OPTIMON; Journot 2011). The optimized monitoring strategy concentrates on the main scientific and regulatory aspects, compliance with requirements for patient consent and serious adverse events (SAE), and the frequency of serious errors concerning the validity of the trial's main results and the trial's eligibility criteria (Chene 2008). Both strategies incorporate central monitoring methods that help to specify the monitoring intervention for each study site within the framework of their assigned risk level.

2. Central monitoring with triggered on-site visits

The triggered on-site monitoring strategy suggested by the Medicines and Healthcare products Regulatory Agency, Medical Research Council (MRC), and UK Department of Health includes an initial risk assessment on the basis of the intervention and design of the trial and a resulting monitoring plan for different trial sites that is continuously updated through centralized monitoring. Over the course of a clinical trial, sites are prioritized for on-site visits based on predefined central monitoring triggers (Meredith 2011; TEMPER: Stenning 2018a).

3. Central and local monitoring

A strategy that is mainly based on central monitoring, combined with a local quality control provided by qualified personnel on-site, is being evaluated in the START Monitoring Substudy (Hullsiek 2015). In this study, continuous central monitoring uses descriptive statistics on the consistency and quality of the data and data completeness. Semi-annual performance reports are generated for each site, focusing on the key variables/endpoints regarding patients' safety (SAEs, eligibility violations) and data quality. This evaluates whether adding on-site monitoring to these procedures leads to differences in the participant-level composite outcome of monitoring findings.

4. Monitoring with targeted or remote source data verification

The monitoring strategy developed for the MONITORING study is characterized by a targeted SDV in which only regulatory and scientific key data are verified (Fougerou-Leurent 2019). This strategy is compared to full SDV and assessed based on final data quality and costs. One pilot study assessed a new strategy of remote SDV where documents were accessed via electronic health records, clinical data repositories, web-based access technologies, or authentication and auditing tools (Mealer 2013).

5. On-site initiation visits upon request

In this monitoring strategy, systematic initiation visits at all sites are replaced by initiation visits that take place only upon investigators' request at a site (Lièvard 2006).

How these methods might work

The intention for risk-based monitoring methods is to increase the efficiency of monitoring and to optimize resource use by directing the amount and content of monitoring visits according to an initially assessed risk level of an individual trial. These new methods should be at least non-inferior in detecting major or critical violation of essential GCP objectives, according to EMA 2017, and might even be superior in terms of prioritizing monitoring content. The risk assessment preceding the risk-based monitoring plan should consider the likelihood of errors occurring in key aspects of study performance, and the anticipated effect of such errors on the protection of participants and the reliability of the trial's results (Landray 2012). Trials within a certain risk category are initially assigned to a defined monitoring strategy, which remains adjustable throughout the conduct of the trial and should always match the needs of the trial and specific trial sites. This flexibility is an advantage, considering the heterogeneity of study designs and participating trial sites. Central monitoring would also allow for continuous verification of data quality based on prespecified triggers and thresholds, and would enable early intervention in cases of procedural or data-recording errors. Besides the detection of missing or invalid data, trial entry procedures and protocol adherence, as well as other performance indicators, can be monitored through a continuous analysis of electronically captured data (Baigent 2008). In addition, comparison with external sources may be undertaken to validate information contained in the data set; and the identification of poorly performing sites would ensure a more targeted application of on-site monitoring resources. Use of methods that take advantage of the increasing use of electronic systems (e.g. electronic case report forms [eCRFs]) may allow data to be checked by automated means and allows the application of entry rules supporting up-to-date, high-quality data. These methods would also ensure patient rights and safety while simultaneously improving trial management and optimizing trial conduct. Adaptations in the monitoring approach toward a reduction of on-site monitoring visits, provided that patient rights and safety are ensured, could allow the application of resources to the most crucial components of the trial (Journot 2011).

In order to evaluate whether these new risk-based monitoring approaches are non-inferior to the traditional extensive on-site monitoring, an assessment of differences in critical and major findings during monitoring activities is essential. Monitoring findings are determined with respect to patient safety, patient rights, and reliability of the data, and classified as critical and major according to the classification of GCP findings described in the *Procedures for reporting of GCP inspections requested by the Committee for Medicinal Products for Human Use* (EMA 2017). Critical findings are conditions, practices, or processes that adversely affect the rights, safety, or well being of the participants or the quality and integrity of data. Major findings are conditions, practices, or processes that might adversely affect the rights, safety, or well being of the participants or the quality and integrity of data.

Why it is important to do this review

There is insufficient information to guide the choice of monitoring approaches consistent with GCP to use in any given trial, and there is a lack of evidence on the effectiveness of suggested monitoring approaches. This has resulted in high heterogeneity in the monitoring practices used by research institutions, especially in the academic setting (Love 2020; Morrison 2011). A guideline describing which type of monitoring strategy is most effective for clinical trials in terms of patient rights and safety, and data quality, is urgently needed for the academic clinical trial setting. Evaluating the benefits and disadvantages of different risk-based monitoring strategies, incorporating components of central or targeted and triggered (or both) monitoring versus intensive on-site monitoring, might lead to a consensus on how effective these new approaches are. In addition, evaluating the evidence of effectiveness could provide information on the extent to which on-site monitoring content (such as SDV or frequency of site visits) can be adapted or supported by central monitoring interventions. In this review, we explored whether monitoring that incorporates central (including statistical) components could be extended to support the overall management of study quality in terms of participant recruitment and follow-up.

The risk-based monitoring interventions that are eligible for this review incorporate on-site and central monitoring components, which may vary extent and procedural structure. In line with the recommendation from the Clinical Trials Transformation Initiative (Grignolo 2011), it is crucial to systematically analyze and compare the existing evidence so that best practices may be established. This review may

facilitate the sharing of current knowledge on effective monitoring strategies, which would help trialists, support units, and monitors to choose the best strategy for their trials. Evaluation of the impact of a change of monitoring approaches on data quality and study cost is relevant for the effective adjustment of current monitoring strategies. In addition, evaluating the effectiveness of these new monitoring approaches in comparison with intensive on-site monitoring might reveal possible methods to replace or support on-site monitoring strategies by taking advantage of the increasing use of electronic systems and resulting opportunities to implement statistical analysis tools.

Objectives

To evaluate the advantages and disadvantages of different monitoring strategies (including risk-based strategies and others) for clinical intervention studies examined in prospective comparative studies of monitoring interventions.

Methods

Criteria for considering studies for this review

Types of studies

We included randomized or non-randomized prospective, empirical evaluation studies that assessed monitoring strategies in one or more clinical intervention studies. These types of embedded studies have recently been called 'studies within a trial' (SWATs) (Anon 2012; Treweek 2018a). We excluded retrospective studies because of their limitations with respect to outcome standardization and variable definitions.

We followed the Cochrane Effective Practice and Organisation of Care (EPOC) Group definitions for the eligible study designs (EPOC 2016).

We applied no restrictions on language or date of publication.

Types of data

We extracted information about monitoring processes as well as evaluations of the comparison and advantages/disadvantages of different monitoring approaches. We included data from published and unpublished studies, and grey literature, that

compared different monitoring strategies (e.g. standard monitoring versus a risk-based approach).

Study characteristics of interest were:

1. Monitoring interventions;
2. Risk assessment characteristics;
3. Finding rates of serious/critical audits;
4. Impact on participant recruitment and follow-up; and
5. Costs.

Types of methods

We included studies that compared:

1. A risk-based monitoring strategy versus an intensive on-site monitoring strategy for prospective intervention studies; or
2. Any other prospective comparison of monitoring strategies for intervention studies.

Types of outcome measures

Specific outcome measures were not part of the eligibility criteria.

Primary outcomes

1. Combined outcome of critical and major monitoring findings in prospective intervention studies. Different error domains of critical and major monitoring findings were combined in the primary outcome measure (eligibility violations, informed-consent violations, findings that raise doubt about the accuracy or credibility of key trial data and deviations of intervention from the trial protocol, errors in endpoint assessment, and errors in SAE reporting).

Critical and major findings were defined according to the classification of GCP findings described in EMA 2017, as follows.

1. Critical findings: conditions, practices, or processes that adversely affected the rights, safety, or well being of the study participants or the quality and integrity of data. Observations classified as critical may have included a pattern of deviations classified either as major, or bad quality of the data or

absence of source documents (or both). Manipulation and intentional misrepresentation of data was included in this group.

2. Major findings: conditions, practices, or processes that might adversely affect the rights, safety, or well being of the study participants or the quality and integrity of data (or both). Major observations are serious deficiencies and are direct violations of GCP principles. Observations classified as major may have included a pattern of deviations or numerous minor observations (or both).

Our protocol stated definitions of combined outcomes of critical and major findings in the respective studies (Table 1) (Klatte 2019).

Secondary outcomes

1. Individual components of the primary outcome:
    1. Major eligibility violations;
    2. Major informed-consent violations;
    3. Findings that raised doubt about the accuracy or credibility of key trial data and deviations of intervention from the trial protocol (with impact on patient safety or data validity);
    4. Errors in endpoint assessment; and
    5. Errors in SAE reporting.
2. Impact of the monitoring strategy on participant recruitment and follow-up.
3. Effect of the monitoring strategy on resource use (costs).
4. Qualitative research data or process evaluations of the monitoring interventions.

Search methods for identification of studies

Electronic searches

We conducted a comprehensive search (May 2019) using a search strategy that we developed together with an experienced scientific information specialist (HE). We systematically searched the Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, and Embase via Elsevier for relevant published literature (PubMed strategy shown below, all searches in full in the Appendix 1). The search strategy for all three databases was peer-reviewed according to PRESS guidelines (McGowan 2016) by the Cochrane information specialist, Irma Klerings (Cochrane

Austria). We also searched the online SWAT repository (go.qub.ac.uk/SWAT-SWAR). We applied no restrictions regarding language or date of publication. Since our original search for the review took place in May 2019, we performed an updated search in March 2021 to ensure that we included all eligible studies up to that date. Our updated search identified no additional eligible studies.

We used the following terms to identify prospective studies that compared different strategies for trial monitoring:

1. triggered monitoring;
2. targeted monitoring;
3. risk-adapted monitoring;
4. risk adapted monitoring;
5. risk-based monitoring;
6. risk based monitoring;
7. centralized monitoring;
8. centralised monitoring;
9. statistical monitoring;
10. on site monitoring;
11. on-site monitoring;
12. monitoring strategy;
13. monitoring method;
14. monitoring technique;
15. trial monitoring; and
16. central monitoring.

The search was intended to identify randomized trials and non-randomized intervention studies that evaluated monitoring strategies in a prospective setting. Therefore, we modified the Cochrane sensitivity-maximizing filter for randomized trials (Lefebvre 2011).

PubMed search strategy:

("on site monitoring"[tiab] OR "on-site monitoring"[tiab] OR "monitoring strategy"[tiab] OR "monitoring method"[tiab] OR "monitoring technique"[tiab] OR "triggered monitoring"[tiab] OR "targeted monitoring"[tiab] OR "risk-adapted monitoring"[tiab] OR "risk adapted monitoring"[tiab] OR "risk-based monitoring"[tiab] OR "risk based monitoring"[tiab] OR "risk proportionate"[tiab] OR "centralized monitoring"[tiab] OR

"centralised monitoring"[tiab] OR "statistical monitoring"[tiab] OR "central monitoring"[tiab]) AND ("prospective" [tiab] OR "prospectively" [tiab] OR randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized [tiab] OR placebo [tiab] OR drug therapy [sh] OR randomly [tiab] OR trial [tiab] OR groups [tiab]) NOT (animals [mh] NOT humans[mh])

Searching other resources

We hand searched reference lists of included studies and similar systematic reviews to find additional relevant study articles (Horsley 2011). In addition, we searched the grey literature (Appendix 2) (i.e. conference proceedings of the Society for Clinical Trials and the International Clinical Trials Methodology Conference), and trial registries (ClinicalTrials.gov, the World Health Organization International Clinical Trials Registry Platform, the European Union Drug Regulating Authorities Clinical Trials Database, and ISRCTN) for ongoing or unpublished prospective studies. Finally, we collaborated closely with researchers of already identified eligible studies (e.g. OPTIMON, ADAMON, INSIGHT START, and MONITORING) and contacted researchers to identify further studies (and unpublished data, if available).

Data collection and analysis

Data collection and analysis methods were based on the recommendations described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2020) and *Methodological Expectations for the Conduct of Cochrane Intervention Reviews* (Higgins 2016).

Selection of studies

After elimination of duplicate records, two review authors (KK and PA) independently screened titles and abstracts for eligibility. We retrieved potentially relevant studies as full-text reports and two review authors (KK and MB) independently assessed these for eligibility, applying prespecified criteria (see: Criteria for considering studies for this review). We resolved any disagreements between review authors by discussion until consensus was reached, or by involving a third review author (CPM). We documented the study selection process in a flow diagram, as described in the PRISMA statement (Moher 2009).

Data extraction and management

For each eligible study, two review authors (KK and MMB) independently extracted information on a number of key characteristics, using electronic data collection forms (Appendix 3). Data were extracted in Epi-Reviewer 4 (Thomas 2010). We resolved any disagreements by discussion until consensus was reached, or by involving a third review author (MB). We contacted authors of included studies directly when target information was unreported or unclear to clarify or complete extracted data. We summarized the data qualitatively and quantitatively (where possible) in the Results section, below. If meta-analysis of the primary or secondary outcomes was not applicable due to considerable methodological heterogeneity between studies, we reported the results qualitatively only.

Extracted study characteristics included the following.

1. General information about the study: title, authors, year of publication, language, country, funding sources.
2. Methods: study design, allocation method, study duration, stratification of sites (stratified on risk level, country, projected enrolment, etc.).
3. Characteristics of clinical trials included in the prospective comparison of monitoring strategies:
    1. Design (randomized or other prospective intervention trial);
    2. Setting (primary care, tertiary care, community, etc.);
    3. National or multinational;
    4. Study population;
    5. Total number of sites randomized/analyzed;
    6. Inclusion/exclusion criteria;
    7. IMP risk category;
    8. Support from clinical trials unit (CTU) or clinical research organization for host trial or evidence for experienced research team; and
    9. Trial phase.
4. Intervention (components related to the applied monitoring strategy, including theoretical basis):
    1. Number of sites randomized/allocated to groups (specifying number of sites or clusters);
    2. Duration of intervention period;
    3. Risk assessment characteristics (follow-up questions)/triggers or thresholds that induce on-site monitoring (follow-up questions);
    4. Frequency of monitoring visits;

5. Extent of on-site monitoring;
6. Frequency of central monitoring reports;
7. Number of monitoring visits per participant;
8. Cumulative monitoring time on-site;
9. Mean number of monitoring visits per site;
10. Delivery (procedures used for central monitoring: structure/components of on-site monitoring/triggers/thresholds);
11. Who performed the monitoring (study team, trial staff; qualifications of monitors);
12. Degree of SDV (median number of participants undergoing SDV); and
13. Co-interventions (site/study-specific co-interventions).

5. Outcomes: primary and secondary outcomes, individual components of combined primary outcome, outcome measures and scales, time points of measurement, statistical analysis of outcome data.
6. Data to assess the risk of bias of included studies (e.g. random sequence generation, allocation concealment, blinding of outcome assessors, performance bias, selective reporting, or other sources of bias).

Assessment of risk of bias in included studies

Two review authors (KK and MMB) independently assessed the risk of bias in each included study using the criteria described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2020) and the Cochrane EPOC Review Group (EPOC 2017). The domains provided by these criteria were evaluated for all included randomized studies and assigned ratings of low, high, or unclear risk of bias. We assessed non-randomized studies using the ROBINS-I tool of bias assessment for non-randomized studies separately (Higgins 2020, Chapter 25).

We assessed the risk of bias for randomized studies as follows.

Selection bias

Generation of the allocation sequence

1. If sequence generation was truly random (e.g. computer generated): low risk.
2. If sequence generation was not specified and we were unable to obtain relevant information from study authors: unclear risk.
3. If there was a quasi-random sequence generation (e.g. alternation): high risk.

4. Non-randomized trials: high risk.

Concealment of the allocation sequence (steps taken prior to the assignment of intervention to ensure that knowledge of the allocation was not possible)

1. If opaque, sequentially numbered envelopes were used or central randomization was performed by a third party: low risk.
2. If the allocation concealment was not specified and we were unable to ascertain whether the allocation concealment had been protected before and until assignment: unclear risk.
3. Non-randomized trials and studies that used inadequate allocation concealment: high risk.

For non-randomized studies, we further assessed if investigators attempted to balance groups by design (control for selection bias) and attempted to control for confounding: high risk according to Cochrane risk of bias tool, but we considered the risk of bias control efforts in our judgment of the certainty of the evidence according to GRADE.

Performance bias

It is not practicable to blind participating sites and monitors to the intervention to which they were assigned because of the procedural differences of monitoring strategies.

Detection bias (blinding of the outcome assessor)

1. If the assessors performing audits had knowledge of the intervention and thus outcomes were not assessed blindly: high risk.
2. If we could not ascertain whether assessors were blinded and study authors did not provide information to clarify: unclear risk.
3. If outcomes were assessed blindly: low risk.

Attrition bias

We did not expect to have missing data for our primary outcome (i.e. the rates of serious/critical audit findings at the end of the host clinical trials; and because missing participants were not audited, missing data in the proportion of critical findings were not expected). However, for the statistical power of the individual study

outcomes, missing data for participants and site accrual could be an issue and is discussed below (Discussion).

Selective reporting bias

We investigated whether all outcomes mentioned in available study protocols, registry entries, or methodology sections of study publications were reported in results sections.

1. If all outcomes in the methodology or outcomes specified in the study protocol were not reported in the results, or if outcomes reported in the results were not listed in the methodology or in the protocol: high risk.
2. If outcomes were only partly reported in the results, or if an obvious outcome was not mentioned in the study: high risk.
3. If information is unavailable on the prespecified outcomes and the study protocol: unclear risk.
4. If all outcomes were listed in the protocol/methodology section and reported in the results: low risk.

Other potential sources of bias

1. If there was one or more important risk of bias (e.g. flawed study design): high risk.
2. If there was incomplete information regarding a problem that may have led to bias: unclear risk.
3. If there was no evidence of other sources of bias: low risk.

We assessed the risk of bias for non-randomized studies as follows.

Pre-intervention domains

1. Confounding – baseline confounding occurs when one or more prognostic variables (factors that predict the outcome of interest) also predicts the intervention received at baseline.
2. Selection bias (bias in selection of participants into the study) – when exclusion of some eligible participants, or the initial follow-up time of some participants, or some outcome events, is related to both intervention and outcome, there will be an association between interventions and outcome even if the effect of interest is truly null.

At-intervention domain

1. Information bias – bias in classification of interventions, i.e. bias introduced by either differential or non-differential misclassification of intervention status.

Post-intervention domains

1. Confounding – bias that arises when there are systematic differences between experimental intervention and comparator groups in the care provided, which represent a deviation from the intended intervention(s).
2. Selection bias – bias due to exclusion of participants with missing information about intervention status or other variables such as confounders.
3. Information bias – bias introduced by either differential or non-differential errors in measurement of outcome data.
4. Reporting bias – bias in selection of the reported result.

Judgment

Risk of bias judgment	**Interpretation**
Low risk of bias	The study was comparable to a well-performed randomized trial with regard to this domain.
Moderate risk of bias	The study was sound for a non-randomized study with regard to this domain but could not be considered comparable to a well-performed randomized trial.
Serious risk of bias	The study had some important problems in this domain.
Critical risk of bias	The study was too problematic in this domain to provide any useful evidence on the effects of intervention.
No information	No information on which to base a judgment about risk of bias for this domain.
From Higgins 2020.	

Measures of the effect of the methods

We conducted a comparative analysis of the impact of different risk-based monitoring strategies on data quality and patient rights and safety measures, for example by the proportion of critical findings.

If meta-analysis was appropriate, we analyzed dichotomous data using a risk ratio with a 95% confidence interval (CI). We analyzed continuous data using mean differences with a 95% CI if the measurement scale was the same. If the scale was different, we used standardized mean differences with 95% CIs.

Unit of analysis issues

Included studies could differ in outcomes chosen to assess the effects of the respective monitoring strategy. Critical/serious audit findings could be reported on a participant level, per finding event, or per site. Furthermore, components of the primary endpoints could vary between studies. We specified the study outcomes as defined in the study protocols or reports, and only meta-analyzed outcomes that were based on similar definitions. In addition, we compared individual components of the primary outcome if these were consistently defined across studies (e.g. eligibility violations).

Cluster randomized trials have been highlighted separately to individually randomized trials. We reported the baseline comparability of clusters and considered statistical adjustment to reduce any potential imbalance. We estimated the intracluster correlation coefficient (ICC), as described by Higgins 2020, using information from the study (if available) or from an external estimate from a similar study. We then conducted sensitivity analyses to explain variation in ICC values.

Dealing with missing data

We contacted authors of included studies in an attempt to obtain unpublished data or additional information of value for this review (Young 2011). Where a study had been registered and a relevant outcome was specified in the study protocol but no results were reported, we contacted the authors and sponsors to request study reports. We created a table to summarize the results for each outcome. We narratively explored the potential impact of missing data in our Discussion.

Assessment of heterogeneity

When we identified methodological heterogeneity, we did not pool results in a meta-analysis. Instead, we qualitatively synthesized results by grouping studies with similar designs and interventions, and described existing methodological heterogeneity (e.g. use of different methods to assess outcomes). If study

characteristics, methodology, and outcomes were sufficiently similar across studies, we quantitatively pooled results in a meta-analysis and assessed heterogeneity by visually inspecting forest plots of included studies (location of point estimates and the degree to which CIs overlapped), and by considering the results of the Chi^2 test for heterogeneity and the I^2 statistic. We followed the guidance outlined in Higgins 2020 to quantify statistical heterogeneity using the I^2 statistic:

1. 0% to 40% might not be important;
2. 30% to 60% may represent moderate heterogeneity;
3. 50% to 90% may represent substantial heterogeneity;
4. 75% to 100%: considerable heterogeneity.

The importance of the observed value of the I^2 statistic depends on the magnitude and direction of effects, and the strength of evidence for heterogeneity (e.g. P value from the Chi^2 test, or a credibility interval for the I^2 statistic). If our I^2 value indicated that heterogeneity was a possibility and either the Tau^2 was greater than zero, or the P value for the Chi^2 test was low (less than 0.10), heterogeneity may have been due to a factor other than chance.

Possible sources of heterogeneity from the characteristics of host trials included:

1. Design (randomized or other prospective intervention trial);
2. Setting (primary care, tertiary care, community, etc.);
3. IMP risk category;
4. Trial phase;
5. National or multinational;
6. Support from a CTU or clinical research organization for host trial or evidence for an experienced research team; and
7. Study population.

Possible sources of heterogeneity from the characteristics of methodology studies included:

1. Study design;
2. Components of outcome;
3. Method of outcome assessment;
4. Level of outcome (participant/site); and
5. Classification of monitoring findings.

Due to high heterogeneity of studies, we used the random-effects method (DerSimonian 1986), which incorporates an assumption that the different studies are estimating different, yet related, intervention effects. As described in Section 9.4.3.1 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2020), the method is based on the inverse-variance approach, making an adjustment to the study weights according to the extent of variation, or heterogeneity, among the varying intervention effects. Due to the small number of studies included into the meta-analyses and the high heterogeneity of the studies in the number of participants or sites included in the analysis we decided to use the inverse variance method. The inverse variance estimates the amount of variation across studies by comparing each study's result with an inverse-variance fixed-effect meta-analysis result. This resulted in a more appropriate weighting of the included studies according to the extent of variation.

Assessment of reporting biases

To decrease the risk of publication bias affecting the findings of the review, we applied various search approaches using different resources. These included grey literature searching and checking reference lists (see Search methods for identification of studies). If 10 or more studies were available for a meta-analysis, we would have created a funnel plot to investigate whether reporting bias may have existed unless all studies were of a similar size. If we noticed asymmetry, we would not have been able to conclude that reporting biases existed, but we would have considered the sample sizes and presence (and possible influence) of outliers and discussed potential explanations, such as publication bias or poor methodological quality of included studies, and performed sensitivity analyses.

Data synthesis

Data were synthesized using tables to compare different monitoring strategies. We also reported results by different study designs. This was accompanied by a descriptive summary in the Results. We used Review Manager 5 to conduct our statistical analysis and undertake meta-analysis, where appropriate (Review Manager 2014).

If meta-analysis of the primary or secondary outcomes was not possible, we reported the results qualitatively.

Two review authors (KK and MB) assessed the quality of the evidence. Based on the methods described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2020) and GRADE (Guyatt 2013a; Guyatt 2013b), we created summary of findings tables for the main comparisons of the review. We presented all primary and secondary outcomes outlined in the Types of outcome measures section. We described the study settings and number of sites addressing each outcome. For each assumed risk of bias cited, we provided a source and rationale, and we implemented the GRADE system to assess the quality of the evidence using GRADEpro GDT software or the GRADEpro GDT app (GRADEpro GDT). If meta-analysis was not appropriate or the units of analysis could not be compared, we presented results in a narrative summary of findings table. In this case, the imprecision of the evidence was an issue of concern due to the lack of a quantitative effect measure.

Subgroup analysis and investigation of heterogeneity

If visual inspection of the forest plots, Chi^2 test, I^2 statistic, and Tau^2 statistic indicated that statistical heterogeneity might be present; we carried out exploratory subgroup analysis. A subgroup analysis was deemed appropriate if the included studies satisfied criteria assessing the credibility of subgroup analyses (Oxman 1992; Sun 2010).

The following was our a priori subgroup: monitoring strategies using very similar approaches and consistent outcomes.

Sensitivity analysis

We conducted sensitivity analyses restricted to:

1. Peer-reviewed and published studies only (i.e. excluding unpublished studies); and
2. Studies at low risk of bias only (i.e. excluding non-randomized studies and randomized trials without allocation concealment; Assessment of risk of bias in included studies).

Results

Description of studies

See: Characteristics of included studies and Characteristics of excluded studies tables.

Results of the search

Our search of CENTRAL, PubMed, and Embase resulted in 3105 unique citations, 3103 citations after removal of duplicates and two additional citations that were identified through reference lists of relevant articles. After screening titles and abstracts, we sought the full texts of 51 records to confirm inclusion or clarify uncertainties regarding eligibility. Eight studies (14 articles) were eligible for inclusion. The results of six of these were published as full papers (Brosteanu 2017b; Fougerou-Leurent 2019; Lienard 2006; Mealer 2013; Stenning 2018b; Wyman 2020), one study was published as an abstract only (Knott 2015), and one study was submitted for publication (Journot 2017). We did not identify any ongoing eligible studies or studies awaiting classification.

See Figure 1 (flow diagram).

Included studies

Seven of the eight included studies were government or charity funded. The other was industry funded (Lienard 2006). The primary objectives were heterogeneous and included non-inferiority evaluations of overall monitoring performance as well as single elements of monitoring (SDV, initiation visit); see Characteristics of included studies table and Table 2.

Overall, there were five groups of comparisons:

1. Risk-based monitoring guided by an initial risk assessment and information from central monitoring during study conduct versus extensive on-site monitoring (ADAMON: Brosteanu 2017b; OPTIMON: Journot 2017);
2. Central monitoring with triggered on-site visits versus regular (triggered) on-site visits (Knott 2015; TEMPER: Stenning 2018b);

3. Central statistical monitoring and local monitoring at sites with annual on-site visits (untriggered) versus central statistical monitoring and local monitoring at sites only (START-MV: Wyman 2020);
4. 100% on-site SDV versus remote SDV (Mealer 2013) or targeted SDV (MONITORING: Fougerou-Leurent 2019); and
5. On-site initiation visit versus no on-site initiation visit (Lièvard 2006).

Since there was substantial heterogeneity in the investigated monitoring strategies and applied study designs, a short overview of each included study is provided below.

General characteristics of individual included studies

1. Risk-based versus extensive on-site monitoring

The ADAMON study was a cluster randomized non-inferiority trial comparing risk-adapted monitoring with extensive on-site monitoring at 213 sites participating in 11 international and national clinical trials (all in secondary or tertiary care and with adults and children as participants) (Brosteanu 2017b). It included only randomized, multicenter clinical trials (at least six trial sites) with a non-commercial sponsor and had standard operating procedures (SOPs) for data management and trial supervision as well as central monitoring of at least basic extent. The prior risk analysis categorized trials into two of three different risk categories and trials were monitored according to a prespecified monitoring plan for their respective risk category. While the RAMP for the highest risk category was only marginally less extensive than full on-site monitoring, risk-based monitoring strategies for the lower risk categories relied on information from central monitoring and previous visits to determine the amount of on-site monitoring. This resulted in a marked reduction of on-site monitoring for sites without noticeable problems, limited to key data monitoring (20% to 50%). Only studies that had been classified as either intermediate risk or low risk based on the trial-specific risk analysis (Brosteanu 2009) were included in the study. From the 11 clinical trials, 156 sites were audited by ADAMON-trained auditors and included in the final analysis. The analysis included a meta-analysis of results obtained within each trial.

The OPTIMON study was a cluster randomized non-inferiority trial evaluating a risk-based monitoring strategy within 22 national and international multicenter studies (Journot 2017). The 22 trials included 15 randomized trials, four cohort studies, and

three cross-sectional studies in the secondary care setting with adults, children, and older people as participants. All trials involved methodology and management centers or CTUs, had at least two years of experience in multicenter clinical research studies, and SOPs in place. A total of 83 sites were randomized to one of two different monitoring strategies. The risk-based monitoring approach consisted of an initial risk assessment with four outcome levels (low, moderate, substantial, and high) and a standardized monitoring plan, where on-site monitoring increased with the risk level of the trial (Journot 2011). The study aimed to assess whether such a risk-adapted monitoring strategy provided results similar to those of the 100% on-site strategy on the main study quality criteria, and, at the same time, improved other aspects such as timeliness and costs (Journot 2017). Only 759 participants from 68 sites were included in the final analysis, because of insufficient recruitment at 15 of the 83 randomized sites. The difference between strategies was evaluated by the proportion of participants without remaining major non-conformities in all of the four assessed error domains (consent violation, SAE reporting violation, eligibility violation, and errors in primary endpoint assessment) assessed after trial monitoring by the OPTIMON team. The overall comparison of strategies was estimated using a generalized estimating equation (GEE) model, adjusted for risk level and intra-site, intra-patient correlation common to all sites.

2. Central monitoring with triggered on-site visits versus regular (untriggered) on-site visits

Knott 2015 was a monitoring study embedded in a large international multicenter trial evaluating the ability of central statistical monitoring procedures to identify sites with problems. Monitoring findings at sites during on-site monitoring visits targeted as a result of central statistical monitoring procedures were compared to monitoring findings at sites chosen by regional coordinating centers. Oversight of the clinical multicenter trial was supported by central statistical monitoring that identified high scoring sites as priority for further investigation and triggered a targeted on-site visit. In order to compare targeted on-site visits with regular on-site visits, high scoring sites, and some low scoring sites in the same countries identified by the country teams as potentially problematic were visited. The decision about which of the low scoring sites would benefit most from an on-site visit was based on prior experience of the regional coordinating centers with the site. Twenty-one sites (12 identified by central statistical monitoring, nine others as comparators) received a comprehensive

monitoring visit from a senior monitor and the number of major and minor findings was compared between the two types of visits (targeted versus regular visit).

The TEMPER study (Stenning 2018b) was conducted in three ongoing phase III randomized multicenter oncology trials with 156 UK sites (Diaz-Montana 2019a). All three included trials were in secondary care settings, were conducted and monitored by the MRC CTU at University College London, and were sponsored by the UK MRC and employed a triggered monitoring strategy. The study used a matched-pair design to assess the ability of targeted monitoring to distinguish sites at which higher and lower rates of protocol or GCP violations (or both) would be found during site visits. The targeted monitoring strategy was based on trial data that were scrutinized centrally with prespecified triggers provoking an on-site visit when certain thresholds had been crossed. In order to compare this approach to standard on-site monitoring, a matching algorithm proposed untriggered sites to visit by minimizing differences in 1. Number of participants and 2. Time since first participant randomized, and by maximizing differences in trigger score. Monitoring data from 42 matched paired visits (84 visits) at 63 sites were included in the analysis of the TEMPER study. The monitoring strategy was assessed over all trial phases and the outcome was assessed by comparing the proportion of sites with one or more major or critical finding not already identified through central monitoring or a previous visit ('new' findings). The prognostic value of individual triggers was also assessed.

3. Central and local monitoring with annual on-site visits versus central and local monitoring only

The START Monitoring Substudy was conducted within one large international, publicly funded randomized clinical trial (START – Strategic Timing of AntiRetroviral Treatment) (Wyman 2020). The monitoring substudy included 4371 adults from 196 secondary care sites in 34 countries. All clinical sites were associated with one of four INSIGHT coordinating centers and central monitoring by the statistical center was done continuously using central databases. In addition, local monitoring of regulatory files, SDV, and study drug management was performed by site staff semi-annually. In the monitoring substudy, sites were randomized to receive annual on-site monitoring in addition to central and local monitoring or to central and local monitoring alone. The composite monitoring outcome consisted of eligibility violations, informed consent violations, intervention (use of antiretroviral therapy as initial treatment not permitted by protocol), primary endpoint and SAE reporting. In

the analysis, a generalized estimation equation model with fixed effects to account for clustering was used and each component of the composite outcome was evaluated to interpret the relevance of the overall composite result.

4. Traditional 100% source data verification versus remote or targeted source data verification

Mealer 2013 was a pilot study on remote SDV in two national clinical trials' networks in which study participants were randomized to either remote SDV followed by on-site verification or traditional on-site SDV. Thirty-two participants in randomized and other prospective clinical intervention trials within the adult trials network and the pediatric network were included in this monitoring study. A sample of participants in this secondary and tertiary care setting, who were due for an upcoming monitoring visit that included full SDV were randomized and stratified at each individual hospital. The five study sites had different health information technology infrastructures, resulting in different approaches to enable remote access and remote data monitoring. Only participants randomized to remote SDV had a previsit remote SDV performed prior to full SDV at the scheduled visit. Remote SDV was performed by validating the data elements captured on CRFs submitted to the coordinating center using the same data verification protocols that were used during on-site visits and remote monitors had telephone access to the local coordinators. The primary outcome was the proportion of data values identified versus not identified for both monitoring strategies. As an additional economic outcome, the total time required for the study monitor to verify a case report item with either remote or on-site monitoring form was analyzed.

The MONITORING study was a prospective cross-over study comparing full SDV, where 100% of data was verified for all participants, and targeted SDV, where only key data were verified for all participants (Fougerou-Leurent 2019). Data from 126 participants from one multinational and five national clinical trials managed by the Clinical Investigation Center at the Rennes University Hospital INSERM in France were included in the analysis. These studies included five randomized trials and one non-comparative pilot single-center phase II study taking place in either tertiary or secondary care units. Key data verified by the targeted SDV included informed consent, inclusion and exclusion criteria, main prognostic variables at inclusion, primary endpoint, and SAEs. The same CRFs were analyzed with full or targeted SDV. SDV of both strategies was followed by the same data-management program,

detecting missing data and checking consistency, on final data quality, global workload, and staffing costs. Databases of full SDV and targeted SDV after the data-management process were compared and identified discrepancies were considered as remaining errors with targeted monitoring.

5. Systematic on-site initiation visit versus on-site initiation visit upon request

Liènard 2006 was a monitoring study within a large international randomized trial of cancer treatment. A total of 573 participants from 135 centers in France were randomized on a center level to receive an on-site initiation visit for the study or no initiation visit. Although the study was terminated early, 68 secondary care centers, stratified by center type (private versus public hospital), had entered at least one participant into the study. The study was terminated because the sponsor decided to redirect on-site monitoring visits to centers in which a problem had been identified. The aim of this monitoring study was to assess the impact of on-site initiation visits on the following outcomes: participant recruitment, quantity and quality of data submitted to the trial coordinating office, and participants' follow-up time. On-site initiation visits by monitors included review of the protocol, inclusion and exclusion criteria, safety issues, randomization procedure, CRF completion, study planning, and drug management. Investigators requesting on-site visits were visited regardless of the allocated randomized group and results were analyzed by randomized group.

Characteristics of the monitoring strategies

There was substantial heterogeneity in the characteristics of the evaluated monitoring strategies. Table 2 summarizes the main components of the evaluated strategies.

Central monitoring components within the monitoring strategies

Use of central monitoring to trigger/adjust on-site monitoring

Central monitoring plays an important role in the implementation of risk-based monitoring strategies. An evaluation of site performance through continuous analysis of data quality can be used to direct on-site monitoring to specific sites or support remote monitoring methods. A reduction in on-site monitoring for certain trials was accompanied by central monitoring which also enabled additional on-site interference in cases of low-quality performance related to data quality, completeness, or patient

rights and safety of specific sites. Six included studies used central monitoring methods to support their new monitoring strategy (ADAMON: Brosteanu 2017b; OPTIMON: Journot 2017; Knott 2015; Mealer 2013; TEMPER: Stenning 2018b; START Monitoring Substudy: Wyman 2020). Four of these studies used central monitoring information to trigger or delegate on-site monitoring. In the ADAMON study, part of the monitoring plan for the lower- and medium-risk studies comprised a regular assessment of the trial sites as 'with' or 'without noticeable problems' (Brosteanu 2017b). Classification as a site 'with noticeable problems' resulted in an increased number of on-site visits per year. In the OPTIMON study, major problems (patient rights and safety, quality of results, regulatory aspects) triggered an additional on-site visit for level B and C sites, or a first on-site visit for level A sites (Journot 2017). All entered data were checked for completeness and consistency for all participants for all sites (OPTIMON study protocol 2008). The TEMPER study evaluated prespecified triggers for all sites in order to direct on-site visits to sites with a high trigger score (Stenning 2018b). A trigger data report based on database exports was generated and used in the trigger meeting to guide the prioritization of triggered sites. Triggers were 'fired' when an inequality rule that reflected a certain threshold of data non-conformities was evaluated as 'true'. Each trigger had an associated weight specifying its importance relative to other triggers, resulting in a trigger score for each site that was evaluated in trigger meetings and guided the prioritization of on-site visits (Diaz-Montana 2019a). In Knott 2015, all sites of the multicenter international trial received central statistical monitoring that identified high scoring sites as priority for further investigation. Scoring was applied every six months and a subsequent meeting of the central statistical monitoring group, including the chief investigator, chief statistician, junior statistician, and head of trial monitoring, and assessed high scoring sites and discussed trigger adjustments. Fired triggers resulted in a score of one and high scoring sites were chosen for a monitoring visit in the triggered intervention group.

Use of central monitoring and remote monitoring to support on-site monitoring

In the ADAMON study, central monitoring activities included statistical monitoring with multivariate analysis, structured telephone interviews, site status in terms of participant numbers (number of included participants, number lost to follow-up, screening failures, etc.) (Brosteanu 2017b). In the OPTIMON study, computerized controls were made on data entered from all participants in all investigation sites to check their completeness and consistency (Journot 2017). Following these controls,

the clinical research associate sent the investigator requests for clarification or correction of any inconsistent data. Regular contact was maintained by telephone, fax, or e-mail with the key people at the trial site to ensure that procedures were observed, and a report was compiled in the form of a standardized contact form.

Use of central monitoring without on-site monitoring

In the START Monitoring Substudy, central monitoring was performed by the statistical center using data in the central database on a continuous basis (Wyman 2020). Reports summarizing the reviewed data were provided to all sites and site investigators and were updated regularly (daily, weekly, or monthly). Sites and staff from the statistical center and coordinating centers also reviewed data summarizing each site's performance every six months and provided quantitative feedback to clinical sites on study performance. These reviews focused on participant retention, data quality, timeliness, and completeness of START Monitoring Substudy endpoint documentation, and adherence to local monitoring requirements. In addition, trained nurses at the statistical center reviewed specific adverse events and unscheduled hospitalizations for possible misclassification of primary START clinical events. Tertiary data, for example, laboratory values, were also reviewed by central monitoring (Hullsiek 2015).

Use of central monitoring for source data verification

In the Mealer 2013 pilot study, remote SDV validated the data elements captured on CRFs submitted to the coordinating center. Data collection instruments for capturing study variables were developed and remote access for the study monitor was set up to allow secure online access to electronic records. The same data verification protocols were used as during on-site visits and remote monitors had telephone access to local coordinators.

Initial risk assessment

An initial risk assessment of trials was performed in the ADAMON (Brosteanu 2017b) and OPTIMON (Journot 2017) studies. The RAS used in the OPTIMON study was evaluated in the validity and reproducibility study, the Pre-OPTIMON study, and was performed in three steps leading to four different risk categories that imply different monitoring plans. The first step related to the risk of the studied intervention in terms of product authorization, invasiveness of surgery technique, CE marking class, and

invasiveness of other interventions, which led to a temporary classification in the second step. In the third step, the risk of mortality based on the procedures of the intervention and the vulnerability of the study population were additionally taken into consideration and may have led to an increase in risk level. The risk analysis used in the ADAMON study also had three steps. The first step involved an assessment of the risk associated with the therapeutic intervention compared to the standard of care. The second step was based on the presence of at least one of a list of risk indicators for the participant or the trial results. In the third step, the robustness of trial procedures (reliable and easy to assess primary endpoint, simple trial procedures) was evaluated. The risk analysis resulted in one of three risk categories entailing different basic on-site monitoring measures in each of the three monitoring classes.

Excluded studies

We excluded 37 studies after full-text screening (Characteristics of excluded studies table). We excluded articles for the following reasons: 21 studies did not compare different monitoring strategies and 16 were not prospective studies.

Risk of bias in included studies

Risk of bias in the included studies is summarized in Figure 2 and Figure 3. We assessed all studies for risk of bias following the criteria described in the *Cochrane Handbook for Systematic Reviews of Interventions* for randomized trials (Higgins 2020). In addition, we used the ROBINS-I tool for the three non-randomized studies (Fougerou-Leurent 2019; Knott 2015; Stenning 2018b; results shown in Appendix 4).

Allocation

Selection bias

Group allocation was at random and concealed in four of the eight studies with low risk of selection bias (Brosteanu 2017b; Journot 2017; Lièvard 2006; Wyman 2020). Three were non-randomized studies; two evaluated triggered monitoring (matched comparator design), where randomization was not practicable due to the dynamic process of the monitoring intervention (Knott 2015; Stenning 2018b), and the other used a prospective cross-over design (the same CRFs were analyzed with full or targeted SDV) (Fougerou-Leurent 2019). Since we could not identify an increased

risk of bias for the prospective cross-over design (intervention applied on same participant data), we rated the study at low risk of selection bias. Although the original investigators attempted to balance groups and to control for confounding in the TEMPER study (Stenning 2018b), we rated the design at high risk of bias according to the criteria described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2020). One study randomly assigned participant-level data without any information about allocation concealment (unclear risk of bias) (Mealer 2013).

Blinding

Performance bias

In six studies, investigators, site staff, and data collectors of the trials were not informed about the monitoring strategy applied (Brosteanu 2017b; Journot 2017; Knott 2015; Lièvard 2006; Stenning 2018b; Wyman 2020). However, blinding of monitors was not practicable in these six studies and thus we judged them at high risk of bias. In two studies, blinding of site staff was difficult because the interventions of monitoring involved active participation of trial staff (high risk of bias) (Fougerou-Leurent 2019; Mealer 2013). It is unclear if the data management was blinded in these two studies.

Detection bias

Although monitoring could usually not be blinded due to the methodological and procedural differences in the interventions, three studies performed a blinded outcome assessment (low risk of bias). In ADAMON, the audit teams verifying the monitoring outcomes of the two monitoring interventions were not informed of the sites' monitoring strategy and did not have access to any monitoring reports (Brosteanu 2017b). Audit findings were reviewed in a blinded manner by members of the ADAMON team and discussed with auditors, as necessary, to ensure that reporting was consistent with the ADAMON audit manuals (ADAMON study protocol 2008). In OPTIMON, the main outcome was validated by a blinded validation committee (Journot 2017). In TEMPER, the lack of blinding of monitoring staff was mitigated by consistent training on the trials and monitoring methods, the use of a common finding grading system, and independent review of all major and critical findings which was blind to visit type (Stenning 2018b). The other five studies provided no information on blinded outcome assessment or blinding of statistical

center staff (unclear risk of bias) (Fougerou-Leurent 2019; Knott 2015; Lièvard 2006; Mealer 2013; Wyman 2020).

Follow up and exclusions

All eight included studies were at low risk of attrition bias (Brosteanu 2017b; Fougerou-Leurent 2019; Journot 2017; Knott 2015; Lièvard 2006; Mealer 2013; Stenning 2018b; Wyman 2020). However, ADAMON reported that "… one site refused the audit, and in the last five audited trials, 29 sites with less than three patients were not audited due to limited resources, in large sites (>45 patients), only a centrally preselected random sample of patients was audited. Arms are not fully balanced in numbers of patients audited (755 extensive on-site monitoring and 863 risk-adapted monitoring) overall" (Brosteanu 2017b). Another study was terminated prematurely due to slow participant recruitment, but the number of centers that randomized participants was equal in both groups (low risk of bias) (Lièvard 2006).

Selective reporting

A design publication was available for one study (START Monitoring Substudy [two publications] Hullsiek 2015; Wyman 2020) and three studies published a protocol (ADAMON: Brosteanu 2017b; OPTIMON: Journot 2017; TEMPER: Stenning 2018b). Three of these studies reported on all outcomes described in the protocol or design paper in their publications (Brosteanu 2017b; Stenning 2018b; Wyman 2020), and one study has not been published as a full report yet, but provided outcomes stated in the protocol in the available conference presentation (Journot 2017). One study has only been published as an abstract to date (Knott 2015), but the study authors communicated results of the prespecified outcomes to us. For the three remaining studies, there were no protocol or registry entries available but the outcomes listed in the methods sections of their publications were all reported in the results and discussion sections (MONITORING: Fougerou-Leurent 2019; Lièvard 2006; Mealer 2013).

Other potential sources of bias

There was an additional potential source of bias for one study (MONITORING: Fougerou-Leurent 2019). If the clinical research assistant spotted false or missing non-key data when checking key data, he or she may have corrected the non-key data in the CRF. This potential bias may have led to an underestimate of the

difference between the two monitoring strategies. The full SDV CRF was considered without errors.

Effects of methods

In order to summarize the results of the eight included studies, we grouped them according to their intervention comparisons and their outcomes.

Primary outcome

Combined outcome of critical and major monitoring findings

Five studies, three randomized (ADAMON: Brosteanu 2017b; OPTIMON: Journot 2017; START Monitoring Substudy: Wyman 2020), and two matched pair (TEMPER: Stenning 2018b; Knott 2015), reported a combined monitoring outcome with four to six underlying error domains (e.g. eligibility violations). The ADAMON and OPTIMON studies defined findings as protocol and GCP violations that were not corrected or identified by the randomized monitoring strategy. The START Monitoring Substudy directly compared findings identified by the randomized monitoring strategies without a subsequent evaluation of remaining findings not corrected by the monitoring intervention. The classification into different severities of findings comprised different categories in three included studies that had different denominations (non-conformity/major non-conformity [Journot 2017], minor/major/critical [Brosteanu 2017b; Stenning 2018b]), but were consistent in the assessment of severity with regard to participant's rights and safety or to validity of study results. Only findings classified as major or critical (or both) were included in the primary comparison of monitoring strategies in the ADAMON and OPTIMON studies. The START Monitoring Substudy only assessed major violations, which constitutes the highest severity of findings with regard to participant's rights and safety or to validity of study results. All three of these studies defined monitoring findings for the most critical aspects in the domains for consent violations, eligibility violations, SAE reporting violations, and errors in endpoint assessment. Since the START Monitoring Substudy focused on only one trial, these descriptions of critical aspects are very trial specific compared to the broader range of critical aspects considered in ADAMON and OPTIMON with a combined monitoring outcome. Critical and major findings are defined according to the classification of GCP findings described in EMA 2017. For detailed information about the classification of monitoring findings in the included studies, see the Additional tables.

1. Risk-based monitoring versus extensive on-site monitoring

ADAMON and OPTIMON evaluated the primary outcome as the remaining combined major and critical findings not corrected by the randomized monitoring strategy. Pooling the results of ADAMON and OPTIMON for the proportion of trial participants with at least one major or critical outcome not corrected by the monitoring intervention resulted in a risk ratio of 1.03 with a 95% CI of 0.80 to 1.33 (below 1.0 would be in favor of the risk-based strategy; Analysis 1.1; Figure 4). However, START Monitoring evaluated the primary outcome of combined major and critical findings as a direct comparison of monitoring findings during trial conduct and the comparison of monitoring strategies differed from the one assessed in ADAMON and OPTIMON. Therefore, we did not include START Monitoring in the pooled analysis, but reported its results separately below.

In the ADAMON study, 59.2% of participants with any major finding not corrected by the randomized monitoring strategy was identified in the risk-based monitoring intervention group compared to 64.2% of participants with any major finding in the 100% on-site group (Brosteanu 2017b). The analysis of the composite monitoring outcome in the ADAMON study using a random-effects model, estimated with logistic regression and with sites as random effects accounting for clustering, resulted in evidence of non-inferiority (point estimates near zero on the logit scale and all two-sided 95% CIs clearly excluding the prespecified tolerance limit) (Brosteanu 2017a).

The OPTIMON study reported the proportions of participants without major monitoring findings (Journot 2017). When considering the proportions of participants with major monitoring findings, 40% of participants in the risk-adapted monitoring intervention group had a monitoring outcome not identified by the randomized monitoring strategy compared to 34% in the 100% on-site group. Analysis of the composite primary outcome via the GEE logistic model resulted in an estimated relative difference between strategies of 8% in favor of the 100% on-site strategy. Since the upper one-sided confidence limit of this difference was 22%, non-inferiority with the set non-inferiority margin of 11% could not be demonstrated.

2. Central monitoring with triggered on-site visits versus regular (untriggered) on-site visits

Two studies used a matched comparator design (Knott 2015; Stenning 2018b). In these new strategies, on-site visits were triggered by the exceeding of prespecified

trigger thresholds. The studies reported the number of triggered sites that had monitoring findings versus the number of control sites that had a monitoring finding.

We pooled these two studies for the primary combined outcome of major and critical monitoring findings including all error domains (Analysis 3.1; Figure 5) and also after excluding re-consent for the TEMPER study (Analysis 4.1; Figure 6). Excluding the error domain "re-consent" gave a risk ratio of 2.04 (95% CI 0.77 to 5.38) in favor of the triggered monitoring while including re-consent findings gave a risk ratio of 1.83 (95% CI 0.51 to 6.55) in favor of the triggered monitoring intervention. These results provide some evidence that the trigger process was effective in guiding on-site monitoring but the differences were not statistically significant.

In the study conducted by Knott and colleagues, 21 sites (12 identified by central statistical monitoring, nine others as comparators) received an on-site visit and 11 of 12 identified by central statistical monitoring had one or more major or critical monitoring finding (92%), while only two of nine comparator sites (22%) had a monitoring finding (Knott 2015). Therefore, the difference in proportions of sites with at least one major or critical monitoring finding was 70%. Minor findings indicative of 'sloppy practice' were identified at 10 of 12 sites in the triggered group and in two of nine in the comparator group. At one site identified by central statistical monitoring, there were serious findings indicative of an underperforming site. These results suggest that information from central statistical monitoring can help focus the nature of on-site visits and any interventions required to improve site quality.

The TEMPER study identified 37 of 42 (88.1%) triggered sites with one or more major or critical finding not already identified through central monitoring or a previous visit and 34 of 42 (81.0%) matched untriggered sites with one of more major or critical finding (difference 7.1%, 95% CI –8.3% to 22.5%; P = 0.365) (Stenning 2018b). More than 70% of on-site findings related to issues in recording informed consent, and 70% of these to re-consent. The prespecified sensitivity analysis excluding re-consent findings demonstrated a clear difference in event rate. When excluding re-consent findings, the numbers reduced to 85.7% for triggered sites and 59.5% for untriggered sites (difference 26.2%, 95% CI 8.0% to 44.4%; P = 0.007). Thus, triggered monitoring in the TEMPER study did not satisfactorily distinguish sites with higher and lower levels of concerning on-site monitoring findings. However, the prespecified sensitivity analysis excluding re-consent findings demonstrated a clear difference in event rate. There was greater consistency between trials in the

sensitivity and secondary analyses. In addition, there was some evidence that the trigger process used could identify sites at increased risk of serious concern: around twice as many triggered visits had one or more critical finding in the primary and sensitivity analyses.

3. Central and local monitoring with annual on-site visits versus central and local monitoring only

The START Monitoring study (Wyman 2020), with 196 sites in a single large international trial, reported a higher proportion of participants with a monitoring finding detected in the on-site monitoring group (6.4%) compared to the group with only central and local monitoring (3.8%), resulting in an odds ratio (OR) of 1.7 (95% CI 1.1 to 2.7; P = 0.03) (Wyman Engen 2020). However, it is not clearly reported if the findings within the groups were identified on-site (on-site visit or local monitoring) or by central monitoring and it was not verified whether central monitoring and local monitoring alone were unable to detect any violations or discrepancies within sites randomized to the intervention group. In addition, relatively few monitoring findings that would have impacted START results were identified by on-site monitoring (no findings of participants who were inadequately consented, no findings of data alteration or fraud).

4. Traditional 100% source data verification versus remote or targeted source data verification

The two studies of targeted (MONITORING: Fougerou-Leurent 2019) and remote (Mealer 2013) SDV reported findings only related to source documents. Different components of source data were assessed including consent verification as well as key data, but findings were reported only as a combined outcome. Minimal relative differences of parameters assessing the effectiveness of these methods in comparison to full SDV were identified in both studies. Both studies only assessed the SDV as the process of double-checking that the same piece of information was written in the study database as well as in source documents. Processes, often referred to as Source Data Review, that confirm that the trial conduct complies with the protocol and GCP and ensure that appropriate regulatory requirements have been followed, are not included as study outcomes.

In the prospective cross-over MONITORING study, comparing the databases of full SDV and target SDV, after the data management process identified an overall error

rate of 1.47% (95% CI 1.41% to 1.53%) and an error rate of 0.78% (95% CI 0.65% to 0.91%) on key data (Fougerou-Leurent 2019). The majority of these discrepancies, considered as the remaining errors with targeted monitoring, were observed on baseline prognostic variables. The researchers further assessed the impact of the two different monitoring strategies on data-management workload. While the overall number of queries was larger with the targeted SDV, there was no statistical difference for the queries related to key data (13 [standard deviation (SD) 16] versus 5 [SD 6]; P = 0.15) and targeted SDV generated fewer corrections on key data in the data-management process step. Considering the increased workload for data management at least in the early setup phase of a targeted SDV strategy, monitoring and data management should potentially be viewed as a whole in terms of efficacy.

The pilot study conducted by Mealer and colleagues assessed the feasibility of remote SDV in two clinical trial networks (Mealer 2013). The accuracy and completeness of remote versus on-site SDV was determined by analyzing the number of data values that were either identical or different in the source data, missing or unknown after remote SDV reconciliated to all data values identified via subsequent on-site monitoring. The percentage of data values that could either not be identified or were missed via remote access were compared to direct on-site monitoring in another group of participants. In the adult network, only 0.47% (95% CI 0.03% to 0.79%) of all data values assigned to monitoring could not be correctly identified via remote monitoring and in the ChiLDReN network, all data values were correctly identified. In comparison, three data values could not be identified in the only on-site group (0.13%, 95% CI 0.03% to 0.37%). In summary, 99.5% of all data values were correctly identified via remote monitoring. Information on the difference in monitoring findings during the two SDV methods was not reported in the publication. The study showed that remote SDV was feasible despite marked differences in remote access and remote chart review policies and technologies.

5. On-site initiation visit versus no on-site initiation visit

There were no data on critical and major findings in Liènard 2006.

Secondary outcomes

Individual components of the primary outcome

Individual components of the primary outcome considered in the included studies were:

- Major eligibility violations;
- Major informed-consent violations;
- Findings that raised doubt about the accuracy or credibility of key trial data and deviations of intervention from the trial protocol (with impact on patient safety or data validity);
- Errors in endpoint assessment; and
- Errors in SAE reporting.

1. Risk-based versus extensive on-site monitoring

In the ADAMON study, there was non-inferiority for all of the five error domain components of the combined primary outcome: informed consent process, patient eligibility, intervention, endpoint assessment, and SAE reporting (Brosteanu 2017a). In the OPTIMON study, the biggest difference between monitoring strategies was observed for findings related to eligibility violations (12% of participants with major non-conformity in eligibility error domain in the risk-adapted group versus 6% of participants in the extensive on-site group), while remaining findings related to informed consent were higher in the extensive on-site monitoring group (7% of participants with major non-conformity in informed consent error domain in the risk-adapted group versus 10% of participants in the extensive on-site group). In the OPTIMON study, consent form signature was checked remotely using a modified consent form and a validated specific procedure in the risk-adapted strategy (Journot 2013). To summarize the domain specific monitoring outcomes of the ADAMON and OPTIMON studies, we analyzed the results of both studies within the four common error domains (Analysis 2.1, including unpublished results from OPTIMON). Pooling the results of the four common error domains (informed consent process, patient eligibility, endpoint assessment, and SAE reporting) resulted in a risk ratio of 0.95 (95% CI 0.81 to 1.13) in favor of the risk-based monitoring intervention (Figure 7).

2. Central monitoring with triggered on-site visits versus regular (untriggered) on-site visits

In TEMPER, informed consent violations were more frequently identified by a full on-site monitoring strategy (Stenning 2018b). During the study, but prior to the first analysis, the TEMPER Endpoint Review Committee recommended a sensitivity analysis to exclude all findings related to re-consent, because these typically communicated minor changes in the adverse effect profile that could have been communicated without requiring re-consent. Excluding re-consent findings to evaluate the ability of the applied triggers to identify sites at higher risk for critical on-site findings resulted in a significant difference of 26.2% (95% CI 8.0% to 44.4%; P = 0.007). Excluding all consent findings also resulted in a significant difference of 23.8% (95% CI 3.3% to 44.4%; P = 0.027).

There were no data on individual components of critical and major findings in Knott 2015.

3. Central and local monitoring with annual on-site visits versus central and local monitoring only

In the START Monitoring Substudy, informed consent violations accounted for most of the primary monitoring outcomes in each group (41 [1.8%] participants in the no on-site group versus 56 [2.7%] participants in the on-site group) with an OR of 1.3 (95% CI 0.6 to 2.7; P = 0.46) (Wyman 2020). The most common consent violation was the most recently signed consent signature page being missing and that the surveillances for these consent violations by on-site monitors varied. Within the START Monitoring Substudy, they had to modify the primary outcome component for consent violations prior to the outcomes assessment in February 2016 because documentation and ascertainment of consent violations were not consistent across sites. This suggests that these inconsistencies and variation between sites could have influenced the results of this primary outcome component. In addition, the follow-up on consent violations by the coordinating centers identified no individuals who had not been properly consented. The largest relative difference was for the findings related to eligibility (1 [0.04%] participant in the no on-site group versus 12 [0.6%] participants in the on-site group; OR 12.2, 95% CI 1.8 to 85.2; P = 0.01), but 38% of eligibility violations were first identified by site staff. In addition, a relative difference was reported for SAE reporting (OR 2.0, 95% CI 1.1 to 3.7; P = 0.02), while the differences for the error domains primary endpoint reporting (OR 1.5, 95% CI 0.7 to 3.0; P = 0.27) and protocol violation of prescribing initial antiretroviral

therapy not permitted by START (OR 1.4, 95% CI 0.6 to 3.4; P = 0.47) as well as for the informed consent domain were small.

4. Traditional 100% source data verification versus remote or targeted source data verification

There were no data on individual components of critical and major findings in MONITORING (Fougerou-Leurent 2019) or Mealer 2013.

5. Systematic on-site initiation visit versus on-site initiation visit upon request

There were no data on individual components of critical and major findings in Lièvard 2006.

Impact of the monitoring strategy on participant recruitment and follow-up

Only two included studies reported participant recruitment and follow-up as an outcome for the evaluation of different monitoring strategies (Lièvard 2006; START Monitoring Substudy: Wyman 2020).

Lièvard 2006 assessed the impact of their monitoring approaches on participant recruitment and follow-up in their primary outcomes. Centers were randomized to receive an on-site initiation visit by monitors or no visit. There was no statistical difference in the number of recruited participants between these two groups (302 participants in the on-site group versus 271 participants in the no on-site group) as well as no impact of monitoring visits on recruitment categories (poor, average, good, and excellent). About 80% of participants were recruited in only 30 of 135 centers, and almost 62% in the 17 'excellent recruiters'. The duration of follow-up at the time of analysis did not differ significantly between the randomized groups. However, the proportion of participants with no follow-up at all was larger in the visited group than in the non-visited group (82% in the on-site group versus 70% in the no on-site group).

Within the START Monitoring Substudy, central-monitoring reports included tracking of losses to follow-up (Wyman 2020). Losses to follow-up were similar between groups (proportion of participants lost to follow-up: 7.1% in the on-site group versus 8.6% in the no on-site group; OR 0.8, 95% CI 0.5 to 1.1), and a similar percentage of study visits were missed by participants in each monitoring group (8.6% in the on-site group versus 7.8% in the no on-site group).

Effect of monitoring strategies on resource use (costs)

Five studies provided data on resource use.

1. Risk-based versus extensive on-site monitoring

The ADAMON study reported that with extensive on-site monitoring, the number of monitoring visits per participant and the cumulative monitoring time on-site was higher compared to risk-adapted monitoring by a factor of 2.1 (monitoring visits) and 2.7 (cumulative monitoring time) (ratios of the efforts calculated within each trial and summarized with the geometric mean) (Brosteanu 2017b). This difference was more pronounced for the lowest risk category, resulting in an increase of monitoring visits per participant by a factor of 3.5 and an increase in the cumulative monitoring time on-site by a factor of 5.2. In the medium-risk category, the number of monitoring visits per participant was higher by a factor of 1.8 and the cumulative monitoring time on-site was higher by a factor of 2.1 for the extensive on-site group compared to the risk-based monitoring group.

In the OPTIMON study, travel costs were calculated depending on the distance and on-site visits were assumed to require two days for one monitor, resulting in monitoring costs of EUR 180 per visit (Journot 2017). The costs were higher by a factor of 2.7 for the 100% on-site strategy when considering travel costs only, and by a factor of 3.4 when considering travel and monitor costs.

2. Central monitoring with triggered on-site visits versus regular (untriggered) on-site visits

There were no data on resource use from TEMPER (Stenning 2018b) or Knott 2015.

3. Central and local monitoring with annual on-site visits versus central and local monitoring only

In the START Monitoring Substudy, the economic consequence of adding on-site monitoring to local and central monitoring was assessed by the person-hours that on-site monitors and coordinating centers spent performing on-site monitoring-related activities and was estimated to be 16,599 person-hours (Wyman 2020). With a salary allocation of USD 75 per hour for on-site monitors, this equated to USD 1,244,925. With the addition of USD 790,467 international travel costs that were allocated for START monitoring, a total of USD 2,035,392 was attributed to on-site monitoring. It

has to be considered that there were four additional visits for cause in the on-site group and six visits for cause in the no on-site group.

4. Traditional 100% source data verification versus remote or targeted source data verification

For the MONITORING study, economic data were assessed in terms of time spent on SDV and data management with each strategy (Fougerou-Leurent 2019). A query was estimated to take 20 minutes to handle for a data manager and 10 minutes for the clinical study coordinator. Across the six studies, 140 hours were devoted by the clinical research associate to the targeted SDV versus 317 hours for the full SDV. However, targeted SDV generated 587 additional queries across studies, with a range of less than one (0.3) to more than eight additional queries per participant, depending on the study. In terms of time spent on these queries, based on an estimate of 30 minutes for handling a single query, the targeted SDV-related additional queries resulted in 294 hours of extra time spent (mean 2.4 [SD 1.7] hours per participant).

For the cost analysis, the hourly costs for a clinical research associate were estimated to be EUR 33.00, a data-manager was EUR 30.50, and a clinical study coordinator was EUR 30.50. Based on these estimates, the targeted SDV strategy provided a EUR 5841 saving on monitoring but an additional EUR 8922 linked to the queries, totaling an extra cost of EUR 3081.

The study on remote SDV by Mealer 2013 only compared time consumed per data item and time per case report form for both included networks. Although there was no relevant difference (less than 30 seconds) per data item between the two strategies, more time was spent with remote SDV. However, this study did not consider travel time for monitors, and the delayed access and increased response time for the communication with study coordinators affected the overall time spent. The authors proposed SOPs for prescheduling times to review questions by telephone and the introduction of a single electronic health record.

For both of the introduced SDV monitoring strategies, a gain of experience with these new methods would most likely translate into improved efficiency, making it difficult to estimate the long-term resource use from these initial studies. For the risk-based strategy in the OPTIMON study, a remote pre-enrollment check of consent forms was a good preventive measure and improved quality of consent forms (80% of non-

conformities identified via remote checking). In general, remote SDV monitoring may reduce the frequency of on-site visits or influence their timing ultimately decreasing the resources needed for on-site monitoring.

5. Systematic on-site initiation visit versus on-site initiation visit upon request

There were no data on resource use from Lièvard 2006.

Qualitative research data or process evaluations of the monitoring interventions

The Mealer 2013 pilot study of traditional 100% SDV versus remote SDV provided some qualitative information. This came from an informal post-study interview of the study monitors and site coordinators. These interviews revealed a high level of satisfaction with the remote monitoring process. None of the study monitors reported any difficulty with using the different electronic access methods and data review applications.

The secondary analyses of the TEMPER study assessed the ability of individual triggers and site characteristics to predict on-site findings by comparing the proportion of visits with the outcome of interest (one major/critical finding) for triggered on-site visits with regular (untriggered) on-site visits (Stenning 2018b). This analysis also considered information of potential prognostic value obtained from questionnaires completed by the trials unit and site staff prior to the monitoring visits. Trials unit teams completed 90/94 pre-visit questionnaires. There was no clear evidence of a linear relationship between the trial team ratings and the presence of major or critical findings, including or excluding consent findings (data not shown). A total of 76/94 sites provided pre-visit site questionnaires. There was no evidence of a linear association between the chance of one major/critical finding and the number of active trials either per site or per staff member (data not shown). There was, however, evidence that the greater the number of different trial roles undertaken by the research nurse, the lower the probability of major/critical findings (number of research nurse roles (grouped) – proportion of one or more major or critical finding within the group, excluding re-consent findings: less than 3: 94%; 4: 94%; 5: 80%; 6: 48% ($P < 0.001$; from Chi^2 test for linear trend) (Stenning 2018b, Online Supplementary Material Table S5).

Discussion

Summary of main results

We identified eight studies that prospectively compared different monitoring interventions in clinical trials. These studies were heterogeneous in design and content, and covered different aspects of new monitoring approaches. We identified no ongoing eligible studies.

Two large studies compared risk-based versus extensive on-site monitoring (ADAMON: Brosteanu 2017b; OPTIMON: Journot 2017), and the pooled results provided no evidence of inferiority of a risk-based monitoring intervention in terms of major and critical findings, based on moderate certainty of evidence (Summary of findings table 1). However, a formal demonstration of non-inferiority would require more studies.

Considering the commonly reported error domains of monitoring findings (informed consent, eligibility, endpoint assessment, SAE reporting), we found no evidence for inferiority of a risk-based monitoring approach in any of the error domains except eligibility. However, CIs were wide. To verify the eligibility of a participant usually requires extensive SDV, which might explain the potential difference in this error domain. We found a similar trend in the START Monitoring Substudy for the eligibility error domain. Expanding processes for remote SDV may improve the performance of monitoring strategies with a larger proportion of central and remote monitoring components. The OPTIMON study used an established process to remotely verify the informed consent process (Journot 2013), which was shown to be efficient in reducing non-conformities related to informed consent. A similar remote approach for SDV related to eligibility before randomization might improve the performance of risk-based monitoring interventions in this domain.

In the TEMPER study (Stenning 2018b) and the START Monitoring Substudy (Wyman 2020), most findings related to documenting the consent process. However, in the START Monitoring Substudy, there were no findings of participants whose consent process was inadequate and, in the ADAMON and the OPTIMON studies, findings in the informed consent process were lower in the risk-adapted groups. Timely central monitoring of consent forms and eligibility documents with adequate anonymization (Journot 2013) may mitigate the effects of many consent form completion errors and identify eligibility violations prior to randomization. This is also

supported by the recently published further analysis of the TEMPER study (Cragg 2021a), which suggested that most visit findings (98%) were theoretically detectable or preventable through feasible, centralized processes, especially all the findings relating to initial informed consent forms, thereby preventing patients starting treatment if there are any issues. Mealer 2013 assessed a remote process for SDV and found it to be feasible. Data values were reviewed to confirm eligibility and proper informed consent, to validate that all adverse events were reported, and to verify data values for primary and secondary outcomes. Almost all (99.6%) data values were correctly identified via remote monitoring at five different trial sites despite marked differences in remote access and remote chart review policies and technologies. In the MONITORING study, the number of remaining errors after targeted SDV (verified by full SDV) was very small for the overall data and even smaller for key data items (Fougerou-Leurent 2019). These results provide evidence that new concepts in the process of SDV do not necessarily lead to a decrease in data quality or endanger patient rights and safety. Processes involved with on-site SDV and often referred to as source data review, that confirm that the trial conduct complies with the protocol and GCP and ensure that appropriate regulatory requirements have been followed, have to be assessed separately. Evidence from retrospective studies evaluating SDV suggest that intensive SDV is often of little benefit to clinical trials, with any discrepancies found having minimal impact on the robustness of trial conclusions (Andersen 2015; Olsen 2016; Tantsyura 2015; Tudur Smith 2012a).

Furthermore, we found evidence that central monitoring can guide on-site monitoring of trial sites via triggers. The prespecified sensitivity analysis of the TEMPER results excluding re-consent findings (Stenning 2018b) and the results from Knott 2015 suggested that using triggers from a central monitoring process can identify sites at higher risk for major GCP violations. However, the triggers used in TEMPER may not have been ideal for all included trials and some tested triggers seemed not to have any prognostic value. Additional work is needed to identify more discriminatory triggers and should encompass work on key performance indicators (Gough 2016) and central statistical monitoring (Venet 2012). Since Knott 2015 focused on one study only, the triggers used in TEMPER were more trial specific. Developing trial specific triggers may lead to even more efficient triggers for on-site monitoring. This may help to distinguish low performing sites from high performing sites and guide monitors to the most urgent problems within the identified site. Study-specific triggers could even provoke specific monitoring activities (e.g. staff turnover indicates

additional training, or data quality issues could trigger SDV activities). Central review of information across sites and time would help direct the on-site resources to targeted SDV and activities best performed in-person, for example, process review or training. We found no evidence that the addition of untriggered on-site monitoring to central statistical monitoring assessed in the START Monitoring Substudy had a major impact on trial results or on participants' rights and safety (Wyman 2020). In addition, there was no evidence that the no on-site group was inferior in the study-specific secondary outcomes including the percentage of participants lost to follow-up, timely data submission and query resolution, and the absolute number of monitoring outcomes in the START Monitoring Substudy was very low (Wyman 2020). This might be due to a study-specific definition of critical and major findings in the monitoring plan and the presence of an established central monitoring system in both intervention groups of the study.

With respect to resource use, both studies evaluating a risk-based monitoring approach showed that considerable resources could be saved with risk-based monitoring (factor three to five; Brosteanu 2017b; Journot 2017). However, the potential increase in resource use at the coordinating centers (including data management) was not considered in any of the analyses. The START Monitoring Substudy reported more than USD 2,000,000 for on-site monitoring, taking into account the monitoring hours as well as the international travel costs (Wyman 2020). In both groups, central and local monitoring by site staff were performed to an equal extent, suggesting that there is no difference in the resources consumed by data management. The MONITORING study reported a reduction in cost of on-site monitoring by the targeted SDV approach, but this was offset by an increase in data management resources due to queries (Fougerou-Leurent 2019). This increase in data management resources may to some degree be due to the inexperience with the new approach of site staff and trial monitors. There was no statistical difference in number of queries related to key data between targeted SDV and full SDV. When an infrastructure for centralized monitoring and remote data checks is already established, a larger difference between resources spent on risk-based compared to extensive on-site monitoring would be expected. Setting up the infrastructure for automated checks, remote processes, and other data management structures as well as the training of monitors and data managers on a new monitoring strategy requires an upfront investment.

Only two studies assessed the impact of different monitoring strategies on recruitment and follow-up. This is an important outcome for monitoring interventions because it is crucial for the successful completion of a clinical trial (Houghton 2020). The START Monitoring study found no significant difference in the percentage of participants lost to follow-up between the on-site and no on-site groups (Wyman 2020). Also, on-site initiation visits had no effect on participant recruitment in Lièvard 2006. Closely monitoring site performance in terms of recruitment and losses to follow-up could enable early action to support affected sites. Secondary qualitative analyses of the TEMPER study revealed that the experience of the research nurse had an impact on the monitoring outcomes (Stenning 2018b). The experience of the study team and the site staff might also be an important factor to be considered in a risk assessment of the study or in the prioritization of on-site visits.

Overall completeness and applicability of evidence

Although we extensively searched for eligible studies, we only found one or two studies for specific comparisons of monitoring strategies. This very limited evidence base stands in stark contrast to the number of clinical trials run each year, each of which needs to perform monitoring in some form. None of the included studies reported on all primary and secondary outcomes specified for this review and most studies reported only a few. For instance, only one study reported on participant recruitment (Lièvard 2006), and only two studies reported on participant retention (Lièvard 2006; Wyman 2020). Some monitoring comparisons were nested in a single clinical trial limiting the generalizability of results (e.g. Knott 2015; START Monitoring: Wyman 2020). However, the OPTIMON (Journot 2017) and ADAMON (Brosteanu 2017b) studies included multiple and heterogeneous clinical trials for their comparison of risk-based and extensive on-site monitoring strategies increasing the generalizability of their results. The risk assessments of the ADAMON and OPTIMON studies differed in certain aspects (Table 2), but the main concept of categorizing studies according to their evaluated risk and adapting the monitoring requirements depending on the risk category was very similar. The much lower number of overall monitoring findings in the START study (based on one clinical trial only) compared with OPTIMON or ADAMON (involving multiple clinical trials) suggests that the trial context is crucial with respect to monitoring findings. Violations considered in the primary outcome of the START Monitoring Substudy were tailored to issues that could impact the validity of the trial's results or the safety of study participants. A definition of assets focused on the most critical aspects of a study that

should be monitored closely is often missing in extensive monitoring plans and allows for some margin of interpretation by study monitors.

The TEMPER study introduced triggers that could direct on-site monitoring and evaluated the prognostic values of these triggers (Stenning 2018b). Only three of the proposed triggers showed a significant prognostic impact across all three included trials. A set of triggers or performance measures of trial sites that are promising indicators for the need of additional support across a wide range of clinical trials are yet to be determined and trigger refinement is still ongoing. Triggers will to some degree always depend on the specific risks determined by the study procedures, management structure, and design of the study at hand. A combination of performance metrics appropriate for a large group of trials and study-specific performance measures might be most effective. Multinational, multicenter trials might benefit the most from the directing of on-site monitoring to sites that show low quality of performance. More studies in trials with large numbers of participants and sites, and trials covering diverse geographic areas, are needed to assess the value of centralized monitoring to assist with the identification of sites where additional support in terms of training is needed the most. This would lead to a more 'needs-oriented' approach, so that clinical routine and study processes in well-performing sites will not be unnecessarily interrupted. An overview of the progress of the ongoing trial in terms of site performance and other aspects such as recruitment and retention would also support the whole complex management processes of trial conduct in these large trials.

Since this review focused on prospective comparisons of monitoring interventions, the evidence from retrospective studies and reports from implementation studies is not included in the above results but is discussed below. We excluded retrospective studies because standardization of extracted data is not possible since data were collected before considering the analysis, especially for our primary outcome. However, trending analyses provide valuable information on outcomes such as improved data quality, recruitment, and follow-up compliance, and thus demonstrate the effect of monitoring approaches on the overall trial conduct and success of the study. We considered the results from retrospective studies in our discussion of monitoring strategies but also pointed out the need to establish more SWAT to prospectively compare methods with a predefined mode of analysis.

Quality of the evidence

Overall, the certainty of this body of evidence on monitoring strategies for clinical intervention studies was low or very low for most comparisons and outcomes (Summary of findings table 1; Summary of findings table 2; Summary of findings table 3; Summary of findings table 4; Summary of findings table 5). This was mainly due to imprecision of effect estimates because of small numbers of observations and indirectness because some comparisons were based on only one study nested in a single trial. The included studies varied considerably in terms of the reported outcomes with most studies reporting only some. In addition, the risk of bias varied across studies. A risk of performance bias was attributed to six of the included studies and was unclear in two studies. Since it was difficult to blind monitors to the different monitoring interventions, an influence of the monitors' performance on the monitoring outcomes could not be excluded in these studies. Two studies were at high risk of bias because of their non-randomized design (Knott 2015; TEMPER: Stenning 2018b). However, since the intervention determined the selection of sites for an on-site visit in the triggered groups, a randomized design was not practicable. In addition, the TEMPER study attempted to balance groups by design and controlled the risk of known confounding factors by using a matching algorithm. Therefore, the judgment of high risk of bias for TEMPER (Stenning 2018b) and Knott 2015 remains debatable. In the START Monitoring Substudy, no independent validation of remaining findings was performed after monitoring intervention. Therefore, it is uncertain if central monitoring without on-site monitoring missed any major GCP violations and chance findings cannot be ruled out. More evidence is needed to evaluate the value of on-site initiation visits. Lièvard 2006 found no evidence that on-site initiation visits affected participant recruitment, or data quality in terms of timeliness of data transfer and data queries. However, the informative value of the study was limited by its early termination and the small number of ongoing monitoring visits. In general, embedding methodology studies in clinical intervention trials provides valuable information for the improvement and adaptation of methodology guidelines and the practice of trials (Bensaaud 2020; Treweek 2018a; Treweek 2018b). Whenever randomization is not practicable in a methodology substudy, the attempt to follow a 'diagnostic study design' and minimize confounding factors as much as possible can increase the generalizability and impact of the study results.

Potential biases in the review process

We screened all potentially relevant abstracts and full-text articles independently and in duplicate, assessed the risk of bias for included studies independently and in duplicate, and extracted information from included studies independently and in duplicate. We did not calculate any agreement statistics, but all disagreements were resolved by discussion. We successfully contacted authors from all included studies for additional information. Since we were unable to extract only the outcomes of the randomized trials included in the OPTIMON study (Journot 2015), we used the available data that included mainly randomized trials but also a few cohort and cross-sectional studies. The focus of this review was on monitoring strategies for clinical intervention studies and including all studies from the OPTIMON study might introduce some bias. With regard to the pooling of study results, our judgment of heterogeneity might be debatable. The process of choosing comparator sites for triggered sites differed between the TEMPER study (Stenning 2018b) and Knott 2015. While both studies selected high scoring sites for triggered monitoring and low scoring sites as control, the TEMPER study applied a matching algorithm to identify sites that resembled the high scoring sites in certain parameters. In Knott 2015, comparator sites from the same countries were identified by the country teams as potentially problematic among the low scoring sites without a pairwise matching to a high scoring site. However, the principle of choosing sites for evaluation based on results from central statistical monitoring closely resembled methods used in the TEMPER study. Therefore, we decided to pool results from TEMPER and Knott 2015.

Agreements and disagreements with other studies or reviews

Although there are no definitive conclusions from available research comparing the effectiveness of risk-based monitoring tools, the OECD advises clinical researchers to use risk-based monitoring tools (OECD 2013). They emphasized that risk-based monitoring should become a more reactive process where the risk profile and performance is continuously reviewed during trial conduct and monitoring practices are modified accordingly. One systematic review on risk-based monitoring tools for clinical trials by Hurley and colleagues summarized a variety of new risk-based monitoring tools for clinical trial monitoring that had been implemented in recent years by grouping common ideas (Hurley 2016). They did not identify a standardized approach for the risk assessment process for a clinical trial in the 24 included risk-based monitoring tools, although the process developed by TransCelerate BioPharma Inc. has been replicated by six other risk-based monitoring tools

(TransCelerate BioPharma Inc 2014). Hurley and colleagues suggested that the responsiveness of the tool depends on their mode of administration (paper-based, powered by Microsoft Excel, or operated as a Service as a system) and the degree of centralized monitoring involved (Hurley 2016). An electronic data capture system is beneficial to the efficient performance of centralized monitoring. However, to support the reactive process of risk-based monitoring, tools should be able to incorporate information on risks provided by on-site experiences from the study monitors. This is in agreement with our findings that a risk-based monitoring tool should support both on-site and centralized monitoring and that assessments are continuously reviewed during study conduct. Monitoring is most efficient when integrated as part of a risk-based quality management system as also discussed by Buyse et al. (Buyse 2020), where a focus on trial aspects that have a potentially high impact on patient safety and trial validity and on systematic errors is emphasized.

From the five main comparisons that we identified through our review, four have also been assessed in available retrospective studies.

Risk-based versus extensive on-site monitoring: Kim and colleagues retrospectively reviewed three multicenter, investigator-initiated trials that were monitored by a modified ADAMON method consisting of on-site and central monitoring according to the risk of the trial (Kim 2021). Central monitoring was more effective than on-site monitoring in revealing minor errors and showed comparable results in revealing major issues such as investigational product compliance and delayed reporting of SAEs. The risk assessment assessed by Higa and colleagues was based on the Risk Assessment Categorization Tool (RACT) originally developed by TransCelerate BioPharma Inc. (TransCelerate BioPharma Inc 2014), and was continuously adopted during the study based on results of centralized monitoring in parallel with site (on-site/off-site) monitoring. Mean on-site monitoring frequency decreased as the study progressed and a Pharmaceutical and Medical Devices Agency inspection after study end found no significant non-conformance that would have affected the study results and patient safety (Higa 2020).

Central monitoring with triggered on-site visits versus regular on-site visits: several studies have assessed triggered monitoring approaches that depend on individual study risks in trending analysis of their effectiveness. Diani and colleagues evaluated the effectiveness of their risk-based monitoring approach in clinical trials involving implantable cardiac medical devices (Diani 2017). Their strategy included a

data-driven risk assessment methodology to target on-site monitoring visits and they found significant improvement in data quality related to the three risk factors that were most critical to the overall compliance of cardiac rhythm management along with an improvement in a majority of measurable risk factors at the worst performing site quantiles. The methodology evaluated by Agrafiotis and colleagues is centered on quality by design, central monitoring, and triggered, adaptive on-site and remote monitoring. The approach is based on a set of risk indicators that are selected and configured during the setup of each trial and are derived from various operational and clinical metrics. Scores from these indicators form the basis of an automated, data-driven recommendation on whether to prioritize, increase, decrease, or maintain the level of monitoring intervention at each site. They assessed the trending impact of their new approach by retrospectively analyzing the change in risk level later in the trials. All 12 included trials showed a positive effect in risk level change and results were statistically significant in eight of them (Agrafiotis 2018). The evaluation of a new trial management method for monitoring and managing data return rates in a multicenter phase III trial performed by Cragg and colleagues adds to the findings of increased efficiency by prioritizing sites for support (Cragg 2019). Using an automated database report to summarize the data return rate, overall and per center, enabled the early notification of centers whose data return rate appeared to be falling, or crossed the predefined acceptability threshold of data return rate. Concentrating on the gradual improvement of centers having persistent data return problems, resulted in an increase in the overall data return rate and return rates above 80% in all centers. These results agree with the evidence we found for the effectiveness of a triggered monitoring approach evaluated in TEMPER (Stenning 2018b) and Knott 2015, and emphasize the need for study-specific performance indicators. In addition, the data-driven risk assessment implemented by Diani 2017 highlighted key focus areas for both on-site and centralized monitoring efforts and enabled an emphasis of site performance improvements where it is needed the most. Our findings agree with retrospective assessments that focusing on the most critical aspects of a trial and guiding monitoring resources to trial sites in need of support may be efficient to improve the overall trial conduct.

Central statistical versus on-site monitoring: one retrospective analysis of the potential of central monitoring to completely replace on-site monitoring performed by trial monitors showed that the majority of reviewed on-site findings could be identified using central monitoring strategies (Bakobaki 2012). One recent scoping review focused on methods used to identify sites of 'concern', at which monitoring activity

may be targeted, and consequently sites 'not of concern', monitoring of which may be reduced or omitted (Cragg 2021b). It included all original reports describing methods for using centrally held data to assess site-level risk described in a reproducible way. Thus, in agreement with our research, they only identified one full report of a study (Stenning 2018b) that prospectively assessed the methods' ability to target on-site monitoring visits to most problematic sites. However, through contacting the authors of Knott 2015, which is only available as an abstract, we gained more detailed information on the methodology of the study and were able to include the results in our review. In contrast to our review, Cragg 2021b included retrospective assessments (in comparison to on-site monitoring, effect on data quality or other trial parameters) as well as case studies, illustrations of methods on data, assessment of methods' ability to identify simulated problem sites, or known problems in real trial data. Thus, it constitutes an overview of methods introduced to the research community, and simultaneously underlines the lack of evidence for their efficacy or effectiveness.

Traditional 100% SDV versus targeted or remote SDV: in addition to these retrospective evaluations of methods to prioritize sites and the increased use of centralized monitoring methods, several studies retrospectively assessed the value and effectiveness of remote monitoring methods including alternative SDV methods. Our findings related to a reduction of 100% on-site SDV in Mealer 2013 and the MONITORING study (Fougerou-Leurent 2019) are in agreement with Tudur Smith 2012b, which assessed the value of 100% SDV in a cancer clinical trial. In their retrospective comparison of data discrepancies and comparative treatment effects obtained following 100% SDV to those based on data without SDV, the identified discrepancies for the primary outcome did not differ systematically across treatment groups or across sites and had little impact on trial results. They also suggested that a focus of SDV on less-experienced sites or sites with differing reporting characteristics of SDV-related information (e.g. SAE reporting compared to other sites), with provision of regular training may be more efficient. Similarly, the study by Anderson and colleagues analyzed error rates of data from three randomized phase III trials monitored with a combination of complete SDV or partial SDV that were subjected to post hoc complete SDV (Andersen 2015). Comparing partly and fully monitored trial participants; there were only minor differences between variables of major importance to efficacy or safety. In agreement with these studies, the study by Embleton-Thirsk and colleagues showed that the impact of extensive retrospective SDV and further extensive quality checks in a phase III academic-led, international,

randomized cancer trial was minimal (Embleton-Thirsk 2019). Besides the potential reduction in SDV, remote monitoring systems for full or partial SDV are becoming more relevant during the COVID-19 pandemic and are currently evaluated in various forms. Another recently published study assessed the clinical trial monitoring effectiveness of remote risk-based monitoring versus on-site monitoring with 100% SDV (Yamada 2021). It used a cloud-based remote monitoring system that does not require site-specific infrastructure for remote monitoring since it can be downloaded onto mobile devices as an application and involves the upload of photographs. Remote monitoring was focused on risk items that could lead to critical data and process errors, determined using the risk assessment and categorization tool developed by TransCelerate BioPharma Inc. (TransCelerate BioPharma Inc 2014). Using this approach, 92.9% (95% CI 68.5% to 98.7%) of critical process errors could be detected by remote risk-based monitoring. With a retrospective review of monitoring reports, Hirase and colleagues supported an increased efficiency of monitoring and resources used by a combination of on-site and remote monitoring using a web-conference system (Hirase 2016).

The qualitative finding in TEMPER (Stenning 2018b) that the experience of the research nurse had an impact on the monitoring outcomes is also reflected in the retrospective study by von Niederhäusern and colleagues, which found that one of the factors associated with lower numbers of monitoring findings was experienced site staff and concluded that the human factor was underestimated in the current risk-based monitoring approach (von Niederhausern 2017).

Authors' conclusions

Implications for systematic reviews and evaluations of healthcare

We found no evidence for inferiority of a risk-based monitoring approach compared to extensive on-site monitoring in terms of critical and major monitoring findings. The overall certainty of the evidence for this outcome was moderate. The initial risk assessment of a study can facilitate a reduction of monitoring. However, it might be more efficient to use the outcomes of a risk assessment to guide on-site monitoring in terms of prioritizing sites with conspicuously low performance quality of critical assets identified by the risk assessment. Some triggers that were used in the TEMPER study (Stenning 2018b) and Knott 2015 could help identify sites that would benefit the most from an on-site monitoring visit. Trigger refinement and inclusion of more trial-specific triggers will, however, be necessary. The development of remote

access to trial documentation may further improve the impact of central triggers. Timely central monitoring of consent forms or eligibility documents with adequate anonymization and data protection may mitigate the effects of many formal documentation errors. More studies are needed to assess the feasibility of eligibility and informed consent-related assessment and remote contact to the site teams in terms of data security and effectiveness without on-site review of documents. The COVID-19 pandemic has resulted in innovative monitoring approaches in the context of restricted on-site monitoring that also includes the remote monitoring of consent forms and other original records as well as compliance to study procedures usually verified on-site. Whereas central data monitoring and remote monitoring of documents were formerly applied to improve efficiency, it now has to substitute on-site monitoring to comply with pandemic restrictions, making evaluated monitoring methods in this review even more valuable to the research community. Both the Food and Drug Administration (FDA) and European Medicines Agency have provided guidance on aspects of clinical trial conduct during the COVID-19 pandemic including remote site monitoring, handling informed consent in remote settings, and the importance of maintaining data integrity and audit trail (EMA 2021; FDA 2020). The FDA has also adopted contemporary approaches to consent involving telephone calls or video visits in combination with a witnessed signing of the informed consent (FDA 2020). Experiences on new informed consent processes and advice on how remote monitoring and centralized methods can be used to protect the safety of patients and preserve trial integrity during the pandemic have been published and provide additional support for sites and sponsors (Izmailova 2020; Love 2021; McDermott 2020). This review may support study teams faced by pandemic-related restrictions with information on evaluated methods that focus primarily on remote and centralized methods. It will be important to provide more management support for clinical trials in the academic setting and develop new recruitment strategies. In our review, low certainty of evidence suggested that initiation visits or more frequent on-site visits were not associated with increased recruitment or retention of trial participants. Consequently, trial investigators should plan for other, more trial-specific strategies to support recruitment and retention. To what extent recruitment or retention can be improved through real-time central monitoring remains to be evaluated. Research has emphasized the need for evidence on effective recruitment strategies (Treweek 2018b), and new flexible recruitment approaches initiated during the pandemic may add to this. During the COVID-19 pandemic, both social media and digital health platforms have been leveraged in novel ways to recruit heterogeneous cohorts of participants (Gaba 2020). In addition, the pandemic

underlines the need for a study management infrastructure supported by central data monitoring and remote communication (Shiely 2021). One retrospective study at the Beijing Cancer Hospital assessed the impact of their newly implemented remote management model on critical trial indicators: protocol compliance rate, rate of loss to follow-up, rate of participant withdrawal, rates of disease progression and mortality, and detection rate of monitoring problems (Fu 2021). The measures implemented after the first COVID-19 outbreak led to significantly higher rates of protocol compliance and significantly lower rates of loss to follow-up or withdrawal after the second outbreak compared to the first, without affecting rates of disease progression or mortality. In general, new experiences with electronic methods initiated throughout the COVID-19 pandemic might facilitate development and even improvement of clinical trial management.

Implications for methodological research

Several new monitoring interventions were introduced in recent years. However, the evidence base gathered for this Cochrane Review is limited in terms of quantity and quality. Ideally, for each of the five identified comparisons (risk-based versus extensive on-site monitoring, central statistical monitoring with triggered on-site visits versus regular [untriggered] on-site visits, central and local monitoring with annual on-site visits versus central and local monitoring only, traditional 100% source data verification [SDV] versus remote or targeted SDV, and on-site initiation visit versus no on-site initiation visit) more randomized monitoring studies nested in clinical trials and measuring effects on all outcomes specified in this review are necessary to draw more reliable conclusions. The development of triggers to guide on-site monitoring while centrally monitoring incoming data is ongoing and different triggers might be used in different settings. In addition, more evidence on risk indicators that help to identify sites with problems or the prognostic value of triggers is needed to further optimize central monitoring strategies. Future methodological research should particularly evaluate approaches with an initial trial-specific risk assessment followed by close central monitoring and the possibility for triggered and targeted on-site visits during trial conduct. Outcome measures such as the impact on recruitment, retention, and site support should be emphasized in further research and the potential of central monitoring methods to support the whole study management process needs to be evaluated. Directing monitoring resources to sites with problems independent of data quality issues (recruitment, retention) could promote the role of experienced study monitors as a site support team in terms of training and advice.

The overall progress in conduct and success of a trial should be considered in the evaluation of every new approach. The fact that most of the eligible studies identified for this review are government or charity funded suggests a need for industry-sponsored trials to evaluate their monitoring and management approaches. This could particularly promote the development and evaluation of electronic case report form-based centralized monitoring tools, which require substantial resources.

Acknowledgements

We thank the monitoring team of the Department of Clinical Research at the University Hospital Basel, including Klaus Ehrlich, Petra Forst, Emilie Müller, Madeleine Vollmer, and Astrid Roesler, for sharing their experience and contributing to discussions on monitoring procedures. We would further like to thank the information specialist Irma Klerings for peer reviewing our electronic database searches.

Contributions of authors

KK, CPM, and MB conceived the study and wrote the first draft of the protocol.

SL, MS, PB, NB, HE, PAJ, and MMB reviewed the protocol and suggested changes for improvement.

HE and KK developed the search strategy and conducted all searches.

KK, CPM, and MB screened titles and abstracts as well as full texts, and selected eligible studies.

KK and MMB extracted relevant data from included studies and assessed risk of bias.

KK conducted the statistical analyses and interpreted the results together with MB and CPM.

KK and MB assessed the certainty of the evidence according to GRADE and wrote the first draft of the review manuscript.

CPM, SL, MS, PB, NB, HE, PAJ, and MMB critically reviewed the manuscript and made suggestions for improvement.

Supporting information

Differences between protocol and review

We did not estimate the intracluster correlation and heterogeneity across sites within the ADAMON and OPTIMON studies as planned in our review protocol (Klatte 2019) due to lack of information. .

We planned in the protocol to assess the statistical heterogeneity of studies in meta-analyses. Due to the small number of included studies per comparison, it was not reasonable to assess heterogeneity statistically.

Planned sensitivity analyses were also not performed because of the small number of included studies.

We removed characteristics of monitoring strategies from the list of secondary outcomes upon request of reviewers and included the information in the section on general characteristic of included studies. We changed the order of the secondary outcomes in an attempt to improve the logical flow of the Results section.

Characteristics of studies

Characteristics of included studies

Brosteanu 2017b

Methods	Design: cluster randomized study Duration of monitoring study: 7 years (due to funding and time limitations, audits were performed in 4 trials after last participant was recruited but before the end of trial; in 2 trials, accrual was still ongoing at the time trial sites were audited; in these cases, audits were restricted to participants having completed their treatment) Support for participating sites: CTU
Data	Monitoring data from 11 randomized trials with trial sites randomized to 2 different monitoring strategies (randomized at the beginning of the trial)
Comparisons	**Intervention:** initial risk assessment according to Brosteanu 2009 with 3 different risk levels and corresponding intensity of on-site monitoring **Control:** extensive on-site monitoring without risk assessment
Outcomes	**Primary outcome:** participant-level composite outcome (informed consent process violation, eligibility criteria violation, SAE reporting violation, errors in endpoint assessment, protocol deviation with impact on patient safety or data validity) **Secondary outcomes:** economic data (mean number of monitoring visits and time spent on-site)
Clinical area and setting of host trial	International and national multicenter trials in secondary and tertiary care in the areas of oncology, neonatology, neurology, intensive care, surgery, and cardiology, including adults and children; involved countries: Germany and the US
Number of patients randomized (analyzed)	1967 randomized (1920 analyzed) participants in 213 randomized (156 analyzed) sites; difference in number of participants randomized and analyzed due to inclusion of sites that did not recruit any participants
Notes	Funding source: German Federal Ministry for Education and Research (non-industry funded) Published as peer-reviewed article in English

Risk of bias table

Item	Authors' judgment	Support for judgment
Selection bias	Low	Randomization of trial sites within participating trials was performed centrally in Leipzig.
Performance bias	High	Quote: "Trial sites were informed by their respective trial sponsor about ADAMON and the planned audits, but not about the assigned monitoring arm. Sponsors, Monitors and ADAMON team were aware of assignment."
Detection bias	Low	Quote: "Audit teams were not informed of the sites' monitoring strategy and did not have access to any

Item	Authors' judgment	Support for judgment
		monitoring reports. Audit findings were reviewed in a blinded manner by members of the ADAMON team and discussed with auditors, as necessary, to ensure that reporting was consistent with the ADAMON audit manuals."
Attrition bias	Low	However: (quote) "… one site refused the audit, and in the last five audited trials, 29 sites with less than three patients were not audited due to limited resources, in large sites (>45 patients), only a centrally preselected random sample of patients was audited. Arms are not fully balanced in numbers of patients audited (755 extensive on-site monitoring and 863 risk-adapted monitoring) overall."
Reporting bias	Low	Protocol available, no indication of selective reporting.
Other bias	Low	

Fougerou-Leurent 2019

Methods	Design: prospective cross-over study Duration of monitoring study: 2 years Support for participating sites: Clinical Investigation Center, INSERM, Rennes, France
Data	Monitoring data from 126 participants in 6 ongoing phase II and phase III randomized trials (selected participants for whom the data monitoring had not started)
Comparisons	**Intervention:** targeted SDV on key data for all participants **Control:** full SDV on 100% of data points for 100% of participants
Outcomes	**Primary outcome:** error rate in the final dataset prepared using the targeted SDV monitoring process, on total data and on key data. **Secondary outcomes:** impact of targeted SDV on the DM workload and the staffing cost of the trial. Secondary endpoints were the number of discrepancies between the datasets prepared using the 2 monitoring strategies at each step, the number of queries issued with each strategy, and the time spent on SDV and DM with each strategy
Clinical area and setting of host trial	National, single center/multicenter trials in secondary and tertiary care settings involving adults (one trial was multinational, the others were national); limited to Rennes, France
Number of patients randomized (analyzed)	126 randomized in the monitoring study (126 analyzed in the monitoring study)
Notes	Funding source: University Hospital Rennes (non-industry funded) Published as peer-reviewed article in English

Risk of bias table

Item	Authors' judgment	Support for judgment
Selection bias	Low	Prospective cross-over design: the same CRFs were analyzed with full or targeted SDV. Participants from Rennes, for whom the data monitoring had not started.
Performance bias	High	It is difficult to blind personnel on full vs partial SDV.
Detection bias	Low	The same DM program (missing data, consistency, protocol deviations) was subsequently implemented in each strategy by central DM staff. No information on blinding.
Attrition bias	Low	All outcomes of methods section included in the outcome data.
Reporting bias	Low	No indication for reporting bias, all outcomes were reported in the methods section.
Other bias	Low	If the CRA spotted a false or missing non-key data when checking a key data, they may have corrected the non-key data in the CRF. This bias may have underestimated the difference between the 2 monitoring strategies. The full SDV CRF was considered without errors.

Journot 2017

Methods	Design: cluster randomized trial Duration of monitoring study: 3 years (OPTIMON staff collected OPTIMON data after completion of monitoring of the trials by the responsible CTU. When the duration for recruitment or main endpoint collection was > 6 months or 1 year, OPTIMON outcome variables were collected at an earlier time point, and only for a certain number of participants) Support for participating sites: clinical research centers
Data	Monitoring data from 22 trials (15 randomized trials, 4 cohort studies, 3 cross-sectional studies) on participants and trial sites (83 proposed) randomized to 2 different monitoring strategies
Comparisons	**Intervention:** initial risk assessment published in Journot 2011 – 4 different risk levels (A, B, C, D) – different degrees of monitoring **Control:** full on-site monitoring (including SDV) without risk assessment

Outcomes	**Primary outcome:** participant-level composite outcome (eligibility violations, informed consent violations, SAE reporting violation, value missing for the primary endpoint) **Secondary outcomes:** economic data (indicators of direct and indirect costs. (The costs directly related to applying each strategy should be taken into account stating: 1. investments necessary in material and training and costs of maintenance, which thus provides the cost of acquisition. Investments classified as redeployable or not, i.e. whether or not limiting the possibility of doing other things in the future and therefore the cost of abandoning; 2. costs related to carrying out the study (if possible, individual per participant); 3. cost of the detection of errors; 4. cost of the consequences of detected and undetected errors; 5. cost of the surveillance of the monitoring strategies) Timeliness, overall data completeness, breakdown of the main judgment criterion according to the type of serious error (proportion of errors related to consent, proportion of errors relating to serious or unexpected adverse events, proportion of errors relating to eligibility criteria, proportion of errors relating to the main judgment criterion of the clinical research study)
Clinical area and setting of host trial	National and international, multicenter trials in secondary care settings and including adults, older people, and children. 19 studies dealt with chronic diseases. 10 studies were on specific populations. 8 studies with risk level A, 4 with risk level B, and 10 with risk level C. Countries involved: France
Number of patients randomized (analyzed)	954 participants randomized in monitoring study (759 analyzed), randomization of 83 sites (68 analyzed); difference in number of participants randomized and analyzed due to inclusion of sites that did not recruit any participants
Notes	Funding source: French National Hospital Clinical Research Program (PHRC) (academic funded) Only published as abstract and conference proceedings, no full report published

Risk of bias table

Item	Authors' judgment	Support for judgment
Selection bias	Low	Randomization by the OPTIMON team's statistician and validated by an independent statistician. Randomization carried out per level in line with the A, B, or C risk levels of the clinical research studies. A complete document describing the randomization procedure (methods, block size, program used) was kept confidentially by the OPTIMON team's statistician. The result of the randomization was automatically sent to the methodology and management center by fax.
Performance bias	High	Randomization was kept confidential and site staff was not informed about assignment. Monitors were not blinded and the same CRA was allowed to performed the monitoring in both arms of the same study.
Detection bias	Low	Assessors were not blinded. However, main outcome was validated by a blinded validation committee.
Attrition bias	Low	No indication of missing data (some sites did not recruit any participants and were not included in the analysis, balanced between groups).
Reporting bias	Low	Protocol available at the study homepage, no full report published yet but data available from conference

Item	Authors' judgment	Support for judgment
		presentations.
Other bias	Low	

Knott 2015

Methods	Design: matched comparator design Duration: 18 months Support for participating sites: Clinical Trial Service Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK
Data	Monitoring data from 21 sites (6 UK sites, 4 China, 11 Scandinavia) of 1 international trial in 245 sites included in analysis
Comparisons	**Intervention:** on-site monitoring visits targeted as a result of high scores determined through central statistical monitoring procedures **Control:** on-site visits in comparator sites chosen by the regional coordinating center among low scoring sites determined through central statistical monitoring procedures
Outcomes	**Primary outcome:** site-level composite outcome. Proportion of sites with ≥ 1 major or serious finding not already identified through central monitoring **Secondary outcomes:** proportion of sites with ≥ 1 minor finding, proportion of sites with ≥ 1 serious finding
Clinical area and setting of host trial	International, multicenter trial; countries involved: UK, China, Scandinavia (Norway, Finland, Sweden, Denmark)
Number of patients randomized (analyzed)	No information on number of participants included in the study (25,673 were randomized in the host trial). 238 sites were considered in the central statistical monitoring procedure and 21 sites were included in the comparison
Notes	Funding source: Clinical Trial Service Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK (non-industry funded) Only published as conference abstract, no full report published

Risk of bias table

Item	Authors' judgment	Support for judgment
Selection bias	High	Non-randomized study. Matched comparator design.
Performance bias	High	Monitors performing the on-site visits were not blinded.
Detection bias	Unclear	No full report published yet.
Attrition bias	Low	No full report published yet, but all available information provided by the study team.
Reporting bias	Low	No full report published yet, but all available information provided by the study team.
Other bias	Low	

Lièvard 2006

Methods	Design: cluster randomized trial Duration: 2 years Support for participating sites: International Drug Development Institute
Data	Monitoring data from 573 participants in 135 participating centers of a large cancer trial
Comparisons	**Intervention:** monitoring strategy where on-site initiation visits were only performed when requested by the investigator **Control:** monitoring strategy that included the routine on-site initiation visits
Outcomes	**Primary outcome:** outcomes of interest to assess the impact of on-site monitoring visits were: number of randomized participants per center, length of participant follow-up in each center, number of CRF pages submitted by each center to the coordinating office, and quality of data assessed by the number of computer-generated data queries for each center (queries per page and queries per participant). Data inserted into Excel **Secondary outcomes:** economic data. Time spent for monitoring was reported in discussion section, but defined as a secondary outcome.
Clinical area and setting of host trial	Multinational, multicenter trial in secondary care centers and involving only adults in the study population; involved countries: only centers in France participated in the methodological substudy of the cancer trial
Number of patients randomized (analyzed)	573 participants randomized (573 analyzed)
Notes	Funding source: the host trial (AERO B-2000) was mainly supported by an unrestricted research grant from Bristol-Myers Squibb France with additional support from Chugai Laboratories (industry-funded) Published as peer-reviewed article in English

Risk of bias table

Item	Authors' judgment	Support for judgment
Selection bias	Low	French centers that had expressed an interest in the trial were randomly allocated by the coordinating office (International Drug Development Institute, Brussels, Belgium)
Performance bias	High	Investigators were not informed that they would be randomized to be visited or not, for such information might have compromised the purpose of the study. They were told that the trial budget would not allow for regular, extensive on-site monitoring visits such as those typically performed in registration trials of new drugs. Investigators requesting on-site visits were visited regardless of the randomized group their center had been allocated to.
Detection bias	Unclear	For the outcome recruitment blinding is not necessary. Unclear if data managers assessing the quality of the data submitted were blinded.
Attrition bias	Low	Data did not appear to have been excluded. However,

Item	Authors' judgment	Support for judgment
		because the study was terminated prematurely, the reported data were incomplete in terms of what was planned for the study. Number of centers that randomized participants was equal in both groups.
Reporting bias	Low	All outcomes reported in the methods section were reported. Some data were incomplete due to premature termination (e.g. participant follow-up). Hours of work for monitoring were reported in the discussion session, but not in the methods or results.
Other bias	Low	

Mealer 2013

Methods	Design: randomized trial Duration: 2 years (pilot study) Support for participating sites: coordinating centers of the ARDS and ChiLDReN networks
Data	Monitoring data from 32 participants in trials from 2 large trial networks
Comparisons	**Intervention:** remote SDV **Control:** full on-site SDV
Outcomes	**Primary outcome:** accuracy and completeness of remote SDV vs on-site monitoring determined by analyzing the number of data values assigned to 4 outcomes: 1. Found-match (data value recorded on the CRF matched the data value in the source document); 2. Found-different (data value recorded on the CRF was different (did not match) the data value in the source document); 3. Missing data (value recorded on the CRF could not be found in the source document); and 4. Unknown (no data on the CRF or in the source document related to a data value that was supposed to be collected) compared to all data values other than those assigned to the "not monitored" outcome. **Secondary outcomes:** economic outcome data – efficiency was measured by analyzing the amount of time it took to complete the SDV tasks by individual data item and by CRF form.
Clinical area and setting of host trial	National, multicenter trials in secondary and tertiary care settings including adults and children in their study population. Involved countries: USA
Number of patients randomized (analyzed)	32 participants randomized (32 analyzed)
Notes	Funding source: NIH/NCATS Colorado CTSI Grant Number UL1 TR000154. The ARDS network was supported by HHSN268200536-179C (MGH) and N01-56167 (University of Colorado). The ChiLDReN network is supported by CCC: 5U01DK062456-11 (University of Michigan) and 2U01DK06243-08 (University of Colorado) (non-industry funded) Published as peer-reviewed article in English

Risk of bias table

Item	Authors' judgment	Support for judgment
Selection bias	Unclear	Quote: "Our study is also limited by the non blinded randomization method chosen."
Performance bias	High	For each research network, the same monitor performed both remote and local monitoring. Remote monitors had telephone access to the same local coordinators who were available during on-site monitoring visits.
Detection bias	Unclear	Monitoring was not performed blindly. Unclear if the analysis was done blinded.
Attrition bias	Low	No attrition reported.
Reporting bias	Low	No indication for reporting bias, all outcomes were reported in the methods section
Other bias	Low	

Stenning 2018b

Methods	Design: prospective matched-pair study Duration: 31 months Support for participating sites: MRC CTU at UCL, Cancer Research UK, UK
Data	Monitoring data from 42 matched paired visits (84 visits) at 63 sites were included in the analysis. The matching algorithm proposed untriggered sites to visit, minimizing differenced in number of participants, and time since first participant randomized and maximizing differences in trigger score
Comparisons	**Intervention:** triggered monitoring strategy in which targeted on-site monitoring based on trial data and conduct that were scrutinized centrally with prespecified triggers for visits to sites **Control:** normal on-site visits to sites without activated triggers
Outcomes	**Primary outcome:** site-level composite outcome (eligibility violations, informed-consent violations, SAE reporting violations, errors in key data and endpoint assessment, errors in pharmacy documents and facilities, and investigator site files). Defined as proportion of sites with ≥ 1 major or critical finding not already identified through central monitoring or a previous visit ('new' findings). **Secondary outcomes:** number of major and critical findings, number of critical findings, proportion of sites with ≥ 1 critical finding and category of major/critical findings
Clinical area and setting of host trial	UK sites in 3 well-established international, multicenter trials cancer trials in secondary care setting including adults only. Involved countries: UK
Number of patients randomized (analyzed)	42 matched paired visits conducted (84 visits) at 63 sites
Notes	Funding source: Cancer Research UK (grant C1495/A13305 from the Population Research Committee); Medical Research Council (MC_EX_UU_G0800814) and the MRC London Hub for Trial Methodology Research (MC_UU_12023/24) (non-industry funded) Published as peer-reviewed article in English

Risk of bias table

Item	Authors' judgment	Support for judgment
Selection bias	High	Non-randomized study. Investigators attempted to balance groups by design and controlled for known confounding factors by using the Microsoft matching algorithm.
Performance bias	High	To ensure visits were arranged and conducted as per normal practice, site staff was not explicitly informed about the TEMPER study or the reason for a monitoring visit. The trials unit staff present at triggered and untriggered visits was not blind to visit type.
Detection bias	Low	Observation bias due to lack of blinding of monitoring staff was mitigated by consistent training on the trials and monitoring methods, the use of a common finding grading system and independent review of all major and critical findings that was blind to visit type.
Attrition bias	Low	All 84 visits were included in the analysis.
Reporting bias	Low	No indication of reporting bias. Scores of matched sites are published in Diaz-Montana 2019a.
Other bias	Low	Exact site selection is not fully reported: chosen sites usually had the highest total trigger scores, but general concerns sometimes led to other sites being prioritized. Visits per site (triggered and control) were not reported. Only that 84 visits were completed in 63 sites (of 156 total).

Wyman 2020

Methods	Design: cluster randomized trial Duration: 5.25 years Support for participating sites: all clinical sites were associated with 1 of 4 international coordinating centers, located in Copenhagen, Denmark; London, UK; Sydney, Australia; and Washington DC, US
Data	Monitoring data from 1 randomized trial in infectious disease with sites randomized to 2 different monitoring strategies; data collection for the monitoring study included 4371 participants (2107 participants in the on-site group, 2264 in the no on-site group) from 196 sites in 34 countries
Comparisons	**Intervention:** central and local monitoring alone **Control:** central, local, and on-site monitoring
Outcomes	**Primary outcome:** participant-level composite outcome (eligibility violations, primary event/SAE not reported within 6 months, informed consent violations, use of antiretroviral therapy not permitted by START, data alteration) **Secondary outcomes:** economic data (person-hours spent conducting on-site monitoring), percentage of participants lost to follow-up, percentage of missed follow-up data collection visits, data submission timelines

Clinical area and setting of host trial	1 international, multicenter trial in infectious disease in a secondary care setting including adults only; involved countries: 34 countries from Europe, North America, South America, Asia, and Africa (Argentina, Australia, Austria, Belgium, Brazil, Chile, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, India, Ireland, Israel, Italy, Luxembourg, Malaysia, Mali, Mexico, Morocco, Nigeria, Norway, Peru, Poland, Portugal, Puerto Rico, South Africa, Spain, Sweden, Switzerland, Thailand, Uganda, the UK, the USA)
Number of patients randomized (analyzed)	4371 participants in 196 sites
Notes	Funding source: National Institute of Allergy and Infectious Disease (non-industry funded) Published as peer-reviewed article in English

Risk of bias table

Item	Authors' judgment	Support for judgment
Selection bias	Low	Site randomization was stratified by country and projected START-MV enrollment (< 15, 15–30, > 30 participants), and was carried out by the statistical center using block randomization prior to the beginning of the substudy.
Performance bias	High	Coordinating centers were informed of the assignments. While sites were not notified of the randomization assignment it was not blinded, as, within the first year, sites randomized to the central + local + on-site monitoring arm were contacted to schedule a monitoring visit. It is unclear if monitors performing the on-site visits were blinded.
Detection bias	Unclear	No indication of blinded outcome assessment. Quote: "A procedure was implemented for statistical center staff to centrally review consent violations found by on-site monitors to determine if the violation met the revised criteria." No information whether statistical center staff were blinded.
Attrition bias	Low	All randomized sites were included in the analysis.
Reporting bias	Low	No indication of selective reporting based on the design paper introducing the INSIGHT START monitoring substudy (Hullsiek 2015).
Other bias	Low	

ARDS network: Acute Respiratory Distress Syndrome network; ChiLDReN: Childhood Liver Disease Research Network; CRA: clinical research associate; CRF: case report form; CTU: clinical trials unit; DM: data management; SAE: serious adverse event; SDV: source data verification.

Characteristics of excluded studies

Study	Reason for exclusion
Agrafiotis 2018	Not a prospective study
Andersen 2015	Not a prospective study
Bailey 2017	No comparison of different monitoring strategies (only abstract available)

Bakobaki 2011	Not a prospective study
Bakobaki 2012	Not a prospective study
Biglan 2016	Not a prospective study (only abstract available).
Collett 2019	No comparison of different monitoring strategies
Cragg 2019	Not a prospective study
Del Alamo 2018	No comparison of different monitoring strategies
Diani 2017	Not a prospective study
Diaz-Montana 2019b	No comparison of different monitoring strategies
Edwards 2014	No comparison of different monitoring strategies
Elsa 2011	No comparison of different monitoring strategies
Fu 2021	Not a prospective study
Hatayama 2020	No comparison of different monitoring strategies
Heels-Ansdell 2010	No comparison of different monitoring strategies
Higa 2020	Not a prospective study
Hirase 2016	Not a prospective study
Jones 2019	Not a prospective study (abstract only)
Jung 2020	No comparison of different monitoring strategies (centralized monitoring used only for medication adherence)
Kim 2011	Not a prospective study (abstract only)
Kim 2021	Not a prospective study
Lane 2013	No comparison of different monitoring strategies
Lim 2017	No comparison of different monitoring strategies
Lindley 2015	No comparison of different monitoring strategies (abstract only)
Miyamoto 2019	No comparison of different monitoring strategies
Morales 2020	No comparison of different monitoring strategies
Murphy 2019	No comparison of different monitoring strategies (abstract only)
Pei 2019	No comparison of different monitoring strategies
Stock 2017	No comparison of different monitoring strategies
Sudo 2017	No comparison of different monitoring strategies
Thom 1996	No comparison of different monitoring strategies
Tudur Smith 2012b	Not a prospective study
von Niederhäusern 2017	Not a prospective study
Yamada 2021	Not a prospective study
Yorke-Edwards 2019	No comparison of different monitoring strategies
Zhao 2013	No comparison of different monitoring strategies

Summary of findings tables

1 Risk-based versus extensive on-site monitoring

Risk-based monitoring compared with extensive on-site monitoring for clinical intervention studies					
Patient or population: **clinical trials in all fields of health care** Settings: **international/national trials** Intervention: **risk-based monitoring strategy** Comparison: **extensive on-site monitoring**					
Outcomes		**Relative effect (95% CI)**	**No of participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Combined outcome of proportion of participants with major or critical monitoring findings		**RR 1.03** (0.80 to 1.33)	2377 (2 studies [nested in 33 clinical trials])	⨁⨁⨁◯ **Moderate**[a]	—
Impact of the monitoring strategy on participant on recruitment		—	—	—	Not reported.
Impact of the monitoring strategy on follow-up		—	—	—	Not reported.
Effect of the monitoring strategy on resource use	**ADAMON:** number of monitoring visits per participant and the cumulative monitoring time	Higher for on-site monitoring by a factor of **2.1 to 2.7** (Ratios of the efforts calculated within each trial and summarized with the geometric mean)	—	⨁⨁◯◯ **Low**[b]	—
	OPTIMON: costs of monitoring	Higher for on-site by a factor of **2.7**			
	OPTIMON: costs of travel and monitoring	Higher for on-site by a factor of **3.4**			
ADAMON: **ADApted MONitoring study;** CI: **confidence interval;** OPTIMON: **Optimisation of Monitoring for Clinical Research Studies;** RR: **risk ratio.**					

GRADE Working Group grades of evidence
High quality: **further research is very unlikely to change our confidence in the estimate of effect.**
Moderate quality: **further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.**
Low quality: **further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.**
Very low quality: **we are very uncertain about the estimate.**

Footnotes

[a]Downgraded one level due to the imprecision of the summary estimate with the 95% confidence interval including the substantial advantages and disadvantages with the risk-based monitoring intervention.
[b]Downgraded two levels due to substantial imprecision; there were no confidence intervals for either of the two estimates on resource use provided in the ADAMON and OPTIMON studies and the two estimates could not be combined due to the nature of the estimate (resource use versus cost calculation).

2 Central monitoring with triggered versus untriggered on-site visits

Central statistical monitoring with triggered on-site visits compared with regular (untriggered) on-site visits for clinical intervention studies				
Patient or population: **clinical trials in all fields of health care** Settings: **international/national trials** Intervention: **triggered on-site visits** Comparison: **regular (untriggered) on-site visits**				
Outcomes	**Relative effect (95% CI)**	**No of participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Sites ≥ 1 major monitoring finding combined outcome	**RR 1.92** (0.40 to 9.17)	105 sites (2 studies)	⨁⨁◯◯ **Low**[a]	—
Impact of the monitoring strategy on participant recruitment	—	—	—	Not reported.
Impact of the monitoring strategy on follow-up	—	—	—	Not reported.
Effect of the monitoring strategy on resource use	—	—	—	Not reported.
***The basis for the** assumed risk **(e.g. the median control group risk across studies) is provided in footnotes. The** corresponding risk **(and its 95% confidence interval) is based on the assumed risk in the comparison group and the** relative effect **of the intervention (and its 95% CI).** CI: **confidence interval;** RR: **risk ratio.**				
GRADE Working Group grades of evidence High quality: **further research is very unlikely to change our confidence in the estimate of effect.** Moderate quality: **further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.** Low quality: **further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.** Very low quality: **we are very uncertain about the estimate.**				

Footnotes

[a]Downgraded one level because both studies were not randomized, and downgraded one level for imprecision.

3 Central and local monitoring only versus central and local monitoring with on-site visits

Central and local monitoring only compared with central and local monitoring with annual on-site visits for clinical trials				
Patient or population: **clinical trials in all fields of health care** Settings: **international/national trials** Intervention: **central and local monitoring only** Comparison: **central and local monitoring with annual on-site visits**				
Outcomes	**Relative effect (95% CI)**	**No of participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Combined outcome of proportion of participants with major or critical monitoring findings	**OR 1.7** (1.1 to 2.7)	4371 (1 study nested in 1 clinical trial)	⨁⨁⨁◯ **Moderate**[a]	Prior defined monitoring findings were very study specific and central monitoring was present in both intervention arms, which might explain the low number of events. Percentages of findings were higher in the on-site group, but the overall impact of these findings on the study was low due to the low absolute number of events.
Impact of the monitoring strategy on participant recruitment	—	—	—	Not reported.
Impact of the monitoring strategy on follow-up	**OR 0.8** (0.5 to 1.1)	4371 (1 study nested in 1 clinical trial)	⨁◯◯◯ **Very low**[b]	—
Effect of the monitoring strategy on resource use	**USD 2,035,392**	—	⨁◯◯◯ **Very low**[c]	—
CI: **confidence interval;** OR: **odds ratio.**				
GRADE Working Group grades of evidence High quality: **further research is very unlikely to change our confidence in the estimate of effect.** Moderate quality: **further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.** Low quality: **further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.** Very low quality: **we are very uncertain about the estimate.**				

Footnotes

[a]Downgraded one level because the estimate was based on a small number of events and because the estimate stemmed from a single study nested in a single trial (indirectness).
[b]Downgraded three levels because the 95% confidence interval of the estimate allowed for substantial benefit as well as substantial disadvantages with the intervention and there were only a small number of events (serious imprecision); in addition, the estimate stemmed from a single study nested in a single trial (indirectness).
[c]Downgraded three levels because the estimate was not accompanied by a confidence interval (imprecision) and because the estimate stemmed from a single study nested in a single trial (indirectness).

4 Remote or targeted source data verification versus 100% source data verification

Remote or targeted SDV compared with traditional 100% SDV for clinical intervention studies					
Patient or population: **clinical trials in all fields of health care** Settings: **international/national trials** Intervention: **remote or targeted SDV** Comparison: **traditional 100% SDV**					
Outcomes		**Relative effect (95% CI)**	**No of participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Monitoring findings	**MONITORING:** overall error rate with targeted SDV	**1.47%** (1.41% to 1.53%)	126 (1 study nested in 6 clinical trials)	⨁◯◯◯ **Low**[a]	—
	MONITORING: error rate on key data with targeted SDV	**0.78%** (0.65% to 0.91%)			
	Mealer et al.: percentage of data values that could not be correctly identified via remote monitoring	**0.47%** (0.03% to 0.79%)	32 (1 study nested in 2 large trial networks)		
Impact of the monitoring strategy on participant recruitment		—	—	—	Not reported.
Impact of the monitoring strategy on follow-up		—	—	—	Not reported.
Effect of the monitoring strategy on resource use	**MONITORING:** saving on monitoring costs by targeted SDV strategy	**EUR 5841**	126 (1 study nested in 6 clinical trials)	⨁◯◯◯ **Very low**[b]	—
	MONITORING: additional cost of data management for targeted SDV (queries)	**EUR 8922**			
	Mealer et al.: time per case report (mean with SD) remote vs on-site	Adult: 4.60 (SD 1.42) min vs 3.60 (SD 0.96) min (P = 0.10);	32 (1 study nested in 2 large trial networks)		

Remote or targeted SDV compared with traditional 100% SDV for clinical intervention studies				
		pediatric: 11.64 (SD 7.54) min vs 6.07 (SD 3.18) min (2-tailed t-test, P = 0.10)		
CI: **confidence interval;** min: **minute;** RR: **risk ratio;** SD: **standard deviation;** SDV: **source data verification.**				
GRADE Working Group grades of evidence High quality: **further research is very unlikely to change our confidence in the estimate of effect.** Moderate quality: **further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.** Low quality: **further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.** Very low quality: **we are very uncertain about the estimate.**				

Footnotes

[a]Downgraded two levels because randomization was not blinded in one of the studies and the outcomes of the two studies could not be combined.
[b]Downgraded by one additional level in addition to (a) for imprecision because there were no confidence intervals provided.

5 Monitoring with versus without initiation visit

No on-site initiation visit compared with on-site initiation visit for clinical intervention studies				
Patient or population: **clinical trials in all fields of health care** Settings: **international/national trials** Intervention: **no on-site initiation visit** Comparison: **on-site initiation visit**				
Outcomes	**Relative effect (95% CI)**	**No of participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Monitoring findings	—	—	—	Not reported.
Impact of the monitoring strategy on participant recruitment **Difference in the number of recruited participants between groups visited vs non-visited**	**302 vs 271** (no statistically significant difference)	573 (1 study nested in 1 clinical trial)	⨁◯◯◯ **Very low**[a]	—
Impact of the monitoring strategy on follow-up **Mean follow-up time, calculated from the date of randomization to the date**	**1.8 (SD 3.2) vs 2.5 (SD 3.6) months**	573 (1 study nested in 1 clinical trial)	⨁◯◯◯ **Very low**[b]	—

No on-site initiation visit compared with on-site initiation visit for clinical intervention studies				
of last form received, visited vs non-visited				
Effect of the monitoring strategy on resource use	—	—	—	Not reported.
CI: **confidence interval;** SD: **standard deviation.**				
GRADE Working Group grades of evidence High quality: **further research is very unlikely to change our confidence in the estimate of effect.** Moderate quality: **further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.** Low quality: **further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.** Very low quality: **we are very uncertain about the estimate.**				

Footnotes

[a]Downgraded three levels because of substantial imprecision (relevant advantages and relevant disadvantages were plausible given the small amount of data), and indirectness (a single study nested in a single trial).

[b]We downgraded by one additional level in addition to (a) for imprecision due to the small number of events.

Additional tables

1 Definitions of combined monitoring outcomes

	ADAMON (translated from German study protocol Brosteanu 2017b)	OPTIMON (Journot 2015)	START (Wyman 2020)	TEMPER (Stenning 2018a)	Knott 2015
General definition (major or critical)	Primary endpoint of the ADAMON study was the proportion of audited participants with ≥ 1 major or critical violation of essential GCP objectives in ≥ 1 of **5 error domains:** informed consent process, participant selection, intervention, endpoint assessment, and SAE reporting. Major or critical GCP violations referred to as 'major audit findings' were determined in independent ADAMON audits at the end of the trial looking at all individual participants in all trial sites. Audit manuals defined trial-specific protocol	The main judgment criterion was the proportion of participants whose observation for the clinical research study contained no serious errors. It was a composite criterion, measured at the individual (participant) level. The errors concerned the following **2 regulatory aspects** – consent and serious or unexpected adverse events – and the following **2 aspects concerning the scientific integrity of the data** – failure to respect eligibility criteria without prior dispensation, and incorrect value or data missing for the main judgment criterion. Considered errors for the analysis (major non-conformities) were protocol or	The primary outcome for the monitoring sub study was a participant-level composite outcome consisting of **6 major components**: major eligibility violations, major informed consent violations, use of ART for initial therapy that is not permitted by the START protocol, ≥ 6-month delay in reporting START primary endpoints or serious events, and data alteration or fraud.	The primary outcome measure was the proportion of sites with ≥ 1 major or critical finding not already identified through central monitoring or a previous visit. Critical findings: those that impact, or potentially could impact, directly on participant safety or confidentiality, or create serious doubt in the accuracy or credibility of trial data. Major findings: included deviations from the protocol that may have resulted in questionable data being obtained, or errors that consisted of a number of minor deviations from regulations, suggesting that procedures were not being followed. Any major finding that was not corrected, or that	The primary outcome measure was the proportion of sites with ≥ 1 major or critical finding not already identified through central monitoring or a previous visit.

	ADAMON (translated from German study protocol Brosteanu 2017b)	OPTIMON (Journot 2015)	START (Wyman 2020)	TEMPER (Stenning 2018a)	Knott 2015
	requirements to be verified and GCP violations to be counted as major ADAMON audit findings. They counted as audit findings only if they still persisted at the time of auditing. GCP violations remedied by appropriate monitoring follow-up actions were not counted.	GCP violations generated by the site, not corrected by the CTU in spite of the randomized monitoring strategy, and validated as such by the validation committee.		recurred after initial notification, was raised to critical status. The Consistency of Monitoring Group (CMG) comprised the Trial Manager or Data Manager(s) (or both) of the trials that take part in the study, the TSMs, and the Clinical Project Manager. The group met 3-monthly to discuss the monitoring findings and reach consensus in consistency in the grading of the findings.	
Informed consent	Informed consent either not available or contains errors (not signed, not dated, date of consent after inclusion of participant). Violation of safety-relevant or effectiveness-relevant eligibility criteria.	**Non-compliance of the participant's consent form** for whatever reason: The consent form could not be found on site; The participant's name was illegible or absent; The participant's signature was missing;	Informed consent violations were **initially** defined as: -Study-specific procedures performed or participant randomized prior to signing the appropriate IRB/ethics committee-approved consent; -Study-specific procedures performed prior to signing new IRB/ethics committee-	All re-consent (e.g. failure to obtain re-consent in a timely manner) Original consent (e.g. missing signatures, missing or incompatible signature dates, incorrect versions used).	Not reported.

ADAMON (translated from German study protocol Brosteanu 2017b)	OPTIMON (Journot 2015)	START (Wyman 2020)	TEMPER (Stenning 2018a)	Knott 2015
	The date of the participant's signature was later than the date at which it should have been signed or it was illegible or absent; 1 of the items that had to be filled in by the investigator was missing or illegible or the date was later than the visit when it was supposed to shown; The name, date, and the participant's signature were visibly not in his/her handwriting.	approved consent (e.g. amendment); -Most recently signed consent not on file; -Signature or date on consent not made by participant or legal representative. The primary outcome component for consent violations was modified in February 2016. For consent prior to randomization: -Participant signed unapproved or incorrect consent or specimens for storage for future research collected prior to obtaining consent. For later consents due to amendments required locally		

	ADAMON (translated from German study protocol Brosteanu 2017b)	OPTIMON (Journot 2015)	START (Wyman 2020)	TEMPER (Stenning 2018a)	Knott 2015
			or by the sponsor: -Participant's signature page was not on file or -Consent form not signed by participant or legal representative.		
Eligibility	1. Approved therapy was altered without urgent medical need. 2. Definition of unacceptable protocol deviation in the therapy of participants documented in the audit manual (e.g. dose deviation, technical deviations during radio therapy).	**Failure to comply with ≥ 1 eligibility criterion** (inclusion or exclusion) **without prior dispensation**. (A request for dispensation was a request, made by the investigator of the investigation site to the methodology and management center, to include a participant for whom an eligibility criterion was not observed.)	Eligibility violations (HIV-negative, lack of 2 CD4+ cell counts > 500 cells/mm^3 within 60 days before randomization, prior ART or interleukin-2 use, or pregnancy).	Source/priority data discrepancy.	Not reported.

	ADAMON (translated from German study protocol Brosteanu 2017b)	OPTIMON (Journot 2015)	START (Wyman 2020)	TEMPER (Stenning 2018a)	Knott 2015
SAE	An SAE was: Not reported; Reported late according to the study protocol; Reported incompletely without timely follow-up; or Reported without enough precision. In clinical studies involving medical compounds without a clear safety profile for the indication of interest, adverse events should be considered in the assessment of monitoring findings.	**Serious or unexpected adverse event not declared** in a way which complied with the regulations in force, while it has been known to the investigator for > 48 hours.	START serious clinical event (grade 4 event or unscheduled hospitalization) not reported within 6 months from occurrence.	Unreported SAE/notable event.	Not reported.
Endpoint	The primary endpoint of the study was: Not collected; Not collected at the required time point	**Value missing for the main judgment criterion** (possibly calculated on part of the monitoring period: see comment 3, section 5 eligibility criteria), whatever the reason, including not	START primary clinical event not reported within 6 months from occurrence (all potential primary endpoints were counted irrespective of later Endpoint Review Committee review).	Unreported endpoint.	Not reported.

	ADAMON (translated from German study protocol Brosteanu 2017b)	OPTIMON (Journot 2015)	START (Wyman 2020)	TEMPER (Stenning 2018a)	Knott 2015
	(protocol deviation); Collected incorrectly or incompletely. (Timely and methodological deviations considered as major in the collection of the primary endpoint were documented in the study-specific audit manual.)	updating a survival criterion. Each file was reviewed by the OPTIMON validation committee (see section 10.4), which confirmed and documented the error without knowing the monitoring strategy applied.			
Interventi on	Observation and follow-up were altered without urgent medical need. Definitions of unacceptable protocol deviation in the observation or follow-up phase were documented in the study-specific audit manual (e.g. unacceptable in terms of validity of study results).	—	Use of ART for initial therapy that was not permitted by START.	—	Not reported.
Others	—	—	—	1. Pharmacy document and facilities. 2. Investigator site files. 3. Source/priority data discrepancy.	Not reported.

2 Method characteristics of monitoring strategies

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
ADAMON (Brosteanu 2017a)	The classification was based on the 3 Components: - The potential risk of the therapeutic intervention evaluated in the trial as compared to standard medical care; - The presence of ≥ 1 of a list of risk indicators for the participant or the trial results; - The robustness of trial procedures (reliable and easy to assess primary endpoint, simple trial procedures). **K1 highest risk – K3 lowest risk**	**K1:** prestudy visit and initiation visit; existence informed consent and all further key data for 100% of participants; 100% SDV was made for 10% of the site's participants, but ≥ 1 participant. *Frequency of on-site visits:* depending on the site's recruitment and the catalogue of monitoring tasks (in general > 6 per year). **K2:** *trial site with noticeable problems:* existence and informed consent for all participants. Further key data for ≥ 50% of the site's participants. *Trial site without noticeable problems:* existence and informed consent for all participants. Further key data for ≥ 20% of the site's participants. All sites: a 100% SDV is made for 1 participant in the site's random sample (to ascertain any	Central monitoring activities: -Statistical monitoring with multivariate analysis, structured telephone interviews, site status in terms of participant numbers (number of included participants, number lost to follow-up, screening failures etc.); -Problems that would have triggered an additional on-site visit as stated in the study protocol included high or low rate of SAEs or late reporting, protocol deviations (procedures), protocol deviations (eligibility, e.g. threshold of relevant laboratory values exceeded), data inconsistencies in comparison to other sites, outstanding study specific documentation (> 50% expected), high data query rate or suspected fraud. (ADAMON study protocol 2008)	Conduct of monitoring was the responsibility of the respective trial sponsor. For each monitoring strategy, disjoint teams of monitors were trained by the ADAMON team. The ADAMON team received the monitoring reports and supervised adherence to the monitoring manuals, providing additional training for monitors if required.

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
		systematic errors). *Frequency of on-site visits:* ≥ 3 per year (sites with problems)/in general ≥ 1 per year (sites without problems) **K3:** for participants recruited so far at the trial site: existence and informed consent for all participants. Further key data for ≥ 20% of the site's participants. *Frequency of on-site visits:* 1 visit at each trial site. If problems or irregularities that exceeded a trial specific predefined tolerance limit were detected at a trial site, a prompt unplanned on-site monitoring visit was made. (Brosteanu 2009)		
OPTIMON (**Journot 2015**)	Classification based on patient risk evaluation (the therapeutic intervention evaluated in the trial as compared to standard medical care –	**Risk level A:** no on-site visit was planned. Remote management of correction requests. Site closure by letter. **Risk level B:** 1 on-site visit, with	- Exhaustive computerized controls on all data from all participants in all investigation sites entered to check their completeness and consistency. - Investigator requests for clarification	Monitors were from the clinical research centers managing the trials; the monitoring

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
	> intermediate risk); and identifying parameters of the intervention or procedures increasing the risk. 1. *At risk procedures* (e.g. risk of mortality or severe morbidity attributable to the procedure). 2. *At-risk investigations* (e.g. use of a radioactive or a relatively undocumented product or product that had not been authorized). 3. *Target population status aggravating risks attributable to the procedure or interventions* (e.g. risk of	verification of 100% of key data was carried out for 10% of participants. Corrections: during each visit concerning key points. Site closure by letter. **Risk level C:** 1 on-site visit, with verification of 100% of key information was carried out for each site on a percentage of participants corresponding to 1 day of monitoring. Corrections: during each visit concerning key points. On-site closure visit. **Risk level A–C:** setting up: before including the first participant. 1. If the investigation site is known and experienced: by telephone. 2. If the investigation site is not known of or not experienced: on-site visit. Consent: blinded copy of the consent form upon inclusion and on-site during	or correction of any inconsistent data. - Regular contact by telephone, fax, or e-mail with the key people in the investigation site to ensure that procedures are observed, and a standardized contact form completed. - Standard operating procedures, in particular for monitoring studies. The following aspects are particularly harmonized. 1. Compiling the protocol and observation file. 2. The form of the information leaflet and consent form. 3. Notification of inclusions and monitoring the rhythm of inclusions. 4. The project team meeting with a predefined agenda, examination of warning signals and taking corrective action. 5. Computer checks, after entry, of 100% of data. 6. Management of error	outcome was validated by a blinded validation committee.

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
	mortality or severe morbidity attributable to a serious pathologic condition or the participant's age, age ≤ 2 ≥ years, age ≥ 80 years, pregnant, parturient, or breastfeeding women). **Lowest risk level A to highest level D**	the following visit or upon site closure. SAE reporting: systematically on-site or remotely. **Risk level D:** full on-site monitoring. Major problems will trigger an additional on-site visit for levels B and C. (Major problem defined as: endangering participant safety [e.g. at-risk intervention/investigation outside the protocol, inclusion of a participant who does not comply with an eligibility criterion]; endangering the quality of results [e.g. allocation of the randomization treatment, unblinding]; endangering participant's rights [e.g. consent, anonymity]; regulatory aspects [e.g. undeclared investigator].)	correction requests. Consent form: the consent form has an additional sheet with a part blinded at the places for the surname and first name of the participant and his/her signature. This sheet must have been faxed to the methodology and management center on pre-inclusion of the participant.	
START (**Wyman 2020**)	No initial risk assessment or triggers, 1 large international study; sites randomized to local.	**Local monitoring:** twice yearly, clinical site staff associated with START carried out specific quality assurance activities and reported findings to the statistical center. - Regulatory files, including informed consent documents for each version of	Central monitoring included regular review of: - Missing data (e.g. missed visits or individual data items); - Timeliness of data submission and	Central monitoring was performed by the statistical center utilizing data in the central database on a continuous basis.

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
		the START protocol. - Study specimen storage and labeling (if specimens were stored or processed [or both] on-site) - Study drug management and accountability (if the site utilized the START central drug repository). - Verified the source documents for eligibility criteria, informed consent, changes in ART, follow-up visits, and reportable START clinical events for a sample of participants (participant charts were prioritized for source document verification if any of the following had occurred since the previous review: 1. START clinical event reported; 2. Participant became newly lost to follow-up or withdrew from the study; 3. Participant transferred from 1 site to another; 4. Participant was previously identified as lost to follow-up	query resolution; data queries; - Discrepancies between specimens stored at the central repository and specimens collected by site as reported on CRFs for each study visit; - Losses to follow-up and withdrawals of consent; - Findings on daily computer edit checks (largely deterministic) that flagged inadmissible values for single items and combinations of items on case report forms (updated regularly (daily, weekly, or monthly). - Review of data summarizing each site's performance every 6 months and provided quantitative feedback to clinical sites on study performance: participant retention, data quality, timeliness, and completeness of START endpoint documentation, and adherence to local monitoring requirements. Trained nurses at the statistical center reviewed grade 4 events and	On-site monitoring of START was performed annually by a coordinating center-designated monitor, whom were either coordinating center staff or staff located in the country of the sites being monitored.

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
		and was still lost.)	unscheduled hospitalizations for possible primary START clinical events and asked sites to submit the appropriate documentation if a possible START primary endpoint was identified.	
MONITORING (**Fougerou-Leurent 2019**)	Key data identified prior to the monitoring intervention (no full risk assessment) The regulatory or scientific key data (or both) verified by the targeted SDV were: informed consent, inclusion and exclusion criteria, main prognostic variables at inclusion (chosen with the principal investigator), primary endpoint, SAEs.	Targeted SDV in which only regulatory or scientific key data (or both) were verified. Cumulative monitoring time on-site reported 140 hours (vs 317 hours for full on-site monitoring).	No central monitoring performed.	A single experienced clinical researcher. A team from the University Hospital Rennes.
Mealer 2013	No initial risk assessment or triggers of monitoring (participants due for an upcoming on-site visit were checked remotely before the on-site visit)	No on-site visit in the intervention group, only remote access. Participants were assigned to having remote SDV performed 2–4 weeks prior to a scheduled on-site visit – 100% remote SDV for 16 participants.	Remote SDV - Validated the data elements captured on case report forms submitted to the coordinating center using the same data verification protocols that were used during on-site visits.	Monitors were from the clinical (ARDS)/data (ChiLDReN) coordinating centers.

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
		Using a time diary that recorded start/stop time intervals, the total Time required for the study monitor to verify a case report form was captured: adult network: 4.60 (SD 1.42) min with no on-site vs 3.60 (SD 0.96) min with on-site (P = 0.10); pediatric: 11.64 (SD 7.54) min with no on-site vs 6.07 (SD 3.18) min with on-site (P = 0.10).	- Remote monitors had telephone access to the same local coordinators that were available during on-site monitoring visits. - To assess the ability of a monitor to verify the data value that was recorded on the study case report form, 6 possible verification outcome states were defined (found-match, found-different, missing, unknown, found match after coordinator query, not monitored). - 'Found-match after coordinator query' represented the case where remote access was insufficient to find a data value that was found during the subsequent on-site inspection.	
Lièvard 2006	No initial risk assessment; however, study was terminated to prioritize certain sites for site initiation visits.	No on-site initiation visit.	—	Monitoring was organized by the International Drug Development Institute.
TEMPER (**Stenning 2018b**)	On-site visits were triggered by the evaluation of trigger scores. Automatic and manual trigger:	Monitoring usually included SDV on a sample of participants and review of consent forms, pharmacy documents and facilities, and investigator site files. The median number of participants	The software system TEMPER-MS was developed in-house at MRC CTU. It comprises a web application developed in ASP.NET web forms, an SQL server database, which stored the	Triggered visits were attended by TEMPER-specific and trial-specific monitors, untriggered visits

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
	SAE rate (high); SAE rate (low); data query rate (specific question); data query rate (overall); data query resolution time; return rate, specific CRF; overall CRF return rate; protocol deviation (eligibility); protocol deviation (withdrawal rate); protocol deviation (treatment); protocol deviation (procedure); general concern; return rate, patient consent form. Triggers listed with abridged narrative in Diaz-Montana et al. (2019). Highly recruiting sites were selected for triggered visits without matching.	undergoing SDV was 4 (IQR 3–5) with triggered vs 4 (IQR 3–5) with untriggered (paired t-test P = 0.08). The frequency of on-site visits was dependent on the evaluation of the trigger site scores in the trigger meetings held 3–6 monthly with the TEMPER team to choose triggered sites for monitoring.	data, generated for TEMPER, reports developed in SQL server reporting services, and data entry screens for collecting monitoring visit data. A data extraction process was run in TEMPER-MS 8data retrieval from the trial database; aggregation per site; further processing to produce trigger data; evaluation of inequality rules (e.g. > 1% of the fields available for data entry were missing or queried: total number of fields available for data entry that were missed or queried/total number of fields available for data entry P > 0.01)) After extraction, a trigger data report was generated and used in the trigger meeting to guide the prioritization of triggered sites. Trigger types included overall CRF return rate, return rate-specific CRF, return rate participant consent form, data query rate (overall), data query rate (specific question), data query resolution time, SAE rate (high), SAE rate (low), protocol deviation (treatment), protocol deviation	only by TEMPER monitors. The same GCP and monitoring training was undertaken both by the trial team members attending visits and the monitors; the latter also received trial-specific training.

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
			(eligibility), protocol deviation (procedure), protocol deviation (withdrawal rate), high recruitment, general concern. 1.The inequality rule was evaluated as either 'true' or 'false' (i.e. is the rule met?). 2. Automatic triggers sometimes had preconditions in their narrative (e.g. an inequality rule might be evaluated only if there were a minimum number of registered participants at the site). 3. Each trigger had an associated weight (default = 1) specifying its importance relative to other triggers. 4. A site score was obtained for each site as the summation of all scores associated with the site. 5. The trigger data report generated for the trigger meeting listed sites sorted by their site score. 6. Some triggers were designed to fire only when their rule was met at	

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
			consecutive trigger meetings (to distinguish sites that were not improving over time from those with temporary problems). 7. The thresholds were based on trial team experience and also considered the time point in the trial progress. For some triggers preconditions (e.g. a minimum number of registered participants at the site) must have been met for trigger data to be generated and some triggers fired only when their rule was met at consecutive trigger meetings to distinguish sites that were not improving over time from those with temporary problems.	
Knott 2015	Indicators included in the trigger score were 'duration of study visit' (time data were entered to form complete), computer times of data entry (patterns), 4 dimension of the low-density lipoprotein measurements (different mean, SD between sites), measurement of non-compliance	In site visits at high scoring sites resembled an extensive on-site visit and in addition directed monitoring on-site based on information from central statistical monitoring (2-day visit).	- All sites of the multicenter international trial received central statistical monitoring that identified high scoring sites as priority for further investigation. - Scoring was applied every 6 months and a following meeting of the central statistical group. - Scores where either 0 or 1, some indicators had thresholds that when exceeded automatically led to a score	1. The central statistical monitoring group, including the chief investigator, chief statistician, and junior statistician, head of trial

Study	Risk assessment characteristics/triggers or thresholds that induce on-site monitoring	On-site monitoring in the intervention group	Central or remote monitoring in the intervention group	People performing the monitoring
	(participant recorded as no longer taking study medication across sites), SAE reporting (reporting times lower than half the median of all sites), percentage of participants reporting muscle symptoms (dropped later), frequency of updates in non-study medication. Fired triggers resulted in a score of 1 and high scoring sites were chosen for a monitoring visit in the triggered intervention group.		of 1. - Indicators included in the trigger score were 'duration of study visit' (time data were entered to form complete), computer times of data entry (patterns), 4 dimension of the low-density lipoprotein measurements (different mean, SD between sites), measurement of non-compliance (participant recorded as no longer taking study medication across sites), SAE reporting (reporting times lower than half the median of all sites), percentage of participants reporting muscle symptoms (dropped later), frequency of updates in non-study medication.	monitoring assessed high scoring sites and discussed trigger adjustments. 2. Monitoring on-site was performed by the head of trial monitoring.

References to studies

Included studies

Brosteanu 2017b

Brosteanu O, Houben P, Ihrig K, Ohmann C, Paulus U, Pfistner B, et al. Risk analysis and risk adapted on-site monitoring in non-commercial clinical trials. Clinical Trials 2009;6:585-96. [CRSREF: 18579046]

*** Brosteanu O, Schwarz G, Houben P, Paulus U, Strenge-Hesse A, Zettelmeyer U, et al. Risk-adapted monitoring is not inferior to extensive on-site monitoring: results of the ADAMON cluster-randomised study. Clinical Trials 2017;14:584-96. [CRSREF: 18579047]**

Study protocol ("Prospektive cluster-randomisierte Untersuchung studienspezifisch adaptierterStrategien für das Monitoring vor Ort in Kombination mit zusätzlichenqualitätssichernden Maßnahmen"). www.tmf-ev.de/ADAMON/Downloads.aspx (accessed prior to 19 August 2021). [CRSREF: 18579048]

Fougerou-Leurent 2019

Fougerou-Leurent C, Laviolle B, Bellissant E. Cost-effectiveness of full versus targeted monitoring of randomized controlled trials. Fundamental & Clinical Pharmacology 2018;32(S1):49 (PM2 - 035). [CRSREF: 16238180]

*** Fougerou-Leurent C, Laviolle B, Tual C, Visseiche V, Veislinger A, Danjou H, et al. Impact of a targeted monitoring on data-quality and data-management workload of randomized controlled trials: a prospective comparative study. British Journal of Clinical Pharmacology 2019;85(12):2784-92. [CRSREF: 16238181; DOI: 10.1111/bcp.14108]**

Journot 2017

Journot V, Perusat-Villetorte S, Bouyssou C, Couffin-Cadiergues S, Tall A, Chene G. Remote preenrollment checking of consent forms to reduce nonconformity. Clinical Trials 2013;10:449-59. [CRSREF: 16238183]

Journot V, Pignon JP, Gaultier C, Daurat V, Bouxin-Metro A, Giraudeau B, et al. Validation of a risk-assessment scale and a risk-adapted monitoring plan for academic clinical research studies – the Pre-Optimon study. Contemporary Clinical Trials 2011;32:16-24. [CRSREF: 16238184]

Journot V. OPTIMON – first results of the French trial on optimisation of monitoring. ssl2.isped.u-bordeaux2.fr/OPTIMON/docs/Communications/2015-Montpellier/OPTIMON%20-%20EpiClin%20Montpellier%202015-05-20%20EN.pdf (accessed 2 October 2019). [CRSREF: 16238185]

*** Journot V. OPTIMON – the French trial on optimization of monitoring. SCT Annual Meeting; 2017 May 7-10; Liverpool, UK. [CRSREF: 16238186]**

Study protocol: evaluation of the efficacy and cost of two monitoring strategies for public clinical research. OPTIMON study: OPTImisation of MONitoring. ssl2.isped.u-bordeaux2.fr/OPTIMON/DOCS/OPTIMON%20-%20Protocol%20v12.0%20EN%202008-04-21.pdf (accessed prior to 19 August 2021). [CRSREF: 16238187]

Knott 2015

Knott C, Valdes-Marquez E, Landray M, Armitage J, Hopewell J. Improving efficiency of on-site monitoring in multicentre clinical trials by targeting visits. Trials 2015;16(Suppl 2):O49. [CRSREF: 16238189]

Lièнard 2006

Liénard JL, Quinaux E, Fabre-Guillevin E, Piedbois P, Jouhaud A, Decoster G, et al. Impact of on-site initiation visits on patient recruitment and data quality in a randomized trial of adjuvant chemotherapy for breast cancer. Clinical Trials 2006;3(5):486-92. [CRSREF: 16238191; DOI: 10.1177/1740774506070807]

Mealer 2013

Mealer M, Kittelson J, Thompson BT, Wheeler AP, Magee JC, Sokol RJ, et al. Remote source document verification in two national clinical trials networks: a pilot study. PloS One 2013;8(12):e81890. [CRSREF: 16238193]

Stenning 2018b

Cragg WJ, Hurley C, Yorke-Edwards V, Stenning SP. Assessing the potential for prevention or earlier detection of on-site monitoring findings from randomised controlled trials: further analyses of findings from the prospective TEMPER triggered monitoring study. Clinical Trials 2021;18(1):115-26. [CRSREF: 18579050; DOI: 10.1177/1740774520972650]

Diaz-Montana C, Choudhury R, Cragg W, Joffe N, Tappenden N, Sydes MR, et al. Managing our TEMPER: monitoring triggers and site matching algorithms for defining triggered and control sites in the temper study. Trials 2017;18:P149. [CRSREF: 18579051]

Diaz-Montana C, Cragg WJ, Choudhury R, Joffe N, Sydes MR, Stenning SP. Implementing monitoring triggers and matching of triggered and control sites in the TEMPER study: a description and evaluation of a triggered monitoring management system. Trials 2019;20:227. [CRSREF: 18579052]

*** Stenning SP, Cragg WJ, Joffe N, Diaz-Montana C, Choudhury R, Sydes MR, et al. Triggered or routine site monitoring visits for randomised controlled trials: results of TEMPER, a prospective, matched-pair study. Clinical Trials 2018;15:600-9. [CRSREF: 18579053]**

Study protocol: TEMPER (TargetEd Monitoring: Prospective Evaluation and Refinement) prospective evaluation and refinement of a targeted on-site monitoring strategy for multicentre cancer clinical trials. journals.sagepub.com/doi/suppl/10.1177/1740774518793379/suppl_file/793379_supp_mat_2.pdf (accessed prior to 19 August 2021). [CRSREF: 18579054]

Wyman 2020

Hullsiek KH, Kagan JM, Engen N, Grarup J, Hudson F, Denning ET, et al. Investigating the efficacy of clinical trial monitoring strategies: design and implementation of the cluster randomized START monitoring substudy. Therapeutic Innovation and Regulatory Science 2015;49:225-33. [CRSREF: 18579056]

*** Wyman Engen N, Huppler Hullsiek K, Belloso WH, Finley E, Hudson F, Denning E, et al. A randomized evaluation of on-site monitoring nested in a multinational randomized trial. Clinical Trials 2020;17(1):3-14. [CRSREF: 18579057; DOI: 10.1177/1740774519881616]**

***** *Main study publication*

Excluded studies

Agrafiotis 2018

Agrafiotis DK, Lobanov VS, Farnum MA, Yang E, Ciervo J, Walega M, et al. Risk-based monitoring of clinical trials: an integrative approach. Clinical Therapeutics 2018;40:1204-12. [CRSREF: 16238204]

Andersen 2015

Andersen JR, Byrjalsen I, Bihlet A, Kalakou F, Hoeck HC, Hansen G, et al. Impact of source data verification on data quality in clinical trials: an empirical post hoc analysis of three phase 3 randomized clinical trials. British Journal of Clinical Pharmacology 2015;79:660-8. [CRSREF: 16238206]

Bailey 2017

Bailey L, Straw FK, George SE. Implementing a risk based monitoring approach in the early phase myeloma portfolio at Leeds CTRU. Trials 2017;18:220. [CRSREF: 16238208]

Bakobaki 2011

Bakobaki J, Rauchenberger M, Kaganson N, McCormack S, Stenning S, Meredith S. The potential for central monitoring techniques to replace on-site monitoring in clinical trials: a review of monitoring findings from an international multi-centre clinical trial. Clinical Trials 2011;8:454-5. [CRSREF: 16238210]

Bakobaki 2012

Bakobaki JM, Rauchenberger M, Joffe N, McCormack S, Stenning S, Meredith S. The potential for central monitoring techniques to replace on-site monitoring: findings from an international multi-centre clinical trial. Clinical Trials 2012;9:257-64. [CRSREF: 16238212]

Biglan 2016

Biglan K, Brocht A, Raca P. Implementing risk-based monitoring (RBM) in STEADY-PD III, a phase III multi-site clinical drug trial for Parkinson disease. Movement Disorders 2016;31(9):E10. [CRSREF: 16238214]

Collett 2019

Collett L, Gidman E, Rogers C. Automation of clinical trial statistical monitoring. Trials 2019;20(Suppl 1):P-251. [CRSREF: 17860912]

Cragg 2019

Cragg WJ, Cafferty F, Diaz-Montana C, James EC, Joffe J, Mascarenhas M, et al. Early warnings and repayment plans: novel trial management methods for monitoring and managing data return rates in a multi-centre phase III randomised controlled trial with paper case report forms. Trials 2019;20:241. [CRSREF: 16238216; DOI: 10.1186/s13063-019-3343-2]

Del Alamo 2018

Del Alamo M, Sanchez AI, Serrano ML, Aguilar M, Arcas M, Alvarez A, et al. Monitoring strategies for clinical trials in primary care: an independent clinical research perspective. Basic & Clinical Pharmacology & Toxicology 2018;123:25-6. [CRSREF: 16238218]

Diani 2017

Diani CA, Rock A, Moll P. An evaluation of the effectiveness of a risk-based monitoring approach implemented with clinical trials involving implantable cardiac medical devices. Clinical Trials 2017;14:575-83. [CRSREF: 16238220]

Diaz-Montana 2019b

Diaz-Montana C, Masters L, Love SB, Lensen S, Yorke-Edwards V, Sydes MR. Making performance metrics work: developing a triggered monitoring management system. Trials 2019;20(Suppl 1):P-63. [CRSREF: 18579059]

Edwards 2014

Edwards P, Shakur H, Barnetson L, Prieto D, Evans S, Roberts I. Central and statistical data monitoring in the Clinical Randomisation of an Antifibrinolytic in Significant Haemorrhage (CRASH-2) trial. Clinical Trials 2014;11:336-43. [CRSREF: 16238222]

Elsa 2011

Elsa VM, Jemma HC, Martin L, Jane A. A key risk indicator approach to central statistical monitoring in multicentre clinical trials: method development in the context of an ongoing large-scale randomized trial. Trials 2011;12:A135. [CRSREF: 16238224]

Fu 2021

Fu ZY, Liu XH, Zhao SH, Yuan YN, Jiang M. A preliminary analysis of remote monitoring practice in clinical trials. Chinese Journal of New Drugs 2021;30(3):209-14. [CRSREF: 17860916]

Hatayama 2020

Hatayama T, Yasui S. Bayesian central statistical monitoring using finite mixture models in multicenter clinical trials. Contemporary Clinical Trials Communication 2020;19:100566. [CRSREF: 17860918]

Heels-Ansdell 2010

Heels-Ansdell D, Walter S, Zytaruk N, Guyatt G, Crowther M, Warkentin T, et al. Central statistical monitoring of an international thromboprophylaxis trial. American Journal of Respiratory and Critical Care Medicine 2010;181:A6041. [CRSREF: 16238226]

Higa 2020

Higa A, Yagi M, Hayashi K, Kosako M, Akiho H. Risk-based monitoring approach to ensure the quality of clinical study data and enable effective monitoring. Therapeutic Innovation and Regulatory Science 2020;54(1):139-43. [CRSREF: 17860920]

Hirase 2016

Hirase K, Fukuda-Doi M, Okazaki S, Uotani M, Ohara H, Furukawa A, et al. Development of an efficient monitoring method for investigator-initiated clinical trials: lessons from the experience of ATACH-II trial. Japanese Pharmacology and Therapeutics 2016;44:s150-4. [CRSREF: 16238230]

Jones 2019

Jones L, Ogburn E, Yu LM, Begum N, Long A, Hobbs FD. On-site monitoring of primary outcomes is important in primary care clinical trials: Benefits of Aldosterone Receptor Antagonism in Chronic Kidney Disease (BARACK-D) trial – a case study. Trials 2019;20(Suppl 1):P-272. [CRSREF: 17860922]

Jung 2020

Jung HY, Jeon Y, Seong SJ, Seo JJ, Choi JY, Cho JH, et al. Information and communication technology-based centralized monitoring system to increase adherence to immunosuppressive medication in kidney transplant recipients: a randomized controlled trial. Nephrology, Dialysis, Transplantation 2020;35(Suppl 3):gfaa143.P1734. [CRSREF: 17860924; DOI: 10.1093/ndt/gfaa143.P1734]

Kim 2011

Kim J, Zhao W, Pauls K, Goddard T. Integration of site performance monitoring module in web-based CTMS for a global trial. Clinical Trials 2011;8:450. [CRSREF: 16238232]

Kim 2021

Kim S, Kim Y, Hong Y, Kim Y, Lim JS, Lee J, et al. Feasibility of a hybrid risk-adapted monitoring system in investigator-sponsored trials in cancer. Therapeutic Innovation and Regulatory Science 2021;55(1):180-9. [CRSREF: 17860926]

Lane 2013

Lane JA, Wade J, Down L, Bonnington S, Holding PN, Lennon T, et al. A Peer Review Intervention for Monitoring and Evaluating sites (PRIME) that improved randomized controlled trial conduct and performance. Journal of Clinical Epidemiology 2011;64:628-36. [CRSREF: 16238234]

Lane JA. Improving trial quality through a new site monitoring process: experience from the Protect Study. Clinical Trials 2008;5:404. [CRSREF: 16238235]

Lane JJ, Davis M, Down E, Macefield R, Neal D, Hamdy F, et al. Evaluation of source data verification in a multicentre cancer trial (PROTECT). Trials 2013;14:83. [CRSREF: 16238236]

Lim 2017

Lim JY, Hackett M, Munoz-Venturelli P, Arima H, Middleton S, Olavarria VV, et al. Monitoring a large-scale international cluster stroke trial: lessons from head position in stroke trial. Stroke 2017;48:ATP371. [CRSREF: 16238238]

Lindley 2015

Lindley RI. Cost effective central monitoring of clinical trials. Neuroepidemiology 2015;45:303. [CRSREF: 16238240]

Miyamoto 2019

Miyamoto K, Nakamura K, Mizusawa J, de Balincourt C, Fukuda H. Study risk assessment of Japan Clinical Oncology Group (JCOG) clinical trials using the European Organisation for Research and Treatment of Cancer (EORTC) study risk calculator. Japanese Journal of Clinical Oncology 2019;49(8):727-33. [CRSREF: 17860928]

Morales 2020

Morales A, Miropolsky L, Seagal I, Evans K, Romero H, Katz N. Case studies on the use of central statistical monitoring and interventions to optimize data quality in clinical trials. Osteoarthritis and Cartilage 2020;28:S460. [CRSREF: 17860930]

Murphy 2019

Murphy J, Durkina M, Jadav P, Kiru G. An assessment of feasibility and cost-effectiveness of remote monitoring on a multicentre observational study. Trials 2019;20(Suppl 1):P-265. [CRSREF: 17860932]

Pei 2019

Pei XJ, Han L, Wang T. Enhancing the system of expedited reporting of safety data during clinical trials of drugs and strengthening the management of clinical trial risk monitoring. Chinese Journal of New Drugs 2019;28(17):2113-6. [CRSREF: 17860934]

Stock 2017

Stock E, Mi Z, Biswas K, Belitskaya-Levy I. Surveillance of clinical trial performance using centralized statistical monitoring. Trials 2017;18:200. [CRSREF: 16238242]

Sudo 2017

Sudo T, Sato A. Investigation of the factors affecting risk-based quality management of investigator-initiated investigational new-drug trials for unapproved anticancer drugs in Japan. Therapeutic Innovation and Regulatory Science 2017;51:589-96. [CRSREF: 17860936; DOI: 10.1177/2168479017705155]

Thom 1996

Thom E, Das A, Mercer B, McNellis D. Clinical trial monitoring in the face of changing clinical practice. The NICHD MFMU Network. Controlled Clinical Trials 1996;17:58S-59S. [CRSREF: 16238246]

Tudur Smith 2012b

Tudur Smith C, Stocken DD, Dunn J, Cox T, Ghaneh P, Cunningham D, et al. The value of source data verification in a cancer clinical trial. PloS One 2012;7(12):e51623. [CRSREF: 18579061]

von Niederhäusern 2017

von Niederhäusern B, Orleth A, Schädelin S, Rawi N, Velkopolszky M, Becherer C, et al. Generating evidence on a risk-based monitoring approach in the academic setting – lessons learned. BMC Medical Research Methodology 2017;17:26. [CRSREF: 16238250]

Yamada 2021

Yamada O, Chiu SW, Takata M, Abe M, Shoji M, Kyotani E, et al. Clinical trial monitoring effectiveness: remote risk-based monitoring versus on-site monitoring with 100% source data verification. Clinical Trials (London, England) 2021;18(2):158-67. [CRSREF: 18579063; DOI: 10.1177/1740774520971254; PubMed: 33258688]

Yorke-Edwards 2019

Yorke-Edwards VE, Diaz-Montana C, Mavridou K, Lensen S, Sydes MR, Love SB. Risk-based trial monitoring: site performance metrics across time. Trials 2019;20(Suppl 1):P-33. [CRSREF: 17860940]

Zhao 2013

Zhao W. Risk-based monitoring approach in practice-combination of real-time central monitoring and on-site source document verification. Clinical Trials 2013;10:S4. [CRSREF: 16238252]

Additional references

ADAMON study protocol 2008

ADAMON study protocol. Study protocol ("Prospektive cluster-randomisierte Untersuchung studienspezifisch adaptierterStrategien für das Monitoring vor Ort in Kombination mit zusätzlichenqualitätssichernden Maßnahmen"). www.tmf-ev.de/ADAMON/Downloads.aspx (accessed prior to 19 August 2021).

Anon 2012

Anon. Education section: Studies Within A Trial (SWAT). Journal of Evidence-based Medicine 2012;5:44-5.

Baigent 2008

Baigent C, Harrell FE, Buyse M, Emberson JR, Altman DG. Ensuring trial validity by data quality assurance and diversification of monitoring methods. Clinical Trials 2008;5:49-55.

Bensaaud 2020

Bensaaud A, Gibson I, Jones J, Flaherty G, Sultan S, Tawfick W, et al. A telephone reminder to enhance adherence to interventions in cardiovascular randomized trials: a protocol for a Study Within A Trial (SWAT). Journal of Evidence-based Medicine 2020;13(1):81-4. [DOI: 10.1111/jebm.12375]

Brosteanu 2009

Brosteanu O, Houben P, Ihrig K, Ohmann C, Paulus U, Pfistner B, et al. Risk analysis and risk adapted on-site monitoring in noncommercial clinical trials. Clinical Trials 2009;6:585-96.

Brosteanu 2017a

Brosteanu O, Schwarz G, Houben P, Paulus U, Strenge-Hesse A, Zettelmeyer U, et al. Risk-adapted monitoring is not inferior to extensive on-site monitoring: results of the ADAMON cluster-randomised study. Clinical Trials 2017;14:584-96.

Buyse 2020

Buyse M, Trotta L, Saad ED, Sakamoto J. Central statistical monitoring of investigator-led clinical trials in oncology. International Journal of Clinical Oncology 2020;25(7):1207-14.

Chene 2008

Chene G. Evaluation of the efficacy and cost of two monitoring strategies for public clinical research. OPTIMON study: OPTImisation of MONitoring. ssl2.isped.u-bordeaux2.fr/OPTIMON/DOCS/OPTIMON%20-%20Protocol%20v12.0%20EN%202008-04-21.pdf (accessed 2 October 2019).

Cragg 2021a

Cragg WJ, Hurley C, Yorke-Edwards V, Stenning SP. Assessing the potential for prevention or earlier detection of on-site monitoring findings from randomised controlled trials: further analyses of findings from the prospective TEMPER triggered monitoring study. Clinical Trials 2021;18(1):115-26. [DOI: 10.1177/1740774520972650; PubMed: 33231127]

Cragg 2021b

Cragg WJ, Hurley C, Yorke-Edwards V, Stenning SP. Dynamic methods for ongoing assessment of site-level risk in risk-based monitoring of clinical trials: a scoping review. Clinical Trials 2021;18(2):245-59. [DOI: 10.1177/1740774520976561; PubMed: 33611927]

DerSimonian 1986

DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials 1986;7(3):177-88.

Diaz-Montana 2019a

Diaz-Montana C, Cragg WJ, Choudhury R, Joffe N, Sydes MR, Stenning SP. Implementing monitoring triggers and matching of triggered and control sites in the TEMPER study: a description and evaluation of a triggered monitoring management system. Trials 2019;20:227.

Duley 2008

Duley L, Antman K, Arena J, Avezum A, Blumenthal M, Bosch J, et al. Specific barriers to the conduct of randomised trials. Clinical Trials 2008;5:40-8.

EC 2014

European Commission. Risk proportionate approaches in clinical trials. Recommendations of the expert group on clinical trials for the implementation of Regulation (EU) No 536/2014 on clinical trials on medicinal products for human use. ec.europa.eu/health/sites/default/files/files/eudralex/vol-10/2017_04_25_risk_proportionate_approaches_in_ct.pdf (accessed 28 July 2021).

EMA 2013

European Medicines Agency. Reflection paper on risk based quality management in clinical trials, 2013.

ema.europa.eu/docs/en_GB/document_library/Scientific_guidelines/2013/11/WC500155491.pdf (accessed 2 July 2021).

EMA 2017

European Medicines Agency. Procedure for reporting of GCP inspections requested by the Committee for Medicinal Products for Human Use, 2017. ema.europa.eu/en/documents/regulatory-procedural-guideline/ins-gcp-4-procedure-reporting-good-clinical-practice-inspections-requested-chmp_en.pdf (accessed 2 July 2021).

EMA 2021

EMA European Medicines Agency. Guidance on the management of clinical trial during the COVID-19 (coronavirus) pandemic. European Medicines Agency 2021;V4(https://ec.europa.eu/health/sites/default/files/files/eudralex/vol-10/guidanceclinicaltrials_covid19_en.pdf (accessed August 2021)).

Embleton-Thirsk 2019

Embleton-Thirsk A, Deane E, Townsend S, Farrelly L, Popoola B, Parker J, et al. Impact of retrospective data verification to prepare the ICON6 trial for use in a marketing authorization application. Clinical Trials 2019;16(5):502-11.

EPOC 2016

Effective Practice Organisation of Care. What study designs should be included in an EPOC review and what should they be called? EPOC resources for review authors, 2016. epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/EPOC%20Study%20Designs%20About.pdf (accessed 2 July 2021).

EPOC 2017

Effective Practice Organisation of Care. Suggested risk of bias criteria for EPOC reviews. EPOC resources for review authors, 2017. epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/Resources-for-authors2017/suggested_risk_of_bias_criteria_for_epoc_reviews.pdf (accessed 2 July 2021).

FDA 2013

US Department of Health and Human Services Food and Drug Administration. Guidance for industry oversight of clinical investigations – a risk-based approach to monitoring. www.fda.gov/downloads/Drugs/Guidances/UCM269919.pdf (accessed 2 July 2021).

FDA 2020

US Food and Drug Administration. FDA guidance on conduct of clinical trials of medical products during COVID-19 public health emergency: guidance for industry, investigators, and institutional review boards, 2020. www.fda.gov/media/136238/download (accessed 19 August 2021).

Funning 2009

Funning S, Grahnén A, Eriksson K, Kettis-Linblad A. Quality assurance within the scope of good clinical practice (GCP) – what is the cost of GCP-related activities? A survey within the

Swedish Association of the Pharmaceutical Industry (LIF)'s members. Quality Assurance Journal 2009;12(1):3-7. [DOI: 10.1002/qaj.433]

Gaba 2020

Gaba P Bhatt DL. The COVID-19 pandemic: a catalyst to improve clinical trials. Nature Reviews. Cardiology 2020;17:673-5.

Gough 2016

Gough J, Wilson B, Zerola M. Defining a central monitoring capability: sharing the experience of TransCelerateBioPharmas approach, part 2. Therapeutic Innovation and Regulatory Science 2016;50(1):8-14. [DOI: 10.1177/2168479015618696; PubMed: 30236019]

GRADEpro GDT

GRADEpro GDT [Computer program]. Version Accessed August 2021. Hamilton (ON): McMaster University (developed by Evidence Prime Inc), 2020. Available at gradepro.org.

Grignolo 2011

Grignolo A. The Clinical Trials Transformation Initiative (CTTI). Annali dell'Istituto Superiore di Sanita 2011;47:14-8. [DOI: 10.4415/ANN_11_01_04; PubMed: 21430332]

Guyatt 2013a

Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, et al. GRADE guidelines: 12. Preparing summary of findings tables – binary outcomes. Journal of Clinical Epidemiology 2013;66:158-72.

Guyatt 2013b

Guyatt GH, Thorlund K, Oxman AD, Walter SD, Patrick D, Furukawa TA, et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles – continuous outcomes. Journal of Clinical Epidemiology 2013;66:173-83.

Hearn 2007

Hearn J, Sullivan R. The impact of the 'Clinical Trials' directive on the cost and conduct of non-commercial cancer trials in the UK. European Journal of Cancer 2007;43:8-13.

Higgins 2016

Higgins JP, Lasserson T, Chandler J, Tovey D, Churchill R. Methodological Expectations of Cochrane Intervention Reviews. London (UK): Cochrane, 2016.

Higgins 2020

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from handbook: training.cochrane.org/handbook/archive/v6.1.

Horsley 2011

Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. Cochrane Database of Systematic Reviews 2011, Issue 8. Art. No.: MR000026. DOI: 10.1002/14651858.MR000026.pub2.

Houghton 2020

Houghton C, Dowling M, Meskell P, Hunter A, Gardner H, Conway A, et al. Factors that impact on recruitment to randomised trials in health care: a qualitative evidence synthesis. Cochrane Database of Systematic Reviews 2020, Issue 10. Art. No.: MR000045. DOI: 10.1002/14651858.MR000045.pub2.

Hullsiek 2015

Hullsiek KH, Kagan JM, Engen N, Grarup J, Hudson F, Denning ET, et al. Investigating the efficacy of clinical trial monitoring strategies: design and implementation of the cluster randomized START monitoring substudy. Therapeutic Innovation and Regulatory Science 2015;49(2):225-33. [DOI: 10.1177/2168479014555912; PubMed: 25973346]

Hurley 2016

Hurley C, Shiely F, Power J, Clarke M, Eustace JA, Flanagan E, et al. Risk based monitoring (RBM) tools for clinical trials: a systematic review. Contemporary Clinical Trials 2016;51:15-27.

ICH 1996

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: guideline for good clinical practice E6 (R2). www.ema.europa.eu/en/documents/scientific-guideline/ich-e-6-r2-guideline-good-clinical-practice-step-5_en.pdf (accessed 28 July 2021).

ICH 2016

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Integrated Addendum to ICH E6(R1): guideline for good clinical practice E6R(2). database.ich.org/sites/default/files/E6_R2_Addendum.pdf (accessed 2 July 2021).

Izmailova 2020

Izmailova ES, Ellis R, Benko C. Remote monitoring in clinical trials during the COVID-19 pandemic. Clinical and Translational Science 2020;13(5):838-41. [DOI: 10.1111/cts.12834]

Journot 2011

Journot V, Pignon JP, Gaultier C, Daurat V, Bouxin-Metro A, Giraudeau B, et al. Validation of a risk-assessment scale and a risk-adapted monitoring plan for academic clinical research studies – the Pre-Optimon study. Contemporary Clinical Trials 2011;32:16-24.

Journot 2013

Journot V, Perusat-Villetorte S, Bouyssou C, Couffin-Cadiergues S, Tall A, Chene G. Remote preenrollment checking of consent forms to reduce nonconformity. Clinical Trials 2013;10:449-59.

Journot 2015

Journot V. OPTIMON – first results of the French trial on optimisation of monitoring. ssl2.isped.u-bordeaux2.fr/OPTIMON/docs/Communications/2015-Montpellier/OPTIMON%20-%20EpiClin%20Montpellier%202015-05-20%20EN.pdf (accessed 28 July 2021).

Landray 2012

Landray MJ, Grandinetti C, Kramer JM, Morrison BW, Ball L, Sherman RE. Clinical trials: rethinking how we ensure quality. Drug Information Journal 2012;46:657-60. [DOI: 10.1177/0092861512464372]

Lefebvre 2011

Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook/archive/v5.1/.

Love 2019

Love SB, Yorke-Edwards V, Lensen S, Sydes MR. Central data monitoring in phase III RCTs in the UKCRC registered clinical trials units: results from a survey. Trials 2019;20(Suppl 1):P-32.

Love 2021

Love SB, Armstrong E, Bayliss C, Boulter M, Fox L, Grumett J, et al. Monitoring advances including consent: learning from COVID-19 trials and other trials running in UKCRC registered clinical trials units during the pandemic. Trials 2021;22:279.

McDermott 2020

McDermott MM, Newman AB. Preserving clinical trial integrity during the coronavirus pandemic. JAMA 2020;323(21):2135-6.

McGowan 2016

McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. Journal of Clinical Epidemiology 2016;75:40-6. [DOI: 10.1016/j.jclinepi.2016.01.021; PubMed: 27005575]

Meredith 2011

Meredith S, Ward M, Booth G, Fisher A, Gamble C, House H, et al. Risk-adapted approaches to the management of clinical trials: guidance from the Department of Health (DH) / Medical Research Council (MRC)/Medicines and Healthcare Products Regulatory Agency (MHRA) Clinical Trials Working Group. Trials 2011;12:A39.

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Journal of Clinical Epidemiology 2009;62:1006-12.

Morrison 2011

Morrison BW, Cochran CJ, White JG, Harley J, Kleppinger CF, Liu A, et al. Monitoring the quality of conduct of clinical trials: a survey of current practices. Clinical Trials 2011;8(3):342-9.

OECD 2013

Organisation for Economic Co-operation and Development. OECD recommendation on the governance of clinical trials. oecd.org/sti/inno/oecdrecommendationonthegovernanceofclinicaltrials.htm (accessed 2 July 2021).

Olsen 2016

Olsen R, Bihlet AR, Kalakou F. The impact of clinical trial monitoring approaches on data integrity and cost? A review of current literature. European Journal of Clinical Pharmacology 2016;72:399-412.

OPTIMON study protocol 2008

OPTIMON study protocol. Study protocol: evaluation of the efficacy and cost of two monitoring strategies for public clinical research. OPTIMON study: OPTImisation of MONitoring. ssl2.isped.u-bordeaux2.fr/OPTIMON/DOCS/OPTIMON%20-%20Protocol%20v12.0%20EN%202008-04-21.pdf (accessed prior to 19 August 2021).

Oxman 1992

Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Annals of Internal Medicine 1992;116:78-84.

Review Manager 2014

Review Manager 5 (RevMan 5) [Computer program]. Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

SCTO 2020

Monitoring Platform of the Swiss Clinical Trial Organisation (SCTO) F dated. Fact sheet: central data monitoring in clinical trials? V 1.0. www.scto.ch/monitoring (accessed 2 July 2021).

Shiely 2021

Shiely F, Foley J, Stone A, Cobbe E, Browne S, Murphy E, et al. Managing clinical trials during COVID-19: experience from a clinical research facility. Trials 2021;22:62.

Stenning 2018a

Stenning SP, Cragg WJ, Joffe N, Diaz-Montana C, Choudhury R, Sydes MR, et al. Triggered or routine site monitoring visits for randomised controlled trials: results of TEMPER, a prospective, matched-pair study. Clinical Trials 2018;15:600-9.

Sun 2010

Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. BMJ 2010;340:c117.

Tantsyura 2015

Tantsyura V, Dunn IM, Fendt K. Risk-based monitoring: a closer statistical look at source document verification, queries, study size effects, and data quality. Therapeutic Innovation and Regulatory Science 2015;49:903-10.

Thomas 2010

EPPI-Reviewer: software for research synthesis. EPPI-Centre Software [Computer program]. Thomas J, Brunton J, Graziosi S. Version 4.0. London (UK): Social Science Research Unit, Institute of Education, University of London, 2010.

TransCelerate BioPharma Inc 2014

TransCelerateBiopharmaInc. Risk-based monitoring methodology. www.transceleratebiopharmainc.com/wp-content/uploads/2016/01/TransCelerate-RBM-Position-Paper-FINAL-30MAY2013.pdf (accessed 28 July 2021).

Treweek 2018a

Treweek S, Bevan S, Bower P, Campbell M, Christie J, Clarke M, et al. Trial Forge Guidance 1: what is a Study Within A Trial (SWAT)? Trials 2018;19:139. [DOI: 10.1186/s13063-018-2535-5]

Treweek 2018b

Treweek S, Pitkethly M, Cook J, Fraser C, Mitchell E, Sullivan F, et al. Strategies to improve recruitment to randomised trials. Cochrane Database of Systematic Reviews 2018, Issue 2. Art. No.: MR000013. DOI: 10.1002/14651858.MR000013.pub6.

Tudur Smith 2012a

Tudur Smith C, Stocken DD, Dunn J, Cox T, Ghaneh P, Cunningham D, et al. The value of source data verification in a cancer clinical trial. PloS One 2012;7:e51623.

Tudur Smith 2014

Tudur Smith C, Williamson P, Jones A, Smyth A, Hewer SL, Gamble C. Risk-proportionate clinical trial monitoring: an example approach from a non-commercial trials unit. Trials 2014;15:127.

Valdés-Márquez 2011

Valdés-Márquez E, Hopewell CJ, Landray M, Armitage J. A key risk indicator approach to central statistical monitoring in multicentre clinical trials: method development in the context of an ongoing large-scale randomized trial. Trials 2011;12(Suppl 1):A135.

Venet 2012

Venet D, Doffagne E, Burzykowski T, Beckers F, Tellier Y, Genevois-Marlin E, et al. A statistical approach to central monitoring of data quality in clinical trials. Clinical Trials 2012;9:705-13.

von Niederhausern 2017

von Niederhausern B, Orleth A, Schadelin S, Rawi N, Velkopolszky M, Becherer C, et al. Generating evidence on a risk-based monitoring approach in the academic setting – lessons learned. BMC Medical Research Methodology 2017;17:26.

Wyman Engen 2020

Wyman Engen N, Huppler Hullsiek K, Belloso WH, Finley E, Hudson F, Denning E, et al. A randomized evaluation of on-site monitoring nested in a multinational randomized trial. Clinical Trials 2020;17(1):3-14. [DOI: 10.1177/1740774519881616; PubMed: 31647325]

Young 2011

Young T, Hopewell S. Methods for obtaining unpublished data. Cochrane Database of Systematic Reviews 2011, Issue 11. Art. No.: MR000027. DOI: 10.1002/14651858.MR000027.pub2.

Other published versions of this review

Klatte 2019

Klatte K, Pauli-Magnus C, Love S, Sydes M, Benkert P, Bruni N, et al. Monitoring strategies for clinical intervention studies [Protocol]. Cochrane Database of Systematic Reviews 2019, Issue 12. Art. No.: MR000051. DOI: 10.1002/14651858.MR000051.

Data and analyses

1 Risk-based versus on-site monitoring – combined primary outcome

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
1.1 Combined outcome of critical and major monitoring findings	2	2377	Risk Ratio (IV, Random, 95% CI)	1.03 [0.81, 1.32]

2 Risk-based versus on-site monitoring – error domains of major findings

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
2.1 Combined outcome of	2	9508	Risk Ratio (IV, Random,	0.95 [0.81, 1.13]

critical and major findings in 4 error domains			95% CI)	
2.1.1 Critical or major finding related to informed consent	2	2377	Risk Ratio (IV, Random, 95% CI)	0.80 [0.63, 1.02]
2.1.2 Critical or major finding related to eligibility	2	2377	Risk Ratio (IV, Random, 95% CI)	1.31 [0.56, 3.07]
2.1.3 Critical or major finding related to endpoint assessment	2	2377	Risk Ratio (IV, Random, 95% CI)	0.91 [0.63, 1.32]
2.1.4 Critical or major finding related to serious adverse effect reporting	2	2377	Risk Ratio (IV, Random, 95% CI)	1.01 [0.83, 1.23]

3 Triggered versus untriggered on-site monitoring

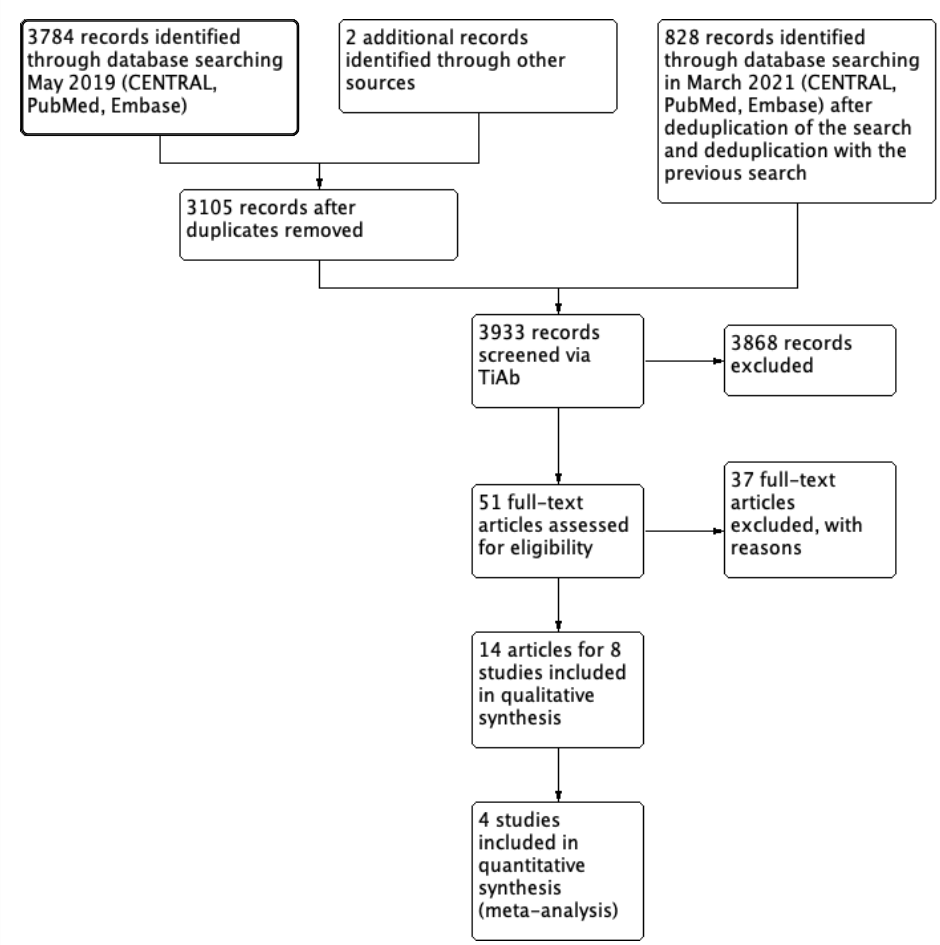
Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
3.1 Sites ≥ 1 major monitoring finding combined outcome	2	105	Risk Ratio (IV, Random, 95% CI)	1.83 [0.51, 6.55]

4 Sensitivity analysis of the comparison: triggered versus untriggered on-site monitoring (sensitivity outcome TEMPER)

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
4.1 Sites ≥ 1 major monitoring finding excluding re-consent	2	105	Risk Ratio (IV, Random, 95% CI)	2.04 [0.77, 5.38]

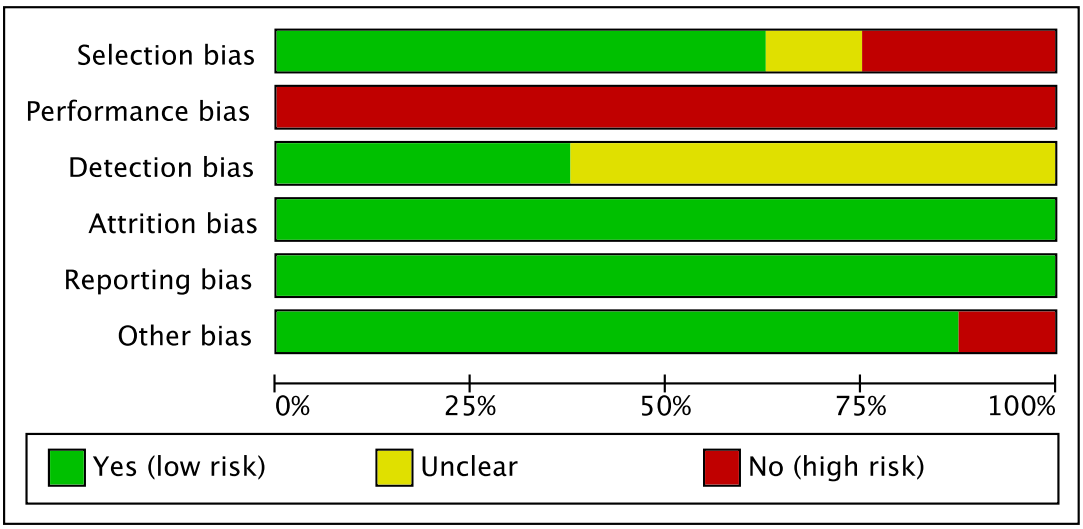
Figures

Figure 1



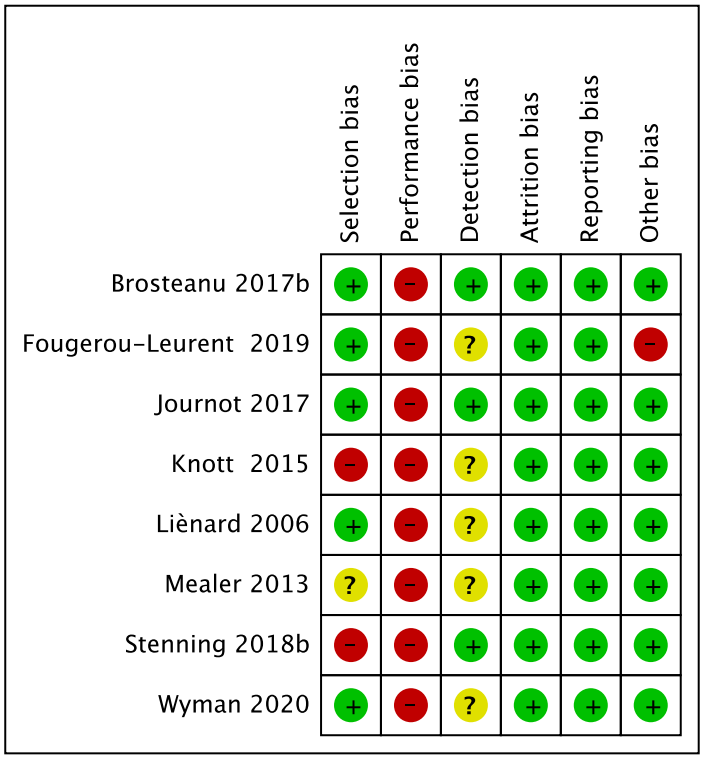
Study flow diagram.

Figure 2



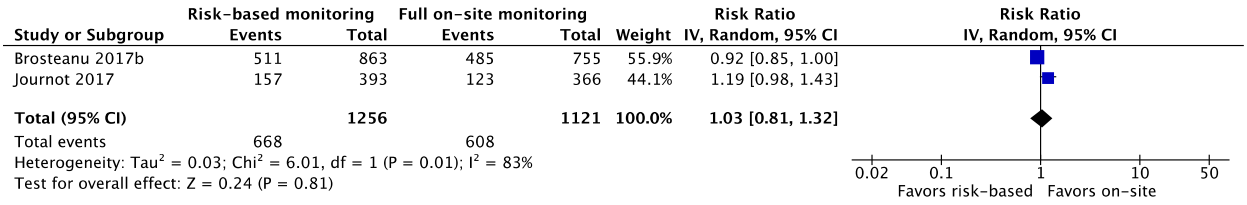
Risk of bias graph: review authors' judgments about each risk of bias item presented as percentages across all included studies.

Figure 3



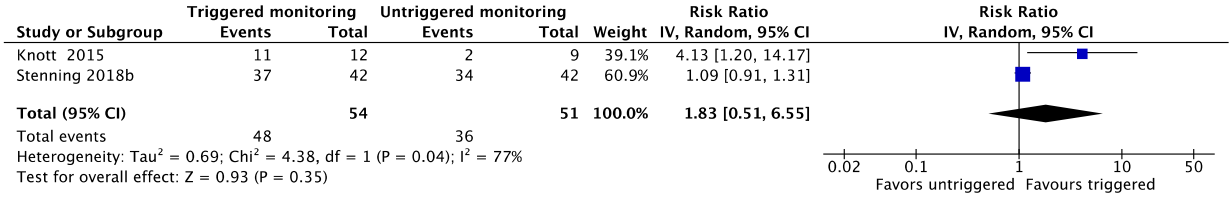
Risk of bias summary: review authors' judgments about each risk of bias item for each included study.

Figure 4 (Analysis 1.1)



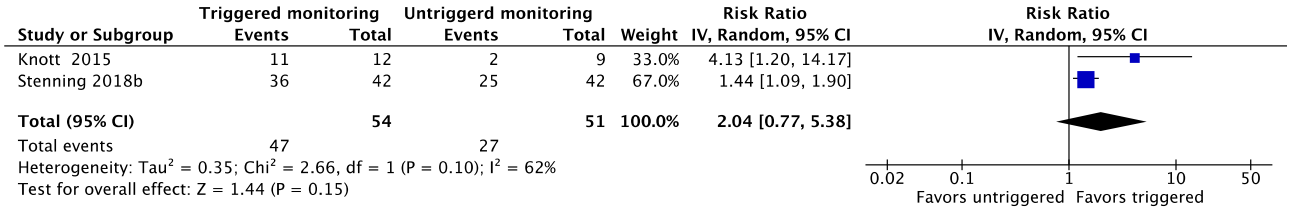
Forest plot of comparison: 1 Risk-based versus on-site monitoring – combined primary outcome, outcome: 1.1 Combined outcome of critical and major monitoring findings.

Figure 5 (Analysis 3.1)



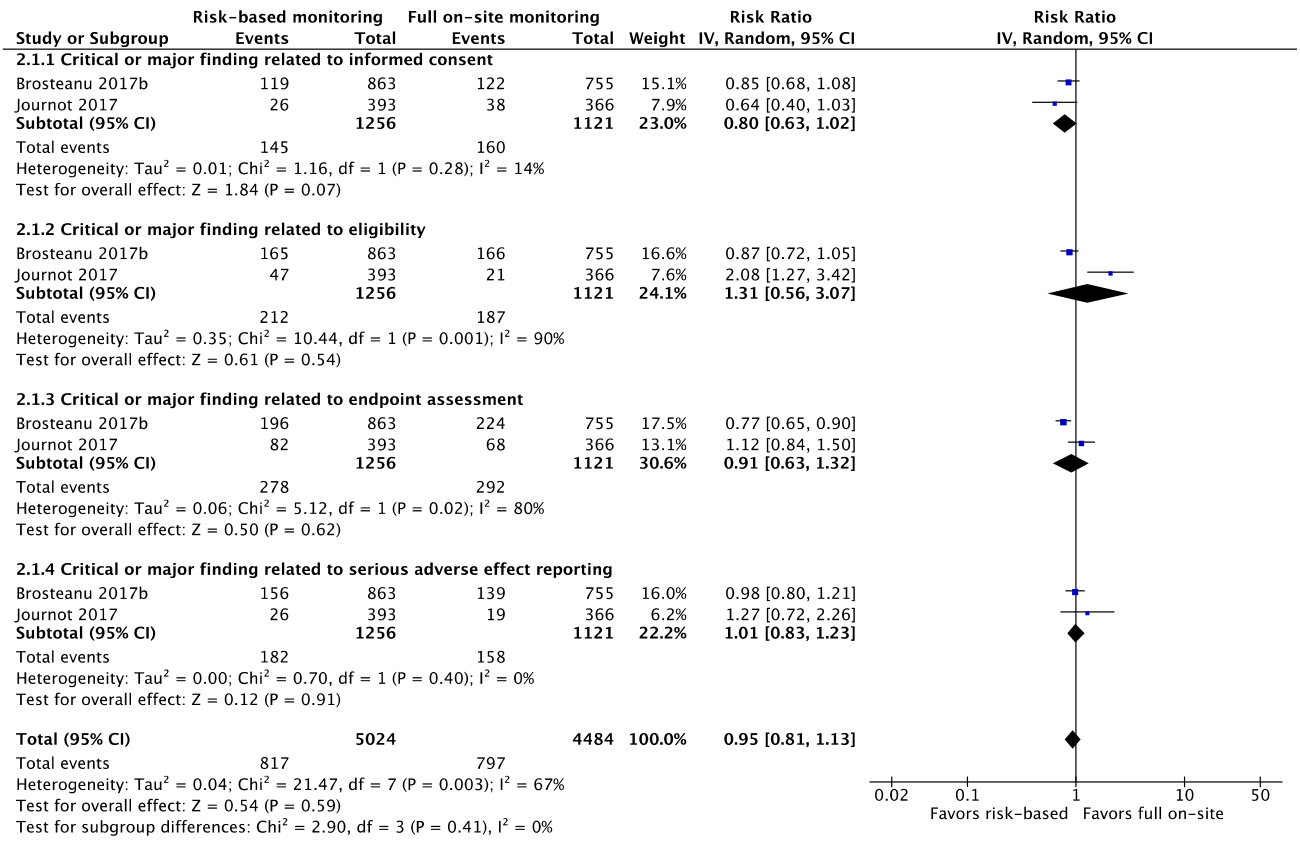
Forest plot of comparison: 3 Triggered versus untriggered on-site monitoring, outcome: 3.1 Sites one or more major monitoring finding combined outcome.

Figure 6 (Analysis 4.1)



Forest plot of comparison: 4 Sensitivity analysis of the comparison: triggered versus untriggered on-site monitoring (sensitivity outcome TEMPER), outcome: 4.1 Sites one or more major monitoring finding excluding re-consent.

Figure 7 (Analysis 2.1)



Forest plot of comparison: 2 Risk-based versus on-site monitoring – error domains of major findings, outcome: 2.1 Combined outcome of major or critical findings in four error domains.

Sources of support

Internal sources

- Department of Clinical Research, Switzerland

  The Department of Clinical Research provided salaries for review contributors.

External sources

- No sources of support provided

Appendices

1 Search strategies CENTRAL, PubMed, and Embase

Cochrane Review on monitoring strategies: search strategies
Terms shown in italics were different compared to the strategy in PubMed.

CENTRAL

3 May 2019: 842 hits (836 trials/6 reviews); Update 16 March 2021: 1044 hits *(monitor* NEAR/2 (site OR risk OR central*))*:ti,ab OR "monitoring strategy":ti,ab OR "monitoring
method":ti,ab OR "monitoring technique":ti,ab OR "triggered monitoring":ti,ab OR "targeted monitoring":ti,ab OR "risk proportionate":ti,ab OR "trial monitoring":ti,ab OR "study monitoring":ti,ab OR "statistical monitoring":ti,ab

PubMed
13 May 2019: 1697 hits; Update 16 March 2021: 2198 hits

("on site monitoring"[tiab] OR "on-site monitoring"[tiab] OR "monitoring strategy"[tiab] OR "monitoring
method"[tiab] OR "monitoring technique"[tiab] OR "triggered monitoring"[tiab] OR "targeted monitoring"[tiab] OR "risk-adapted monitoring"[tiab] OR "risk adapted monitoring"[tiab] OR "risk-based
monitoring"[tiab] OR "risk based monitoring"[tiab] OR "risk proportionate"[tiab] OR "centralized
monitoring"[tiab] OR "centralised monitoring"[tiab] OR "statistical monitoring"[tiab] OR "central monitoring"[tiab] OR “trial monitoring”[tiab] OR “study monitoring”[tiab]) AND ("Clinical Studies as
Topic"[Mesh] OR (("randomized controlled trial"[pt] OR controlled clinical trial[pt] OR trial*[tiab]
OR study[tiab] OR studies[tiab]) AND (conduct*[tiab] OR practice[tiab] OR manag*[tiab] OR standard*[tiab] OR harmoni*[tiab] OR method*[tiab] OR quality[tiab] OR performance[tiab])))

Embase (via Elsevier)
13 May 2019: 1245 hits; Update 16 March 2021: 1494 hits
('monitoring strategy':ti,ab OR 'monitoring method':ti,ab OR 'monitoring technique':ti,ab OR 'triggered monitoring':ti,ab OR 'targeted monitoring':ti,ab OR 'risk-adapted monitoring':ti,ab OR 'risk adapted monitoring':ti,ab *OR 'risk based monitoring'/exp* OR 'risk proportionate':ti,ab OR 'trial monitoring':ti,ab OR 'study monitoring':ti,ab OR 'statistical monitoring':ti,ab *OR (monitor* NEAR/2 (site OR risk OR central*)):ti,ab)*
AND
('clinical trial (topic)'/exp OR ((trial* OR study OR studies) NEAR/3 (conduct* OR practice OR manag* OR standard* OR harmoni* OR method* OR quality OR performance)):ti,ab)

2 Grey literature search

Sources:

OpenSIGLE; British Library; Direct Plus; BIOSIS databases (www.biosis.org/); Web of Science; Citation Index; Conferences; Web of Science (Core Collection) Proceedings Paper, Meeting Abstracts; Handsearch of References in identifies articles;WHO Registry (ICTRP portal); ECRIN Risk-based Monitoring Toolbox

3 Data collection form content

1. General Information

Name of person extracting data, report title, report ID, publication type, study funding source, possible conflicts of interest.

2. Methods and study population (trials)

Study design, duration study, design of host trials, characteristics of host trials (primary care, tertiary care, allocated), total number of sites randomized, total number of sites included in the analysis, stratification of sites. Example: stratified on risk level, country, projected enrolment etc., inclusion /exclusion criteria for host trials.

3. Risk of bias assessment

Random sequence generation, allocation concealment, blinding of outcome assessment, performance bias, incomplete outcome data, selective outcome reporting, other bias, validated outcome assessment – grading of findings (minor, major, critical).

4. Intervention groups

Number randomized to group, duration of intervention period, was there an initial risk assessment preceding the monitoring plan?, classification of trials/sites, risk assessment characteristics, differing monitoring plan for risk classification groups, what was the extent of on-site monitoring in the risk-based monitoring group?, triggers or thresholds that induced on-site monitoring, targeted on-site monitoring visits or according to the original trials monitoring plan?, timing (frequency of monitoring visits, frequency of central/remote monitoring), number of monitoring visits per participant, cumulative monitoring time on-site, mean number of monitoring visits per site, delivery (procedures used for central monitoring structure/components of on-site monitoring triggers/thresholds), who performed the monitoring (part of study team, trial staff – qualification of monitors), degree of source data verification (median number of participants undergoing source data verification), co-interventions (site/study-specific co-interventions).

5. Outcomes

Primary outcome, secondary outcomes, components of primary outcome (finding error domains), predefined level of outcome variables (major, critical, others, upgraded)?, time points measured (end of trial/during trial), factors impacting the outcome measure, person performing the outcome assessment, was outcome/tool validated?, statistical analysis of outcome data, imputation of missing data.

6. Results

Comparison of interventions, outcome, subgroup (error domains), postintervention or change from baseline?, unit of analysis, statistical methods used and appropriateness of these methods.

7. Other information (key conclusions of study authors).

4 Risk of bias assessment for non-randomized studies

Domain	Study	Judgment	Support for judgment
Preintervention			
Confounding	Stenning 2018b	Low risk of bias	Decision for on-site visit dependent on the same triggers within 1 study. Confounding was minimized by matched pair design.
	Knott 2015	Moderate risk of bias	No matching of sites, confounding by other factors possible.
	Fougerou-Leurent 2019	Low risk of bias	Same CRF was analyzed with different methods.
Selection bias	Stenning 2018b	Low risk of bias	Matching of comparator sites by algorithm. Same triggers used for all sites within 1 study.
	Knott 2015	Serious risk of bias	Choice of comparator only matched by region, choice not entirely dependent on trigger scores.
	Fougerou-Leurent 2019	Low risk of bias	Prospective cross-over design: the same case report forms were analyzed with full or targeted source data verification.
Information bias	Stenning 2018b	Moderate risk of bias	Monitoring was not blinded to intervention.
	Knott 2015	Moderate risk of bias	Monitoring was not blinded to intervention.
	Fougerou-Leurent 2019	Serious risk of bias	Monitoring was not blinded. If the clinical research associate spotted false or missing non-key data when checking a key data, they may have corrected the non-key data in the case report form. This bias may have led to underestimate the difference between the 2 monitoring strategies. The full source data verification case report form was considered without errors.
Postintervention			
Confounding	Stenning 2018b	Low risk of bias	The same monitoring extend was performed in both groups, no sign for non-adherence to the intervention.
	Knott 2015	Low risk of bias	The same monitoring extend was performed in both groups, no sign for non-adherence to the intervention.
	Fougerou-Leurent 2019	Low risk of bias	Cross-over design, time factor did not influence results.
Selection bias	Stenning 2018b	Low risk of bias	All follow-up considered.
	Knott 2015	Low risk of bias	All follow-up considered.
	Fougerou-Leurent 2019	Low risk of bias	All follow-up considered.
Information bias	Stenning 2018b	Moderate risk of bias	Judgment of findings not blinded.

Domain	Study	Judgment	Support for judgment
	Knott 2015	Moderate risk of bias	Judgment of findings not blinded.
	Fougerou-Leurent 2019	Moderate risk of bias	The same data management program (missing data, consistency, protocol deviations) was subsequently implemented in each strategy by central data management staff. No information on blinding.
Reporting bias	Stenning 2018b	Low risk of bias	Several reports published, all outcomes reported.
	Knott 2015	Moderate risk of bias	No published protocol and no full report published.
	Fougerou-Leurent 2019	Low risk of bias	Full report published, all outcomes of method section reported.

2.2 Manuscript II: Development of a Risk-tailored Approach and Dashboard for Efficient Management and Monitoring of Investigator-Initiated Trials

Katharina Klatte[1], Suvitha Subramaniam[1], Pascal Benkert[1], Alexandra Schulz[1], Klaus Ehrlich[1], Astrid Rösler[1], Mieke Deschodt[2,3], Thomas Fabbro, Christiane Pauli-Magnus[1]*, Matthias Briel[1,4]*

[1] Department of Clinical Research, University Hospital Basel and University of Basel, Basel, Switzerland

[2] Department of Public Health & Primary Care, KU Leuven, Leuven, Belgium

[3] Competence Centre of Nursing, University Hospitals Leuven, Belgium

[4] Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

* Shared senior authorship

Corresponding author

Katharina Klatte

Department of Clinical Research, University Hospital Basel, Spitalstrasse 12, CH-4031 Basel, Switzerland

Telephone: +41 61 328 50 92

Email: Katharina.Klatte@usb.ch

Abstract

Background: Most randomized controlled trials (RCTs) in the academic setting have limited resources for clinical trial management and monitoring. Inefficient conduct of trials was identified as an important source of waste even in well-designed studies. A comprehensive approach identifying and continuously checking critical trial risks (e.g. insufficient recruitment, missing data) to allow the timely initiation of corrective action is, therefore, needed. We developed a risk-tailored approach with an initial risk assessment of an individual trial that informs the compilation of monitoring and management procedures in a trial dashboard.

Methods: We performed a systematic literature review to identify risk indicators and trial monitoring approaches followed by a contextual analysis involving local, national and international stakeholders. Based on this work we developed a risk-tailored management approach with integrated monitoring for RCTs and including a visualizing trial dashboard. We piloted the approach and refined it in an iterative process based on feedback from stakeholders and performed formal user testing with investigators and staff of two clinical trials.

Results: The developed risk assessment comprises four domains (patient safety and rights, overall trial management, intervention management, trial data). An accompanying manual provides rationales and detailed instructions for the risk assessment. We programmed two trial dashboards tailored to one medical and one surgical RCT to manage identified trial risks based on daily exports of accumulating trial data. We made the code for a generic dashboard available on GitHub that can be adapted to individual trials.

Conclusions: The presented trial management approach with integrated monitoring enables user-friendly, continuous checking of critical elements of trial conduct to support trial teams in the academic setting.

Keywords: Clinical trial; trial management; risk-tailored monitoring; trial dashboard

Introduction

Randomized controlled trials (RCTs) are the gold standard for assessing the effects of medical interventions. However, they are typically resource intense and pose various organisational challenges. [1-3] Inefficient management and monitoring of RCTs have been identified as an important source of waste.[1-5] Monitoring efforts are traditionally quite generic and extensive, but problems such as slow participant recruitment, considerable losses to follow-up, or poor data quality are often recognized too late during trial conduct delaying necessary adjustments of processes or the protocol. In addition, resources for clinical trial monitoring and management are usually scarce in the academic setting and sophisticated commercial solutions can be costly. [6 7]

Organisational challenges and critical factors jeopardizing trial integrity and quality may vary considerably across trials; therefore, a risk assessment conducted prior to trial initiation or at certain intervals during trial conduct may yield different risk profiles for individual trials. Trial monitoring protects the safety and rights of participants, ensures data are accurate, complete and verifiable, and that the trial follows the principles of good clinical practice [8 9] Currently recommended risk-based trial monitoring allows for an adaptation of the monitoring intensity according to an initial risk assessment of a trial and has been developed to reduce resource intense onsite visits with source data verification for non-high-risk trials.[1-3 10-12 13 14] However, this approach typically does not consider individual risk profiles of RCTs, but rather classifies trials by generic risk categories.[14] To accommodate individual trial risks, a monitoring strategy may include several components such as centralized monitoring (evaluation of accumulated trial data performed in a timely manner at a central location), onsite monitoring (performed at investigator sites with source data verification and review of protocol-specified processes), or remote monitoring (same tasks as onsite monitoring but performed away from investigator sites)[13 15 16].

Trial management should provide for smooth and reliable trial procedures including participant recruitment, randomisation, intervention application, data collection, and data cleaning.[17 18] Data cleaning and checking of recruitment and retention rates, for instance, need to be performed in a timely fashion, so that corrective measures can be taken early on and detrimental effects on the trial can be avoided.[19] Trial monitoring is most effective when performed on cleaned data, because incorrect processes may be missed due to poor data quality and monitoring efforts are wasted on individual data errors. Therefore, trial management and monitoring ideally are

integrated tasks that make use of accumulating data during trial conduct, i.e. continuously keeping oversight of complex study processes and performing centralized data monitoring. [20-22]

The objective of this project was to develop a risk-tailored approach that integrated trial management and monitoring in investigator-initiated RCTs. We closely collaborated with relevant stakeholders (trial coordinators, principal investigators, data managers, trial monitors, statisticians) to create a user-friendly dashboard that efficiently visualizes data on critical processes of individual trials.

Methods

Overview of research process

In the first phase of this user-centred project,[23] we developed a concept of a risk-tailored trial monitoring and management approach with corresponding trial dashboard (**Figure 1**). The development involved relevant stakeholder groups and was based on the results of systematic literature reviews on existing monitoring strategies,[13] and a contextual analysis to identify current practices and needs of various user groups. The concept and dashboard were piloted and refined in an iterative process involving different end users and other stakeholder groups. In the second phase, we performed formal user testing of the developed risk assessment and dashboard. Experiences of investigators and trial staff of one medical and one surgical investigator-initiated RCT were gathered using semi-structured interviews to further refine the concept and dashboard.

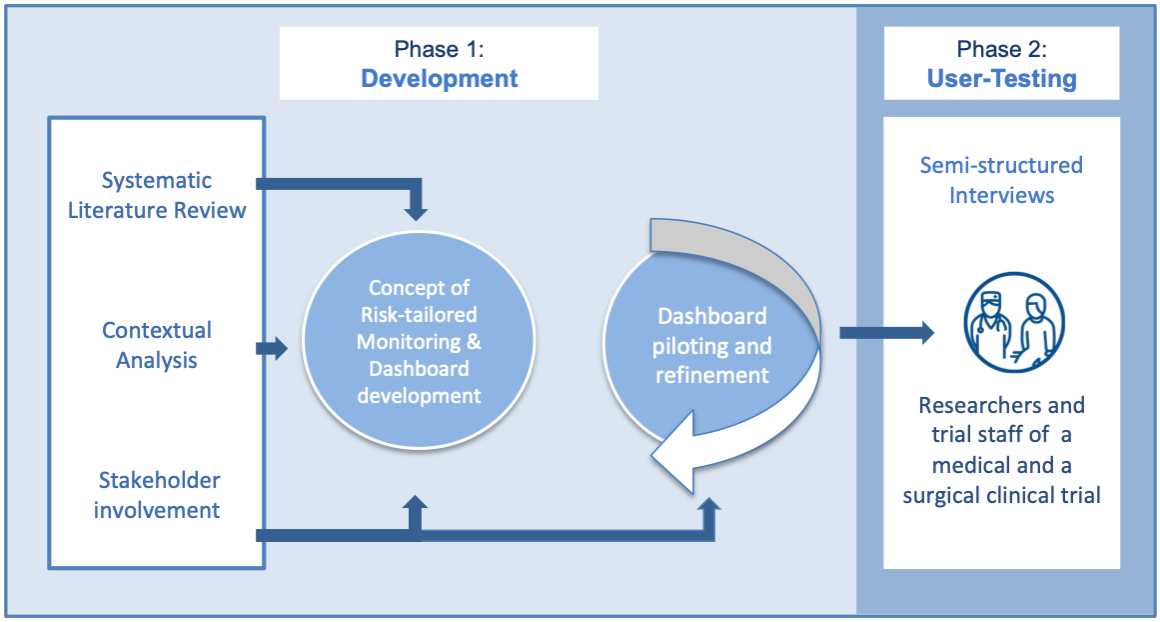


Figure 1: Overview of the two phases of the development and user-testing of the risk-tailored approach and trial dashboard

Systematic Literature Review

To identify and structure components for the initial risk assessment of individual trials, we systematically searched for published risk assessment approaches and risk indicators used to support trial oversight and to identify centres in need for support. We considered different components and qualitative evidence from process evaluations of tested monitoring strategies summarized in a previously conducted systematic review.[10] We further considered the guideline of the European Clinical Research Infrastructure Network (Ecrin) [14] and the risk assessment guideline developed by the Swiss Clinical Trial Organization [24], TransCelerate metrics [25 26], Whitham metrics [27], and the trial specific metrics used by the Medical Research Council (MRC) Clinical Trials Unit (CTU) at University College London (UCL) Trial specific metrics [28]. Results from this literature review are summarizes in **Supplementary Table 1**.

Stakeholder involvement

We set up a local, multidisciplinary working group including end users and representatives of different stakeholder groups within the Department of Clinical Research (DKF) and associated research groups at the University Hospital Basel. At this local level, we involved members from the Data Science and Data Management Teams of the DKF experienced in central monitoring, R shiny applications,

dashboard development, data base structures and exports; we involved trial monitors with experience in on-site and remote monitoring, knowledge of study site structures and processes; study coordinators and investigators experienced in managing RCTs. Stakeholder meetings with all members of these groups provided an additional opportunity for feedback and exchange of information on the risk assessment and dashboard development as well as on the application strategy. In order to get input from a national group of stakeholders in Switzerland, we contacted the national platform of the Swiss Clinical Trial Organisation for trial monitoring. Finally, we gathered experiences from international methodological research groups and UK-based CTUs using risk-based approaches or study dashboards to support trial conduct. The different activities with stakeholders at all levels are summarized in **Supplementary Table 2**.

Contextual Analysis

Gathering contextual input from various end users and the above-mentioned stakeholders guided the development of the risk-tailored approach and helped to determine relevant domains and applications to be considered in the initial risk assessment. We structured the identified stakeholder needs into content related factors such as the inclusion of the follow-up visits into the risk assessment, and design related factors such as the suggested separation of severity and likelihood in the assessment or the colour code for the status of queries visualized in the dashboard (**Supplementary Table 3)**. In terms of content of the risk assessment, it became clear, for instance, that the assessment covers a wide spectrum of risks applicable to a large variety of RCTs. The design of the risk assessment guide should support the intuitive assessment by different end user groups (monitors, study managers, principal investigators). The study dashboard should reflect the outcome of the risk assessment and the design of the dashboard should enable an efficient navigation within the routine study procedure by end-users. The findings of the contextual analysis are summarized in **Supplementary Table 3.**

Development and piloting of the concept and dashboard

Based on the systematically reviewed literature, our contextual analysis and stakeholder input, we drafted a generic risk-assessment template. We then created trial-specific dashboards for a medical and a surgical multicentre trial that differed in their risk profile, but both comprised complex study procedures and data collection. The risk-tailored approach continued to evolve as we gathered contextual information, detected gaps in the assessment procedure, and identified critical components of study management. We developed R code to extract data values from exported data tables and summarized, compared, and calculated relevant information to create pathways for the identified risks. The output of these operations was then visualized in the trial dashboard. The piloting and refinement was an iterative process incorporating repeated feedback from the end-users and the stakeholder representatives in the project group on dashboard content, structure, user-friendly interface, and visualization of critical study data.

User testing

The aim of the user testing was to identify challenges in the routine use of the dashboard experienced by different user groups. Each of the six users (i.e. 2 trial managers, 2 monitors, 2 principal investigators) received a detailed manual of the features and operation mode of the study dashboard.
We interviewed users 6-12 weeks after using the study dashboard in daily trial routine. We followed a semi-structured interview guide, which allowed for expansion on topics that emerged during the interview. All interviews took approximately 30 minutes. The interviewer (KK) transcribed the recorded interviews and extracted suggestions for improvement. We then updated the trial dashboard based on the feedback of the users and provided the adapted version for further use and evaluation.

Results

The final concept consisted of the following three steps: trial-specific risk assessment prior to study start, selection and development of data-based pathways to address identified risks, and visualization of pathways output in a trial dashboard.

Trial-specific risk assessment

The trial-specific risk assessment comprised four domains (participant safety and rights, overall study management, device/medication management, study data), and each domain contained several risk elements (**Table 1).** To better assess if these elements are critical for a specific trial and which trial components are at particular risk, we determined trial assets and corresponding risk scenarios. Trial assets are conditions essential for the successful and proper conduct of a trial, e.g. visits must be scheduled and take place in the required timeframe, SAE have to be reported on time and need to be closely followed over the whole study conduct. If a trial includes many follow-up visits over a long follow-up time and assessment have to take place in a very narrow time window, this asset would be considered at risk (example shown in **Table 2**, Part A). Other assets, for example SAE reporting and oversight, are essential for all clinical trials and, thus, are considered as a risk that applies to all trials (marked in red, Example shown in **Table 2**, Part B). The identified risks are then analysed in terms of severity and likelihood. For example, if many follow-up visits need to be coordinated but the time window of the endpoint assessment is wide the severity is rated as less critical. The likelihood is highly influenced by the experience of the trial team and participating centres with similar trials, training and experience of all involved staff members, and the resources available for the study.
The complete list of assets, as well as the corresponding risk scenarios, is provided in the full risk assessment in **Supplementary Table 4**. The first risk assessment should be performed before the start of the trial and based on the study protocol, Case Report Forms (CRFs), the planned and actual budget of the study, expected recruitment rates for all participating centres, information on the trial intervention, and information about planned study staff (see Appendix for detailed **Manual**).

Table 1: Domains and their attributed risk elements

Domain	Risk Elements
Participant Safety and Rights	Informed consent AE/SAE reporting and documentation Inclusion/exclusion
Overall Study Management	Recruitment Retention Study procedures and endpoint assessment (e.g. bio sampling, imaging quality) Participant schedule (e.g. timeframe of visits) AE/SAE management
Device/ Medication Management	Administration Accountability/ storage
Study Data	Data quality – completeness, consistency, timeliness Documentation/ storage

Abbreviations: AE, adverse events; SAE, serious adverse events

Table 2: Example of assets and risk scenarios for risk elements in the domain Overall Study Management (Part A) and Participant Safety and Rights (Part B). Assets that apply to all trials are marked in red.

A)

Domain	Risk element	Asset	Risk scenario
Overall Study Management	**Participant Schedule**	Visits/Phone calls must be within the given Timeframe	(A) Time point of visit is critical for the endpoint assessment of the study
			(B) Large number of visits are difficult to organize and coordinate between centres and patients

B)

Domain	Risk element	Asset	Risk scenario
Participant Safety and Rights	**SAE/AE**	SAE have to be reported and documented correctly in the required timeframe	Complexity of CRF or missing SOPs for SAE Reporting leads to (A) Incorrect documentation and (B) Delayed reporting of SAEs

Abbreviations: CRF, case report form; SOPs, standard operating procedures; SAE, serious adverse events

Pathways to manage identified risks

In order to continuously manage identified risks, we created pathways that eventually allowed for tailored visualization of accumulating trial data and implemented action at suitable time intervals (e.g., email reminders, staff overviews) in a study dashboard. The operations applied to the exported data tables via R code are dependent on the specific information needed to provide a clear oversight on identified risk elements. The code is structured into modules that contain the operations of all pathways visualized in one dashboard tab (e.g. SAE management). For example, the module SAE contains operations that count the number of SAEs, determine the number of patients with SAE and calculate the ratio SAEs per patient randomized. In addition, information like severity, causality and outcome are extracted from the SAE form data table and percentages of value options (e.g. SAE outcome: Continuing, Resolved with sequel, Resolved with sequel, others) are calculated and graphically displayed (**Figure 2**, Panel A and B). The developed study dashboards contain tabs that visualize the output of created pathways reflecting identified study-specific risks. These tabs are based on the R modules containing the pathways as well as the code required for a clear visual presentation (value boxes, graphs, lists). When pilot testing our risk assessment guide, it became apparent that some risks apply to almost all trials (marked in red in the full risk assessment **Supplementary Table 4**). The management of these risks is, thus, based on tabs classified as “generic” in the study dashboard, while other, more seldom and study-specific risks are considered in “optional” tabs (**Table 3**). The content of generic tabs can also be adapted depending on, for instance, the complexity or time point of outcome assessment in a trial. The

generic dashboard template is freely available on GitHub (https://github.com/CTU-Basel/viewTrial).

Visualization of data based pathways

The output of the pathways is visualized in the corresponding tabs in the study dashboard. The arrangement of the tabs within the study dashboard can be determined by study teams; a division into study management related tabs and oversight/study progress tabs might provide a better overview for the different user groups (principal investigator, study manager, and trial monitor). The main tabs can also contain sub-tabs. For example, the number of due visits is displayed under the visits tab in the sub-category “due visits”. In this context, the definitions of due, overdue, and missed visits are dependent on the specific timeframes of the study protocol. Total numbers are provided as well as a list of the patient ID and a direct link to the corresponding eCRF in the database (**Figure 2**, Panel A). Each tab or sub-tab can represent several pathway outputs displayed in form of value boxes, graphical presentations, or lists of relevant patients. For example, the SAE management tab provides an overview on SAE prevalence in boxes, and in additional panels the user can switch between the graphical representation of SAE severity, causality, and outcome. Additionally, a list of patients with SAE is provided below, displaying information on SAE status (e.g. ongoing/closed) and a short description of the event (**Figure 2**, Panel B). The information is provided for the overall study, including all randomized patients as numbers and percentages in boxes, while graphs differentiating between centres are provided to better assess which centres are in need for support in a certain aspect of the study conduct. In addition, the dashboard allows filtering for specific centres and time ranges of interest or choosing particular study visits from drop down menus to provide users with more detailed information (see **Supplementary Figure 1** for an example). The output of the pathways visualized in the dashboard is based on a daily export of trial data and, thus, includes up-to-date information on randomised patients and entered data. The generic and some of the optional tabs are listed in **Table 3**. Examples of the tabs from the two study dashboards are provided in **Supplementary Figures 2-5**. The generic dashboard is accessible via GitHub and generic data is provided to test the different code modules behind each tab (examples provided in **Supplementary Figure 6 and 7**).

A)

B)

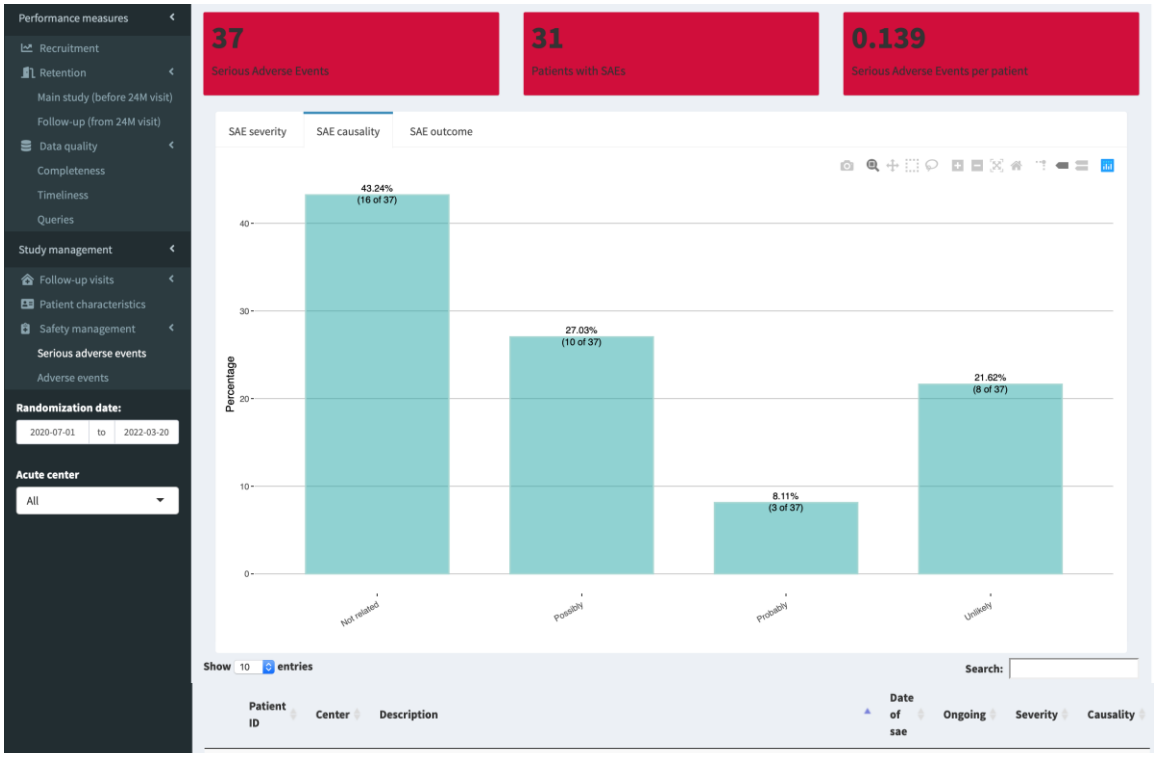


Figure 2: Dashboard screenshots of the Visits tab, sub-tab “Due visits” (Panel A), and the Safety management tab, sub-tab “Serious adverse events” (Panel B)

Table 3: Structure and content of dashboard tabs

Domain	Risk Elements	Example Tabs	Content of Tab	Functionality/Purpose	Generic/ Optional
Participant Safety and Rights	Informed consent	**Informed consent**	In case of a re-consent this tab can provide an overview of patients who have not yet signed the re-consent form	To ensure patient rights and support of re-consent process through site-specific reminders, list of patients that still need a re-consent.	Optional
	AE/SAE reporting and documentation	**AE/SAE**	Provides an overview of timeliness and completeness of AE/SAE entries	To ensure that all AE/SAE forms are complete and that the date of first entry is within the required reporting timeframe	Generic
	Inclusion/exclusion	**Safety**	In case of safety-relevant inclusion or exclusion criteria, a verification of relevant information available in the data base can provide additional security (e.g. blood pressure has to be within a certain range – check for the entry of blood pressure in the data base)	To provide the option for additional checks for inclusion/ exclusion criteria besides the marked list of criteria in the eCRF	Optional
Overall Study Management	Recruitment	**Recruitment**	Recruitment trajectories for expected and actual recruitment in total and per centre (**Supplementary Figure 2)**	To monitor the progress of participant recruitment enabling early action in case of slow recruitment.	Generic
		Patient Characteristics	Relevant patient characteristics are summarized and presented (e.g. gender, age, background of treatment)	To inform the study team on the accuracy of inclusion/exclusion criteria and provide an overview of the sample population in terms of relevant characteristics	Generic
	Retention	**Retention**	Patients who have ended the study resulting in missing outcome data, reasons for leaving the	To monitor the progress of participant retention, consider reasons for ending study in recruitment. Time point of	Generic

Domain	Risk Elements	Example Tabs	Content of Tab	Functionality/Purpose	Generic/ Optional
			study, kind of data collected before study end (Primary outcome data available) (**Supplementary Figure 3**)	ending the study important for amount of data analysable.	
	Study procedures and endpoint assessment	**Bio sampling (e.g. blood samples)**	Overview of samples taken and availability of sample results	To support sample management in terms of localization and status of bio sample. Important for biomarker determination.	Optional
		Imaging quality	Automated and visual verification of imaging data quality, e.g., for MRI or CT	To enable early adjustments in case of low quality imaging data and ensure that the imaging data is analysable.	Optional
	Participant schedule:	**Follow-up visits**	Overview of follow-up visits with a particular focus on visits where primary outcome data is collected. **(Figure 3, Panel A**)	To assist in integrating follow-up visits on time into the daily clinical routine might be difficult for trial sites. Support through reminders for due visits can be initiated through the dashboard.	Optional
	AE/SAE management	**Safety management (SAEs, AEs)**	The Safety tab provides an overview of SAEs and AEs that have been reported in the study and information on severity and outcome of SAEs/AEs (**Figure 3,** Panel B)	To estimate potential safety issues (e.g. SAEs occurring more often in one study arm, number of SAEs in total, number of patients with SAE)	Generic
Device/ Medication Management	Administration Accountability/ storage	**Medication**	Overview of medication consumption based on number of patients and their current position in the medication plan per protocol and comparison with IMP stock at sites	To assist in the managing of IMP stock overview and enable reminders for restocking	Optional

Domain	Risk Elements	Example Tabs	Content of Tab	Functionality/Purpose	Generic/ Optional
Study Data	Data quality – completeness, consistency, timeliness Documentation/ storage	**Data Quality**	Completeness of forms (Primary end point, secondary endpoint, SAE/AE forms) Timeliness of data entry, Number of queries, status of queries (open, resolved) (**Supplementary Figures 4,5**)	To increase awareness of items missing in the data base Trial sites may have different challenges when integrating a trial in their daily clinical routine and therefore need support in different aspects of the study conduct. Completeness and timeliness of data entry as well as query management constitute indicators for need of support. Query status helps the study monitor to decide which centre needs more assistance/ on-site visit.	Generic

Abbreviations: AE, adverse events; CT, computerized tomography ;IMP, investigational medicinal product; MRI, magnetic resonance imaging; SAE, serious adverse events

User testing

The users testing of our study dashboards provided positive feedback in terms of improved study oversight and facilitated conduct. Trial monitors and study staff agreed that the initial risk assessment was beneficial, because it increased the awareness of critical processes in the collection of outcome data, enabling corrective measures at an early time point, e.g. adaptation of database structures. A clear benefit perceived by all user groups was the more frequent and improved communication with trial sites; sites were better prepared for remote or on-site monitoring visits, because many issues were recognized and solved in advance. In addition, users made several suggestions for further elements to be included in the dashboard. A detailed summary of the results from the user testing is provided in **Supplementary Table 5**.

Discussion

Using a systematic approach involving relevant stakeholder groups, we developed a concept of risk-tailored trial monitoring and management that focuses on the identification and control of trial specific risks during trial conduct. The continuous evaluation of most important risks provides important information about the study progress, e.g. in terms of recruitment, endpoint assessment, as well as in terms of data management and data quality, e.g. CRF completion, timeliness of follow-up visits. Completeness of essential data points as the basis for analysable patient data is continuously evaluated and trial monitors and study managers maintain an overview of visit timeframes, SAE reporting, and query management.

Strengths and limitations

Strengths of our study are the systematic and structured process of development of the risk assessment and the trial dashboard, which included the involvement of all local stakeholder groups and the performance of a comprehensive contextual analysis. In addition, the development was based on prior evidence gathered through systematic literature searches and exchange with international stakeholder groups. Directly involving end users in developing and evaluating the usability of our tool may facilitate the implementation process, promote wider adoption, maintain involvement, and increase user satisfaction with the concept as well as the tool.[29] Providing an R code repository for other study teams that can be adapted and applied to differently

structured data bases, constitutes a software-independent, affordable approach for the limited budget of investigator-initiated trials.

Our study has the following limitations: First, we performed user testing in two ongoing RCTs only, and, thus, the spectrum of feedback may have been limited and may compromise the extrapolation of mentioned benefits and disadvantages to other trials. Both RCTs had already started participant recruitment when the dashboard was implemented. This allowed for a qualitative comparison of management and monitoring processes without and with the dashboard tool in place. However, it will be crucial to subsequently evaluate the impact and value of the study dashboard during the entire course of a clinical trial. Since both RCTs are still ongoing, we could not evaluate the impact of the tool on participant safety and overall trial success, including the percentage of analysable data, at the end of a trial. Lastly, we have not yet evaluated any cost-effectiveness of our developed approach, e.g. assessing whether the dashboard has the potential to reduce monitoring and management hours needed to ensure a safe and successful trial conduct. While some users felt that our dashboard would only be worthwhile for multicentre trials, others found that the costs of providing a study dashboard will always depend on the needs and preferences of the study team and the complexity of the study.

Comparison with similar studies and frameworks

Following the recommendations of the Clinical Trials Transformation Initiative (CTTI), effective and efficient monitoring and management needs to first determine what matters for a specific trial and focus on areas of highest risk for generating errors that matter.[30 31] With our risk assessment guide and the study dashboard we address the need for this focus and provide a tool that supports the continuous oversight of the quality of the trial conduct.

Dashboards that visualize time-dependent parameters have recently met a growing acceptance in medical and administrative health care settings.[32-39] Dashboards have been introduced to support various aspects of clinical trials, including web applications for eligibility screening and overview of the enrolment progress [37], web-based support of recruitment management and communication; [38] graphical summaries and diagrams of the progress of patient accrual and form completion [39], feedback on data completeness by using a traffic light system[40], and automated reports of data compliance, protocol adherence and safety [41]. These available dashboards typically focus on specific elements of trial conduct and communication with trial sites; however, our dashboard provides a comprehensive overview of all

elements of a trial identified as critical. In addition, tables and graphical representations are often limited to certain time intervals. [37] The daily export of trial data providing up-to-date trial information is part of the core idea of our approach as it enables immediate actions and improves communication with site staff.

Several commercial solutions supporting the overall trial conduct in various aspects are readily available)[6 42-46], but for investigator-initiated trials with tight budgets such software packages typically remain unaffordable. We wanted to provide a comprehensive and affordable option for investigator-initiated trials that can be adapted to individual needs and preferences and further developed by the research community. Therefore, we transparently present all details of the structured risk assessment and manual as well as the generic code for our dashboard in publicly accessible repositories via GitHub. We invite users to report difficulties or suggestions for improvement for consideration in future modifications of the generic dashboard via GitHub.

Implications

Besides the emphasis on the feasibility and design of clinical trials, measures to increase the efficiency of clinical trial conduct are needed. [47] Current challenges include premature discontinuation of a significant proportion of clinical trials, and inflated costs mainly due to delayed recruitment and organisational issues.[47] We propose a comprehensive approach integrating management and monitoring of a clinical trial into one risk management tool supporting the conduct of investigator-initiated trials.

Overseeing the progress of a trial in each centre based on up-to-date information, provides the opportunity for trial monitors to prioritize centres for on-site visits or remote interactions, tailor their action to the specific issues of a centre, and guide decisions on where resources and training is needed the most. In addition, providing automated reminders for upcoming visits or sampling, overview of investigational medicinal product supply, overview of patients who need a re-consent, overview of ongoing SAEs, etc. can increase the efficiency of the trial management processes. The tool further provides the opportunity to improve the overall communication between the study team and trial sites and may increase motivation through the involvement of sites in the trial progress and the option to compliment active participation in the trial. Feedback from the user testing also revealed a positive perception of study managers and investigators to improved data quality visible in the dashboard: “If incomplete is empty, I am at ease”.

The impact of this tool is largely dependent on the successful implementation into clinical trial practice. The perception of benefits and opportunities by stakeholders and end-users have been collected while the effectiveness of the tool in terms of analysable data collected, timeline of recruitment, conformity of SAE/AE reporting and documentation, support of the overall study management still have to be evaluated.

The next step is now to implement the risk assessment as a routine step in the joint planning of clinical trials with the respective study teams. The timely generation of a dashboard on the basis of the generic template and further study-specific risks has to be organized. Strategies to further evaluate this implementation process as well as the effectiveness of this new approach in studies of different design and structure have to be developed. This evaluation will provide more information on the feasibility of study-specific dashboards supporting trial monitoring and management in the heterogeneous field of clinical trials.

Conclusion

In summary, the presented risk-assessment guide and dashboard tool provide a systematically developed and user-tested instrument for the risk-tailored support of trial monitoring and trial management. Feedback from the user testing of the instrument revealed many benefits for the involved stakeholder groups. However, the effectiveness of the dashboard in terms of a safe trial conduct and overall support for a successful completion of clinical trials needs to be further evaluated.

References

1. Yusuf S. Randomized clinical trials: slow death by a thousand unnecessary policies? *CMAJ* 2004;171(8):889-92; discussion 92-3. doi: 10.1503/cmaj.1040884 [published Online First: 2004/10/13]
2. Eisenstein EL, Lemons PW, 2nd, Tardiff BE, et al. Reducing the costs of phase III cardiovascular clinical trials. *Am Heart J* 2005;149(3):482-8. doi: 10.1016/j.ahj.2004.04.049 [published Online First: 2005/05/03]
3. Eisenstein EL, Collins R, Cracknell BS, et al. Sensible approaches for reducing clinical trial costs. *Clin Trials* 2008;5(1):75-84. doi: 10.1177/1740774507087551 [published Online First: 2008/02/20]
4. Kasenda B, von Elm E, You J, et al. Prevalence, characteristics, and publication of discontinued randomized trials. *JAMA* 2014;311(10):1045-51. doi: 10.1001/jama.2014.1361 [published Online First: 2014/03/13]
5. Al-Shahi Salman R, Beller E, Kagan J, et al. Increasing value and reducing waste in biomedical research regulation and management. *Lancet* 2014;383(9912):176-85. doi: 10.1016/S0140-6736(13)62297-7 [published Online First: 2014/01/15]
6. CluePoints.Risk-basedmonitoring. http://cluepoints.com/risk-based-monitoring. 2016 [accessed 31st July 2022.
7. Medidata.Strategic monitoring https://www.medidata.com/wp-content/uploads/2018/12/Strategic-Monitoring_2018Medidata_Fact-Sheet.pdf2018 [accessed 31st July 2022.
8. International Conference on Harmonisation of technical requirements for pharmaceuticals for human use (ICH). Guideline for good clinical practiceE6(R2) Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-6-r2-guideline-good-clinical-practice-step-5_en.pdf2018 [accessed 21st July 2022.
9. Medicines and Healthcare products Regulatory Agency. Good clinicalpractice guide: . The Stationery Office, 2012.
10. Califf RM. Clinical trials bureaucracy: unintended consequences of well-intentioned policy. *Clin Trials* 2006;3(6):496-502. doi: 10.1177/1740774506073173 [published Online First: 2006/12/16]

11. Sertkaya A, Wong HH, Jessup A, et al. Key cost drivers of pharmaceutical clinical trials in the United States. *Clin Trials* 2016;13(2):117-26. doi: 10.1177/1740774515625964 [published Online First: 2016/02/26]
12. Funning S GA, Eriksson K, Kettis-Lindblad A Quality assurance within the scope of Good Clinical Practice (GCP) - What is the cost of GCP-related activities? A survey within the Swedish Association of the Pharmaceutical Industry (LIF)'s members. *Qual Assur J* 2009;12, 3-7.
13. Klatte K, Pauli-Magnus C, Love SB, et al. Monitoring strategies for clinical intervention studies. *Cochrane Database Syst Rev* 2021;12:MR000051. doi: 10.1002/14651858.MR000051.pub2 [published Online First: 2021/12/09]
14. Ecrin. Risk-Based Monitoring Toolbox Available from: https://ecrin.org/tools/risk-based-monitoring-toolbox2021 [
15. Baigent C, Harrell FE, Buyse M, et al. Ensuring trial validity by data quality assurance and diversification of monitoring methods. *Clin Trials* 2008;5(1):49-55. doi: 10.1177/1740774507087554 [published Online First: 2008/02/20]
16. Love SB, Armstrong E, Bayliss C, et al. Monitoring advances including consent: learning from COVID-19 trials and other trials running in UKCRC registered clinical trials units during the pandemic. *Trials* 2021;22(1):279. doi: 10.1186/s13063-021-05225-5 [published Online First: 2021/04/16]
17. Farrell B, Kenyon S, Shakur H. Managing clinical trials. *Trials* 2010;11:78. doi: 10.1186/1745-6215-11-78 [published Online First: 2010/07/16]
18. Council MR. Clinical Trials for Tomorrow: An MRC Review of Randomised Control Trials: Medical Research Council 2003.
19. Love SB, Yorke-Edwards V, Diaz-Montana C, et al. Making a distinction between data cleaning and central monitoring in clinical trials. *Clin Trials* 2021;18(3):386-88. doi: 10.1177/1740774520976617 [published Online First: 2021/03/04]
20. Organisation S-MPotSCT. F dated. Fact sheet: Central Data Monitoring in Clinical Trials – V 1.0. available from www.scto.ch/monitoring SCTO; [accessed May 3rd 2022.
21. Venet D, Doffagne E, Burzykowski T, et al. A statistical approach to central monitoring of data quality in clinical trials. *Clin Trials* 2012;9(6):705-13. doi: 10.1177/1740774512447898 [published Online First: 2012/06/12]
22. Agrafiotis DK, Lobanov VS, Farnum MA, et al. Risk-based Monitoring of Clinical Trials: An Integrative Approach. *Clin Ther* 2018;40(7):1204-12. doi: 10.1016/j.clinthera.2018.04.020 [published Online First: 2018/08/14]

23. Walden A, Garvin L, Smerek M, et al. User-centered design principles in the development of clinical research tools. *Clin Trials* 2020;17(6):703-11. doi: 10.1177/1740774520946314 [published Online First: 2020/08/21]
24. Organisation S-SCT. Risk assessment Form for Clinical Research Projects available from: https://www.sctoplatforms.ch/en/tools/risk-assessment-form-for-clinical-research-projects-30.html2018 [accessed May 3rd 2022.
25. Wilson B, Provencher T, Gough J, et al. Defining a Central Monitoring Capability:Sharing the Experience of TransCelerate BioPharma's Approach, Part 1. *Therapeutic Innovation & Regulatory Science* 2014;48(5):529-35. doi: 10.1177/2168479014546335
26. Gough J, Wilson B, Zerola M, et al. Defining a Central Monitoring Capability:Sharing the Experience of TransCelerate BioPharma's Approach, Part 2. *Therapeutic Innovation & Regulatory Science* 2016;50(1):8-14. doi: 10.1177/2168479015618696
27. Whitham D, Turzanski J, Bradshaw L, et al. Development of a standardised set of metrics for monitoring site performance in multicentre randomised trials: a Delphi study. *Trials* 2018;19(1):557. doi: 10.1186/s13063-018-2940-9 [published Online First: 2018/10/18]
28. Diaz-Montana C, Cragg WJ, Choudhury R, et al. Implementing monitoring triggers and matching of triggered and control sites in the TEMPER study: a description and evaluation of a triggered monitoring management system. *Trials* 2019;20(1):227. doi: 10.1186/s13063-019-3301-z [published Online First: 2019/04/19]
29. Goodman MS, Sanders Thompson VL. The science of stakeholder engagement in research: classification, implementation, and evaluation. *Transl Behav Med* 2017;7(3):486-91. doi: 10.1007/s13142-017-0495-z [published Online First: 2017/04/12]
30. CTTI. RECOMMENDATIONS: EFFECTIVE AND EFFICIENT MONITORING AS A COMPONENT OF QUALITY ASSURANCE IN THE CONDUCT OF CLINICAL TRIALS Available at:

https://ctti-clinicaltrials.org/wp-content/uploads/2021/06/CTTI_Monitoring_Recs.pdf2021 [accessed 21st July 2022.
31. CTTI. QUALITY BY DESIGN PROJECT - CRITICAL TO QUALITY (CTQ) FACTORS PRINCIPLES DOCUMENT Available at:

https://ctti-clinicaltrials.org/wp-content/uploads/2021/07/CTTI_QbD_Workshop_Principles_Document.pdf 2021 [accessed 21st July 2022.

32. Simpao AF, Ahumada LM, Desai BR, et al. Optimization of drug-drug interaction alert rules in a pediatric hospital's electronic health record system using a visual analytics dashboard. *J Am Med Inform Assoc* 2015;22(2):361-9. doi: 10.1136/amiajnl-2013-002538 [published Online First: 2014/10/17]
33. Hartzler AL, Izard JP, Dalkin BL, et al. Design and feasibility of integrating personalized PRO dashboards into prostate cancer care. *J Am Med Inform Assoc* 2016;23(1):38-47. doi: 10.1093/jamia/ocv101 [published Online First: 2015/08/12]
34. Dowding D, Randell R, Gardner P, et al. Dashboards for improving patient care: review of the literature. *Int J Med Inform* 2015;84(2):87-100. doi: 10.1016/j.ijmedinf.2014.10.001 [published Online First: 2014/12/03]
35. Waitman LR, Phillips IE, McCoy AB, et al. Adopting real-time surveillance dashboards as a component of an enterprisewide medication safety strategy. *Jt Comm J Qual Patient Saf* 2011;37(7):326-32. doi: 10.1016/s1553-7250(11)37041-9 [published Online First: 2011/08/09]
36. Nelson SD, Del Fiol G, Hanseler H, et al. Software Prototyping: A Case Report of Refining User Requirements for a Health Information Exchange Dashboard. *Appl Clin Inform* 2016;7(1):22-32. doi: 10.4338/ACI-2015-07-CR-0091 [published Online First: 2016/04/16]
37. Arab L, Hahn H, Henry J, et al. Using the web for recruitment, screen, tracking, data management, and quality control in a dietary assessment clinical validation trial. *Contemp Clin Trials* 2010;31(2):138-46. doi: 10.1016/j.cct.2009.11.005 [published Online First: 2009/11/21]
38. Mattingly WA, Kelley RR, Wiemken TL, et al. Real-Time Enrollment Dashboard For Multisite Clinical Trials. *Contemp Clin Trials Commun* 2015;1:17-21. doi: 10.1016/j.conctc.2015.09.001 [published Online First: 2016/02/16]
39. Toddenroth D, Sivagnanasundaram J, Prokosch HU, et al. Concept and implementation of a study dashboard module for a continuous monitoring of trial recruitment and documentation. *J Biomed Inform* 2016;64:222-31. doi: 10.1016/j.jbi.2016.10.010 [published Online First: 2016/10/23]
40. Improving data entry and study compliance efficiently using immediate audit and feedback tools 5th International Clinical Trials Methodology Conference (ICTMC); 2019. Trials.

41. Automation of clinical trial statistical monitoring. 5th International Clinical Trials Methodology Conference (ICTMC); 2019. Trials.
42. PerkinElmer. Risk-based monitoring.: Available at: http://www.perkinelmer.com/product/risk-based-monitoring-rbm.; [accessed August 18th 2022.
43. IQVIA. The next wave of central monitoring. Available at: https://www.iqvia.com/library/white-papers/the-next-wave-of-centralized-monitoring. [accessed August 18th 2022.
44. Intelligence TR. Ttirials: a complete risk-based monitoring solution Available at: https://tritrials.com [accessed August 18th 2022.
45. Cyntegrity. Risk-based monitoring solutions Available at: https://cyntegrity.com [accessed August 18th 2022.
46. Plc I. Patient centric monitoring: preventing and learning from mistakes Available at: http://www.iconplc.com/icon-views/blog/2015/05/21/patient-centric-monitorin [accessed August 18th 2022.
47. Sciences S-SAoM. White Paper: Clinical Research. *Swiss Academies Communications* 2021; 16 (4)

Supporting Information

Supplementary Table 1: Summary of risk indicators identified through systematic literature review

Domain	Risk indicators
Recruitment	• Total number of patients enrolled at a site / the recruitment target set for the trial at this site • Total number of patients who consented/ total number of patients eligible for the trial
Retention	• Number of patients lost to follow-up or who withdrew from the study • Reasons for losses to follow-up/withdrawals

Data Quality	• Systematic errors
	  ○ Abnormal trend in data, specific data item always missing
	• CRF completion
	  ○ Average delta of visit date and data entry (CRF completion) for each centre compared to the average of all centres
	  ○ Percentage of missing data items (Total number of fields available for data entry that are missed or queried / Total number of fields available for data entry)
	• Entry Management
	  ○ Number of queries
	  ○ Data query response rate
	  ○ Number of overdue queries
	  ○ Percentage of patients with open queries
	  ○ Discrepancies between SAE event and day of reporting (more than 7 days)

Follow-up visits	• Are visits taking place in the required timeframe • Number of visits overdue • Number of visits missed (out of timeframe stated in study protocol)
Informed Consent and Eligibility	• Number of patients with informed consent date/ Number of patients randomized • Delta between consent date and randomization date • All eligibility criteria fulfilled based on protocol (If not covered by CRF rules)

Supplementary Table 2: Overview of stakeholder activities supporting the contextual analysis

Stakeholder activity	Participants involved	N	Mode of stakeholder activities	Purpose of stakeholder activities
Local project group meetings	Representatives from 6 different groups: Data science, Data management, Monitoring, Trial management, principle investigators	40 meetings between May 2019 and February 2021, with the number of participants ranging from 5-7 persons	In-person meetings	• To help the project team stay informed on project progress; • Identify relevant barriers; • Inform stakeholders and discuss their input and concerns related to various project components.
Interviews with local Data management/Data science experts	Data management and data science team members	1 meeting each with 6 experts	In-person interviews	• To identify current practices and, work that has already been done in terms of central monitoring, • To identify needs and suggestions for improvement
Interviews with local Monitoring team	Monitoring team members	1-2 meetings with each monitor – 4 members	In-person interviews	• To discuss current practice • To identify needs and suggestions for improvement
Department of Clinical Research team meeting	Department of Clinical Research (all divisions including the Data Science, Data management, Monitoring)	2 meetings with 35-40 persons	Presentation	• To verify if all needs and concerns of the stakeholder groups have been considered in the concept development

Stakeholder activity	Participants involved	N	Mode of stakeholder activities	Purpose of stakeholder activities
Meetings with local Principal investigators of clinical trials	2 Principal investigators of ongoing trials at the University Hospital Basel	Meeting once a month	In-person meeting	• To discuss needs and suggestions with principle investigators: What would help them in terms of management support and performance overview
National SCTO monitoring Platform	SCTO monitoring platform members	One meeting in Bern, email contact	In-person meeting, email correspondence	• To get input on the concept and discuss needs and suggestions with monitoring teams in Switzerland.
International Clinical Trials Methodology Conference	International Methodology Research Community	3 day conference in Brighton with short oral & poster presentation and informal meetings	Presentation & In-person meeting	• To discuss the preliminary results of our literature review • To learn about evaluated risk indicators for trial and site performance.
International collaboration with trial monitoring experts at the MRC Clinical Trials Unit UCL	2 Senior statisticians from the MRC UCL	2 meetings, email correspondence over 2 years	In-person meetings, email correspondence	• Exchange about ongoing refinement of risk indicators, • To learn about central monitoring of clinical trials established in the UK

Stakeholder activity	Participants involved	N	Mode of stakeholder activities	Purpose of stakeholder activities
International collaboration with monitoring dashboard developers in Aberdeen	Senior statisticians	1 phone call, email correspondence over 3 months	Telephone call, email correspondence	• Exchange about study dashboard development and content

Abbreviations: SCTO, Swiss Clinical Trial Organisation; MRC UCL, Medical Research Council of the University College London

Supplementary Table 3: Overview of stakeholder input as part of the contextual analysis

Step of Development	Content	Structure and Design
Risk assessment	• Include patient safety as well as procedural risks endangering successful trial completion • Include the management of participant schedule per protocol (e.g. timeframe of visits) • Include Data collection and storage related items • Include handling of Investigational Medicinal Product handling • Include safety management (e.g. SAE reporting, status) • Include complex informed consent processes (Re-consent) • Consider complex sampling or imaging during conduct • Take into account additional factors like experience of staff, budget	• Structured into four domains of Participant Safety and Rights, Overall Study Management, Device/Medication Management, Study Data • Structured into risk elements (e.g. Participant Schedule) and assets (a standard requirement that provides the basis for safety and accuracy of a clinical trial, e.g. visits have to take place in the required timeframe) • Provide possible risk scenarios to better apply the assessment to a study • Document rationale for rating • Assess severity and likelihood of risks
Study Dashboard	• Need for standardized central data monitoring (**Supplementary Table 2:** Summary of current practice) • A strong need for assistance and overview of management elements essential for study conduct (e.g. overview of follow-up visits done,	• Drop-down menu to be able to choose single visits • Colour code status of queries should be equal to database • Colour code for problems to be addressed

Step of Development	Content	Structure and Design
	recruitment curves). • A need to assess compare study progress in each site • Visualize ongoing SAEs, status of SAEs, timely reporting of SAEs • Visualize form completeness of primary outcome • Visualize patients who need a re-consent • Include status of queries • A need for site-specific information based on the entered study data to assist and guide the on-site monitoring visits • Trigger (red value boxes) for immediate phone call or email reminder to enable early resolution of a problem independent of an on-site visit • Differentiate between participants ending the study with or without the primary outcome data collected • Show analysis for reasons of ending the study (retention)	immediately should be red • Listing the patient for whom a correction/action is needed and a link to the eCRF • Provide the option to choose patients in a specific randomization time-frame (e.g. during Covid-Pandemic) • Option to choose specific centres • Retention before and after primary outcome reached

Abbreviations: IMP, Investigational Medicinal Product; SAE, Serious Adverse Event; eCRF, electronic Case Report Form

Supplementary Table 4: Full risk assessment guide. Generic content that applies to all trials is marked in red. In order to make the risk assessment operational, refer to the manual.

Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
Participant Safety and Rights	Informed consent	Condition/Characteristics of subject population and complexity of informed consent process must be considered	(A) Subject population is vulnerable (emergency situation, children, patient not able to consent)	☐	☐	☐	
			(B) Multiple informed consent processes lead to delayed, incorrect informed consent process (e.g. pre-screening, sub-studies, Re-consent in case of next-of-kin consent)	☐	☐	☐	
		Latest version of Informed consent must be obtained and documented according to GCP guidelines	Patient is submitted to study procedure before the informed consent is obtained (e.g. Condition of patients or emergency situations aggravate timely informed consent process)	☐	☐	☐	
			(B) Informed consent is not signed	☐	☐	☐	

Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
			or dated correctly				
			(C) Older versions of informed consent documents, that do not include latest protocol amendments are signed by participants	☐	☐	☐	
	Inclusion/Exclusion	Safety relevant inclusion/exclusion criteria must be considered	Only a subgroup of the population is suitable for the study - safety relevant criteria for inclusion/exclusion	☐	☐	☐	
	SAE/AE	Causality and medical evaluation of Adverse Events have to be performed thoroughly by qualified staff	(A) Inexperience with study drug (outside authorized indication), drug dose or drug not tested in study population causes misjudgement of Adverse Events (e.g. Serious drug reaction and device effect are not considered for specific study population)	☐	☐	☐	
			(B) The potential for an interaction of basic or background therapies, prescribed, recommended or allowed by the protocol is not	☐	☐	☐	

Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
			considered in the evaluation of Adverse Events				
		SAE have to be reported and documented correctly in the required timeframe	Complexity of CRF or missing SOPs for SAE Reporting leads to (A) incorrect documentation and (B) delayed reporting of SAEs	☐	☐	☐	
Overall Study Management	Recruitment	Monthly recruitment should follow the recruitment schedule	A) Actual participant recruitment is distinctly lower than the estimated recruitment rate	☐	☐	☐	
			(B) An alternative treatment competes with participant recruitment	☐	☐	☐	
			C) Pre-feasibility assessment of the study recruitment is not based on reliable sources (clinical department activity, pre-screening registry, pilot study)	☐	☐	☐	

Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
		Condition of subject population must be considered recruitment estimations	Difficulty in obtaining informed consent because of the high morbidity of the patient	☐	☐	☐	
	Retention	Loses to follow-up have to be minimized and reasons have to be evaluated	(A) Long follow up times may lead to decreased number of participants (death, unwillingness)	☐	☐	☐	
			(B) Follow-up visits cannot be scheduled if prior visits are not reported in a timely manner - study schedule is delayed	☐	☐	☐	
		Condition of subject population must be considered	High drop-out rate because of high numbers of SAE due to severe condition of study population	☐	☐	☐	
	Study procedures and endpoint assessments	Randomization has to take place accurately and within the given Timeframe	(A) Randomization does not take place in the given Timeframe	☐	☐	☐	
			B) Randomization ID documentation is prone to error	☐	☐	☐	
		Interim Analysis has to take place within the given Timeframe	Interim Analysis difficult to coordinate (international, multicentre)	☐	☐	☐	

Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
		Procedures for Blinding and unblinding have to be clear	(A) Missing objectivity in the assessment of the primary and main secondary outcomes by unblinded outcome assessors	☐	☐	☐	
			(B) No adequate procedures for unblinding in place	☐	☐	☐	
		Complexity of study procedures must be considered and Study conduct must adhere to the protocol procedures	(A) Complexity of study design and treatment schedule increases the risk of non-adherence to the study protocol	☐	☐	☐	
			(B) Techniqual requirements - e.g. critical handling of samples/ new assessment tools	☐	☐	☐	
			(C) Trial specific knowledge or training required	☐	☐	☐	
		Study Procedures and Conduct must be documented accurately within the given timeframe	(A) Numerous source systems (Electronic and paper source systems)	☐	☐	☐	
			(B) Different/Additional data collected in CRF as described in trial protocol/ trial schedule	☐	☐	☐	
			(C) Risk for slow data entry in the database	☐	☐	☐	

Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
	Endpoint assessment	Complexity of primary endpoint must be considered	Complex assessment procedure are necessary to obtain the primary endpoint or assessment is not robust (subjective, unblinded, patient-reported)	☐	☐	☐	
	Participant Schedule	Visits/Phone calls must be within the given Timeframe	(A) Time point of visit is critical for the endpoint assessment of the study	☐	☐	☐	
			(B) Large number of visits are difficult to organize and coordinate between centres and patients	☐	☐	☐	
		Interventions/medication must be verified and within the given Timeframe	(A) Medication/Intervention at several time points during the day	☐	☐	☐	
			(B) Severe condition of patient	☐	☐	☐	
			(C) Daily medication not verified by a second person	☐	☐	☐	
		Concomitant therapy should be consistent over the conduct of the study	(A) Heterogeneity of participants morbidities that require different concomitant treatment	☐	☐	☐	
			(B) Heterogeneity in procedures between study sites	☐	☐	☐	
Device/ Medication	Storage/ Accountability	Drug supply must be guaranteed over the	(A) Drug supply channel's validation is not up-to-date, timely	☐	☐	☐	

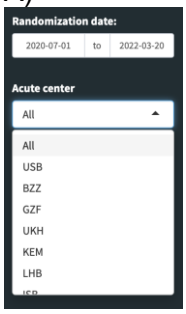
Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
Management		whole study period (Production Schedule/ Stock at cites/ Central)	supply of medication is endangered.				
			(B) Complex IMP shipping process	☐	☐	☐	
		Drug accountability has to be verified	IMP Handling/preparation/administration has potential for dosing errors, temperature deviations (e.g. self-administration)	☐	☐	☐	
	Administration	Correct IMP administration must be guaranteed and monitored	Complex IMP handling requirements (e.g. Temperature sensitive, small timeframe till expiration date of the medication (Short shelf-life))	☐	☐	☐	
Study Data	Data Quality	Study data has to be complete and up to date	(A) Complex assessment procedure are necessary to obtain the primary endpoint or secondary endpoint	☐	☐	☐	
			(B) Many data points have to be entered into the CRF	☐	☐	☐	

Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
		Information entered in the electronic CRF must be identical to the source data (Source Data Verification)	(A) No double-data entry implemented	☐	☐	☐	
			(B) No source data verification possible	☐	☐	☐	
		Study data must be accurate	Staff is not adequately trained in the generation or documentation of the study data	☐	☐	☐	
	Data storage	Study data must be stored in a safe place.	A) No backup, no audit trail --> loss of data	☐	☐	☐	
			B) Source data not locked away --> loss of Source data	☐	☐	☐	

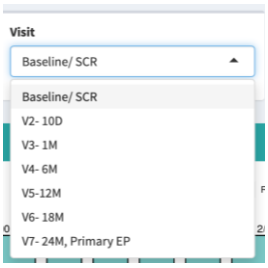
Abbreviations: AE, adverse events; GCP, Good Clinical Practice; IMP, investigational medicinal product; CRF, case report form; SAE, serious adverse events SOPs, standard operating procedures

Supplementary Figure 1: Example filter (Panel A), example drop down menu (Panel B)

A)

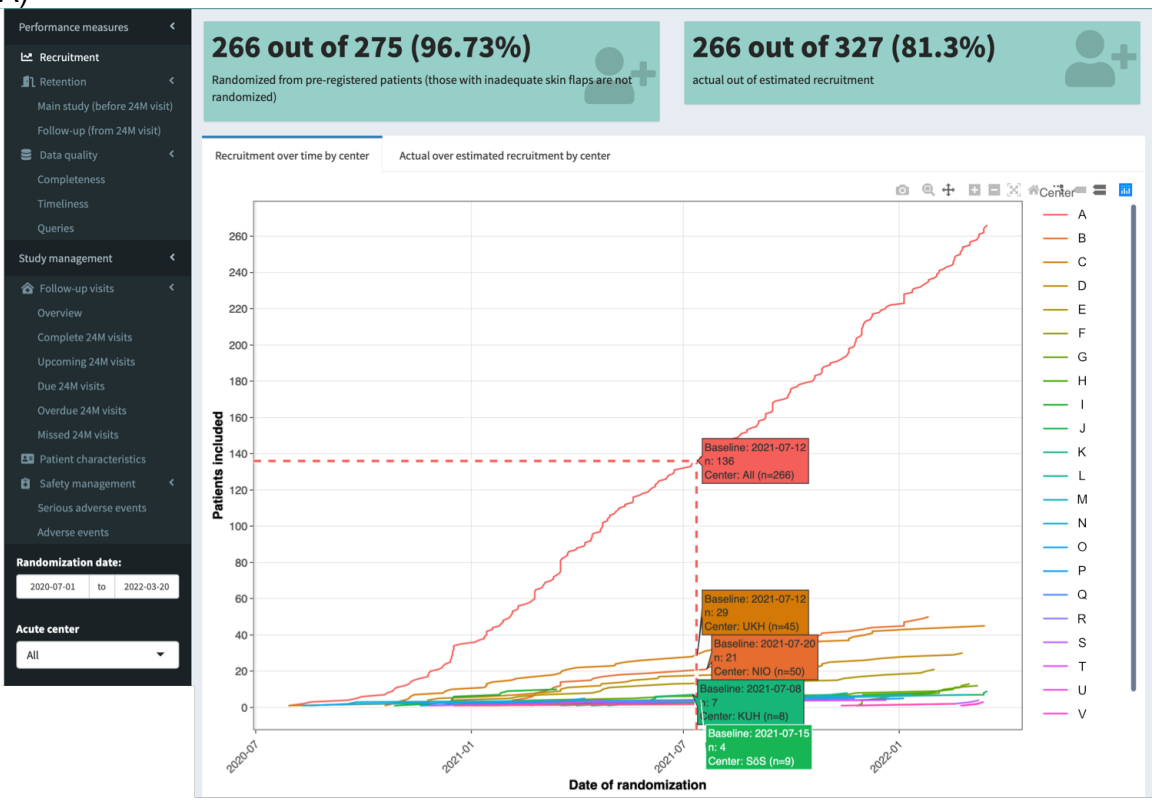


B)

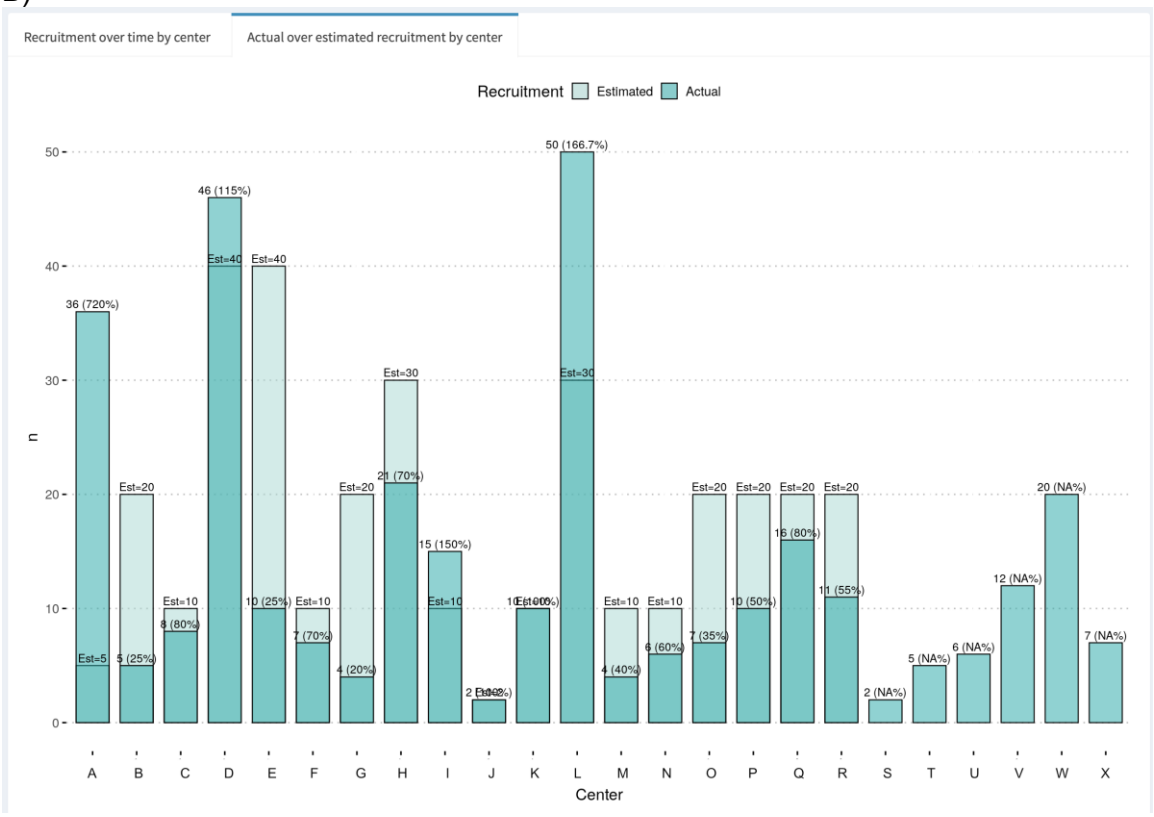


Supplementary Figure 2: Recruitment Dashboard – Development (Panel A) and Actual over estimated recruitment (Panel B)

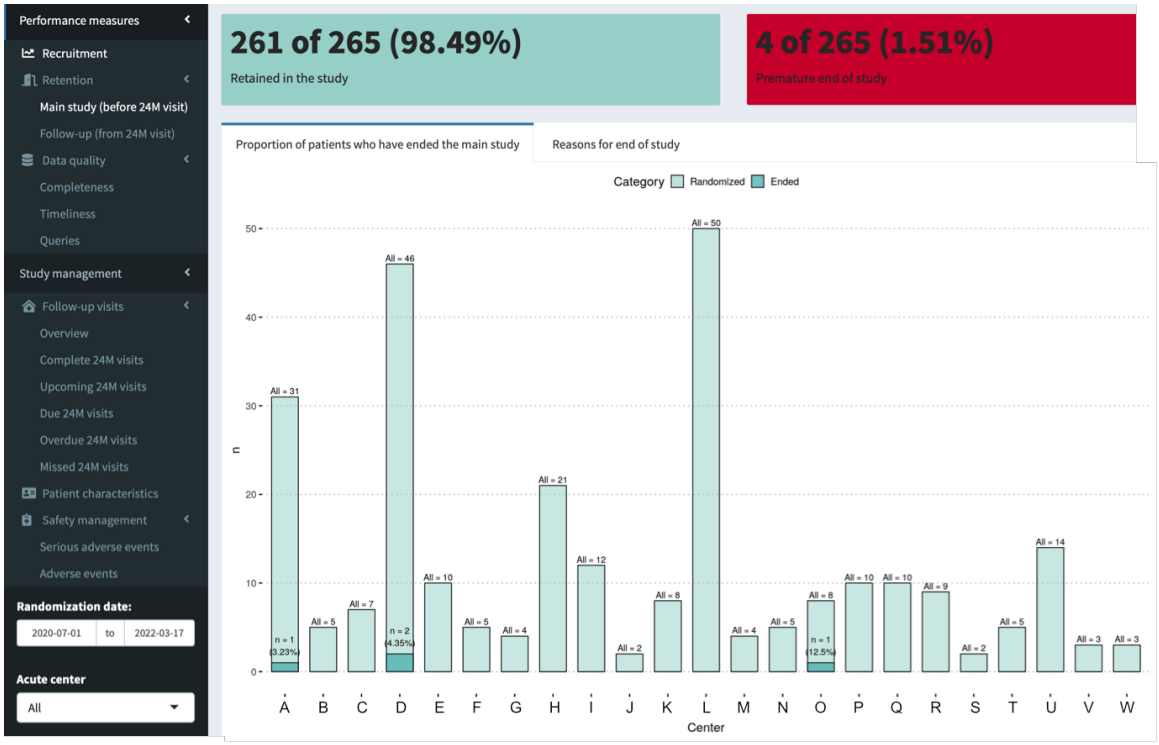
A)



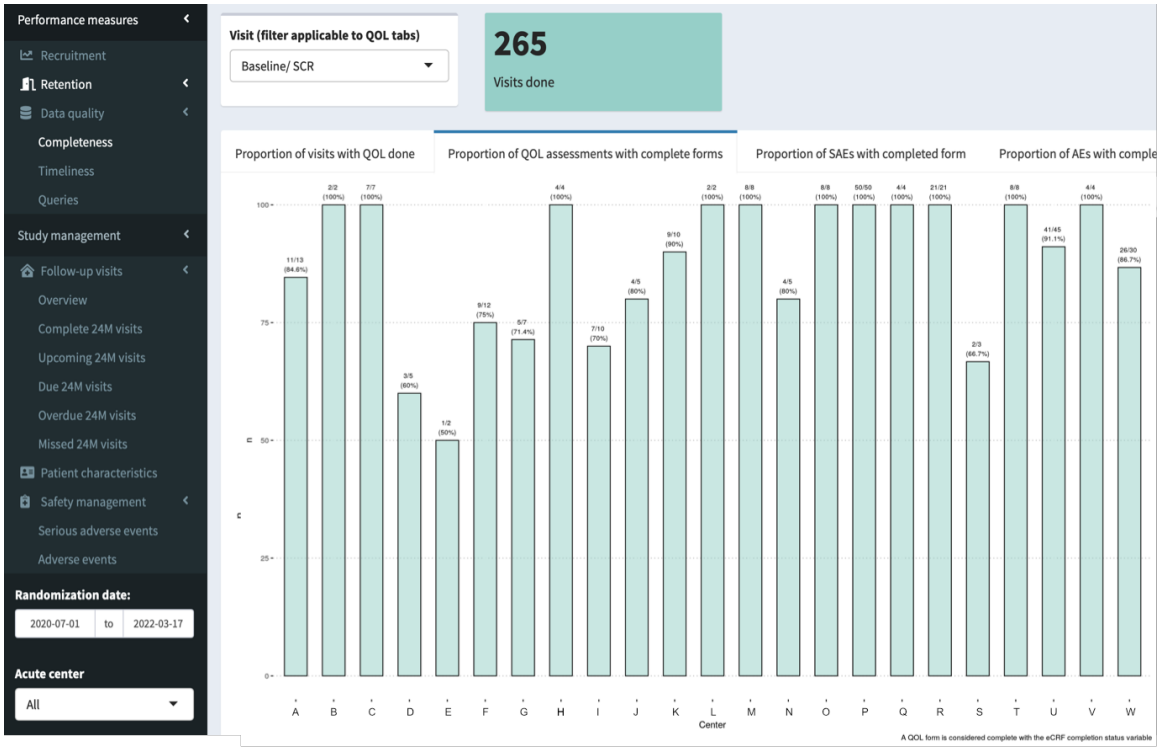
B)



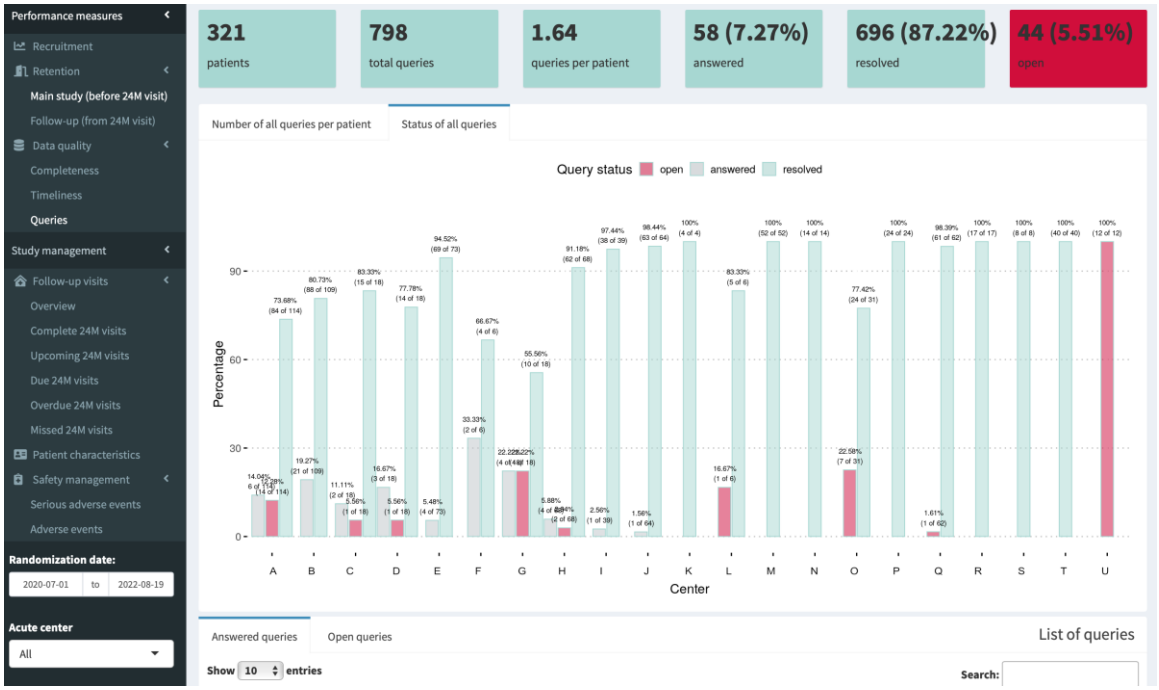
Supplementary Figure 3: Retention before Primary Endpoint Assessment



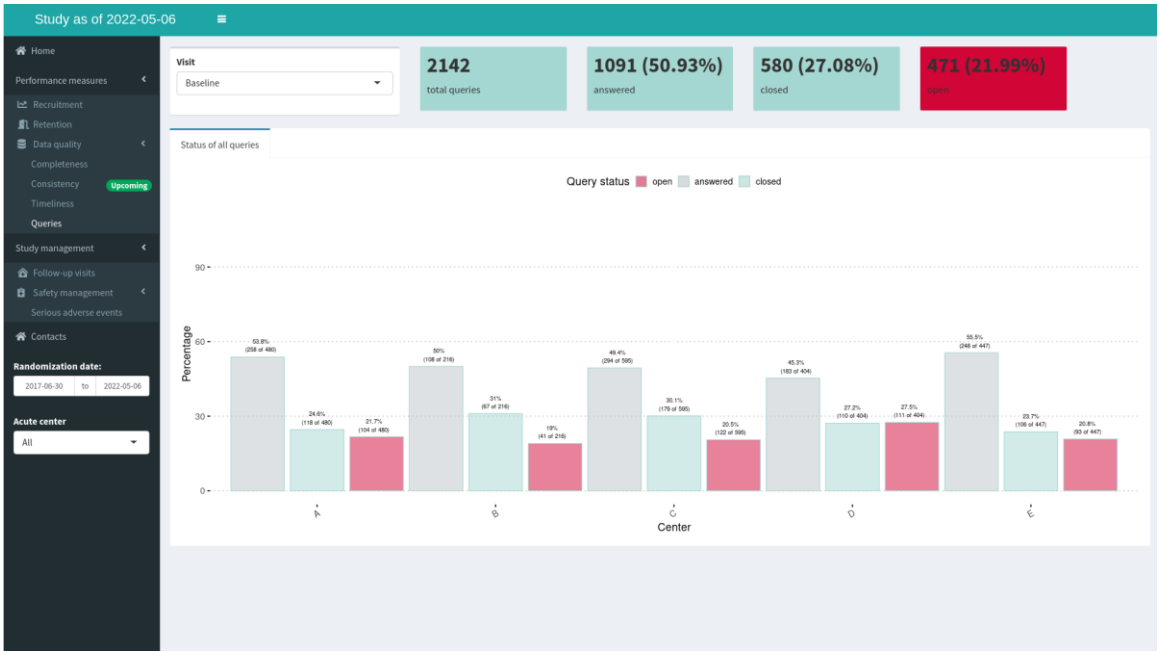
Supplementary Figure 4: Data Quality - Completeness of primary endpoint forms



Supplementary Figure 5: Status of queries



Supplementary Figure 6: Example of a query status tab generated from the generic code available on Github.



Supplementary Figure 7: Example of a safety management tab generated from the generic code available on Github.

Supplementary Table 5: Summary of results from the user testing

Position	General Feedback			Specific feedback		
	Content Risk assessment/dashboard (Positive and negative comments)	Overall Rating 0-100*	Suggested Applicability	Structure and Design	Additional Tabs	Further suggestions for improvement
Trial monitors (n=2)	Dashboard should visualize what the monitoring does not cover. Taking monitoring plan into consideration.	60-80	Dashboard for studies with 3 or more sites	Queries – Resolved – should be colour coded in yellow to resemble database colour code Spacing numbers/sites	MONITORING tab Which patients have been monitored – partly monitored is the interesting variable CONSENT tab – Already program re-consent button into data base in case of amendments – re-consent – overview of which patient signed which version	Check if all visits have been done in the required timeframe- display of delta Verify entry, informed consent with delegation log
Trial managers (n=2)	- Preventive – When sites are primed and many problems are solved in advance of the on-site visit– - Increased awareness of data management (what do we need, what is missing)	65-85	All studies, depending on the complexity of the intervention and	Like driving a car – one red lamp – auto repair shop All red – not able to drive	VISITS Include all visits, not just primary outcome visits ITEM INTERVENTION (E.g. early discharge needs special attention)	SAE – include narrative of comment field Login once in secutrial – direct connection to sheet

Position	General Feedback				Specific feedback	
	Content Risk assessment/dashboard (Positive and negative comments)	Overall Rating 0-100*	Suggested Applicability	Structure and Design	Additional Tabs	Further suggestions for improvement
	- Improves communication – option for compliments needed -Early awareness of data/ intervention issues - Positive feedback for study coordinator (It is reassuring when the Incomplete box shows 0 patients) - Continuous analysis of SAEs (how many, status) - Overview on data completeness - “Like Central Data cleaning” (which data is missing, identifying key elements of data sheets, number of outcomes analysable)		infrastructure		Drug compliance Premature discontinuation OUTCOMES Include secondary outcomes (e.g. recurrent stroke) HEALTH ECONOMIC OUTCOMES (E.g. length of stay Important for Biomarker SAMPLE MANAGEMENT Biomarker status and results (Tracking of sample status) INFORMED CONSENT Displaying which protocol and consent versions were effective	of interest (e.g. SAE form) App cleaning: differentiate between nothing can be done/ problem that needs to be solved - Corrected lists or colour coded list - Better overview Differentiate between patients died/ withdrawn Should include more project management aspects – making Excel sheets superfluous Option for users to

Position	General Feedback				Specific feedback	
	Content Risk assessment/dashboard (Positive and negative comments)	Overall Rating 0-100*	Suggested Applicability	Structure and Design	Additional Tabs	Further suggestions for improvement
					at which time point in the study	enter data and comments to the dashboard)
Principal Investigators (n=2)	- Provides an overview over patients with incomplete endpoints, differentiate between outcome measures forever missed and outcomes that may still be possible to extract from source data – list of patients to contact centres - Overview of outcome measures for statistical calculations - 90 % useable outcome - Identify issues that are more prevalent in specific centres. - Very efficient in terms of data completeness - Provides a systematic overview of patient recruitment - Filters are very useful, e.g. show only patients randomized during the COVID pandemic	95	All studies (Dependent on cost benefit relation) Useful for registers/ cohort studies	Very well designed	Secondary outcomes / other variables / Patient characteristics	Provide basic package for studies to choose the amount and area of support needed

* Rating from 0 to 100, 100 represents the best evaluation

Risk-tailored Approach for Efficient Management and Monitoring of Investigator-Initiated Trials

MANUAL

Intention

The main focus of the risk-tailored approach for efficient trial management is to support the conduct of clinical trials by providing a continuous oversight on the correct implementation, progress and accuracy of the most important elements of a clinical trial. We define an asset of a clinical trial as a standard requirement that provides the basis for safety and accuracy of a clinical trial. An example for an ASSET in a clinical trial could be the condition that the recruitment of participants should follow the expected recruitment rate or that visits must be scheduled and take place in the required timeframe.
In the first step, the developed risk assessment guide supports the Principle Investigator (PI) or project manager in the identification of the critical elements of an individual clinical trial. Factors that potentially impact the likelihood or severity of the risk have to be considered in the risk analysis step. Once study elements that are potentially at risk have been identified, CONTROL PATHWAYS are developed and implemented in order to provide a continuous oversight of these elements. The continuous overview of these elements supports the trial management and enables preventive measures and early interference. Oversight on data completeness, query status, SAE reporting and Informed consent status can also support and guide the on-site monitoring by prioritizing sites, and participants.

The risk-tailored approach consists of 4 steps, which may be repeated or updated throughout the study conduct:

(1) Identification of study-specific risks
(2) Analysis of the risks
(3) Development of control pathways
(4) Implementation of control pathways

Requirements and Basic Elements

In order to perform the risk assessment of a particular study, the following resources are required:

- Study Protocol
- Case Report Form (CRF) structure
- Established contact to responsible data manager of the study
- Overview of staff assigned to the study (e.g. Delegation Log, Curriculum Vitae and research experience of study staff)
- Information on the planned and actual budget of the study (Feasibility report)
- Expected recruitment for all participating centres
- Investigational Medicinal Product (IMP) information e.g. information about the handling of the IMP in terms of storage, transport and expiration date, information on supply chain (more than one supplier), organization of IMP distribution to participating centres.
- Established contact with the principal investigator

1st step: Identification of study-specific risks

Study elements are categorized in 4 domains, which represent different aspects of the clinical trial conduct:

I. Participant Safety and Rights

II. Overall Study Management (Study procedures, Participant Schedule)

III. Device/ Medication Management

IV. Study Data

Table 1: Risk elements of the 4 domains.

Domain	Risk Elements
Participant Safety and Rights	Informed consent AE/SAE reporting and documentation Inclusion/exclusion
Overall Study Management	Recruitment Retention Study procedures and endpoint assessment (e.g. bio sampling, imaging quality) Participant schedule (e.g. timeframe of visits) AE/SAE management
Device/ Medication Management	Administration Accountability/ storage
Study Data	Data quality – completeness, consistency, timeliness Documentation/ storage

Abbreviations: SAE, Serious Adverse Events

These domains contain risk elements covering most important elements of trial conduct. In order to identify the CRITICAL ASSETS of a trial, essential information has to be extracted and collected from the listed sources. To evaluate the applicability of an ASSET, the structure of the risk assessment guide includes one or more possibly applicable RISK SCENARIOS that endanger the ASSET or its accuracy (Example provided in **Figure 1**). In the process of identifying study specific risks, all listed ASSETS and the corresponding RISK SCENARIOS in the template are being evaluated with respect to the particular clinical trial. Every RISK SCENARIO that might apply to this trial should be marked (example shown in **Figure 2**).

Asset: Value of standard requirements that provide the basis for safety and accuracy of a clinical trial

Risk: Definition of risk including the source of the risk (Risk scenario)

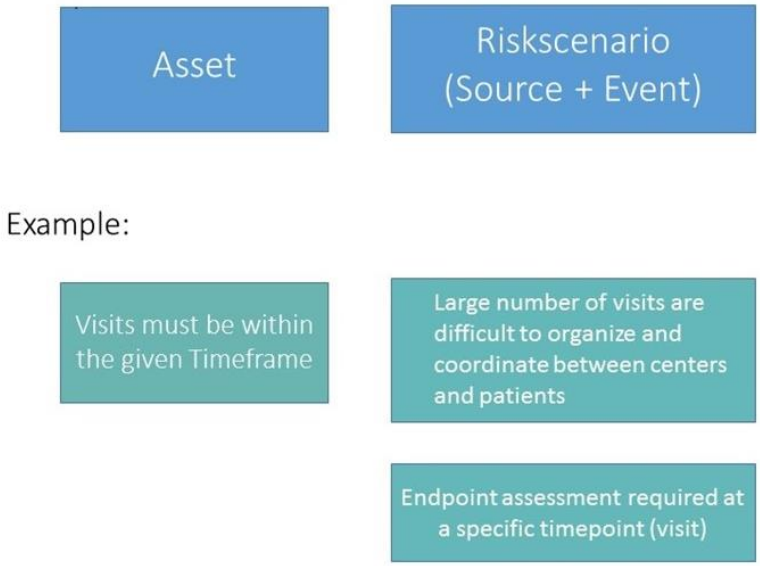


Figure 1: Example of RISK SCENARIOS for an ASSET

Domain	Risk element	Asset	Risk scenario	Promotor*	Likely	Critical	Rationale for Rating
Overall Study Management	Participant Schedule	Visits/Phone calls must be within the given Timeframe	(A) Time point of visit is critical for the endpoint assessment of the study	☐	☐	☐	
			(B) Large number of visits are difficult to organize and coordinate between centres and patients	☐	☐	☐	

Fig 2: Structure of Risk assessment –Example of an asset from the Overall Study Management domain.

* Check for applicability of one of the Promotors for the identified risk

(1) Experience with similar study or pilot study conducted - Similar study was successfully conducted in the same setting (similar infrastructure, Procedures, Intervention); (2) Well-trained, experienced, and dedicated principal investigators and study staff present Number of successful clinical trials completed by the PI/ Number of clinical trials supported by Study Nurse; (3) Adequate Budget – Feasibility report available.

2nd step: Analysis of risks

In order to estimate the risk for a specific scenario/asset, further information on the study may have to be obtained from the PI or study manager, e.g. information on the infrastructure of participating centres, actual budget etc., and considered in the analysis. The impact of the following three PROMOTORS will be assessed and considered in the LIKELIHOOD rating of each trial-specific RISK SCENARIO.

Promotors:

(1) Experience with similar study or pilot study conducted

- Similar study was successfully conducted in the same setting (similar infrastructure, Procedures, Intervention)

(2) Well-trained, experienced, and dedicated principal investigators and study staff present

Define experience:

- Number of successful clinical trials completed by the PI
- Number of clinical trials previously supported by designated study staff

(3) Adequate budget

- Check with estimations in the budget plan (large discrepancy vs. small discrepancy)

Likelihood:

The evaluation of the likelihood of a risk scenario for most of the assets is highly influenced by the applicability of the three promotors. If the budget of the study is adequate and as planned as well as well trained, experienced staff is present or a similar study has already been conducted in the same setting, the likelihood for the risk scenario decreases.

Likelihood description: Likely (including possible, almost certain) vs. Unlikely (including rare)

The estimation of the likelihood must be updated regularly during trial conduct, e.g. in case of staff fluctuations or changes in the funding/budget situation

The SEVERITY will be mainly influenced by the CONSEQUENCES of the specific RISK SCENARIO on the overall study.

Consequence: The impact of the risk when the critical asset is not met

Severity: Assessing the question: How critical is the consequence for the overall study conduct, study outcome, and patients' rights and safety. In order to categorize consequences and enable an analysis of the risk, critical and non-critical consequences are defined according to the classification of Good Clinical Practice /GCP) findings described in "Procedure for reporting of GCP inspections requested by the Committee for Medicinal Products for Human Use (CHMP)" by the European Medicines Agency in 2017.
(Available under https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/ins-gcp-4-procedure-reporting-good-clinical-practice-inspections-requested-chmp_en.pdf).

Definition for critical: Conditions, practices, or processes that endanger the rights, safety or well being of the participants or the protocol-conform collection of outcome data.
Possible consequences: Rejection of data and/or legal action is required
Definition for non-critical: Conditions, practices, or processes that deviate from the planned conduct, but are not expected to endanger the safety or well being of the participants or the protocol-conform collection of outcome data.
Possible consequences: Data might be rejected or the duration of the study might be extended. Non-critical applicable risk scenarios still indicate the need for improvement of conditions, practices and processes.

For each applicable ASSET, a rating will be performed for the SEVERITY and LIKELIHOOD based on the CONSEQUENCES of the RISK SCENARIO and the applicability of PROMOTORS for this ASSET. Other factors such as experience with a similar intervention or process might also influence the rating; therefore it is important to provide a short rationale for the rating. All assets that are rated "likely" and/or "critical" will then be included in the

development and implementation of control pathways. Examples of the evaluation of assets are provided in **Table 2**.

Table 2: Example of a risk assessment and risk analysis of four different assets

Asset	Risk Scenario	Promotor	Likely	Critical	Rationale for Rating
Study data has to be complete and up to date	Incomplete Data and incorrect Data transfer - Compromising the assessment of the primary and secondary endpoints	None applicable	■	■	Endpoint is assessed at Rehabilitation centre, slow or incomplete data transfer (input of many data points into CRF) is likely
Monthly recruitment should follow the recruitment schedule	Insufficient participant recruitment - Extended study duration and increased costs	Inexperienced centres included	■	■	Many smaller inexperienced centres and a competing alternative treatment
Visits/phone calls must happen within the given timeframe	Impacts the analysis of the primary endpoint	No study nurse present	■	■	Primary endpoint is assessed at 3-month visit, which takes place at the rehabilitation centres (no experienced staff present).
Complexity of primary endpoint must be considered	Validity of primary outcome is low	Second blinded assessment by experienced PI	□	■	Endpoint assessment is complex but staff is specially trained and there is a second blinded assessment by the PI.

Abbreviations: PI, Principle Investigator, CRF, Case Report Form.

3rd step: Development of Control Pathways

All ASSETS and the corresponding RISK SCENARIOS will be included in the development of PATHWAYS. PATHWAYS are a set of operations applied to the data set (collected through the CRF) and supporting information (expected recruitment, shipping information etc.) and will generate an OUTPUT that will be visualized (Graph/color-coded panels) for clinical monitors and study staff in the dashboard application. An example for a PATHWAY is provided in **Table 3**.

Table 3: Example of pathways for Likely and/or Critical Risk scenarios

Risk scenario	Pathway	Dashboard ID
Time point of visit is critical for the endpoint assessment of the study	Calculation of visit plan based on randomization date of patient -> Continuous overview of upcoming visits (possibility for reminders) Due visits (+/- 14 days) – number of patients (per centre) and list of patients Overdue visits (+/- x days - depending on the timeframe stated in the study protocol) - number of patients (per centre) and list of patients Missed visits (out of acceptable timeframe) – number of patients (per centre) and list of patients	FOLLOW-UP VISITS

4th step: Implementation of Control Pathways

Dashboard Approach:

Each pathway is assigned to a DASHBOARD ID/TAB. The visualization of the control pathway OUTPUT is accessible under this ID/TAB in the dashboard. The OUTPUT is based on daily data exports from the trial database. Examples of the visualized OUTPUT in the dashboard are provided in **Figure 2** and **Figure 3**. A basic structure of the pathways of generic Tabs is provided in form of modules (mod) in the Github environment

(https://github.com/CTU-Basel/viewTrial) and listed in **Table 4**. The template structures can be downloaded and adapted to the specific database format and structure of exported tables containing relevant variables.

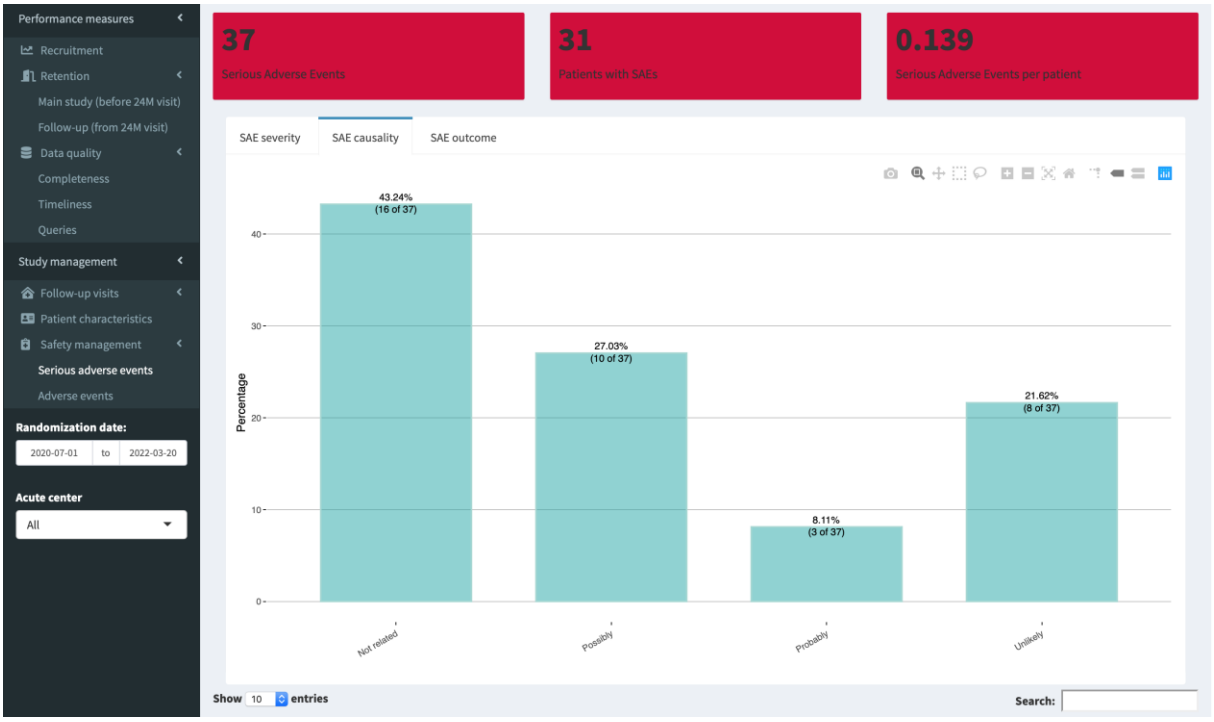


Fig 2: Example of the visualization control pathway OUTPUT in the Safety management TAB

Fig 3: Example of the visualization control pathway OUTPUT in the Visits Tab

Table 4: List of Dashboard ID/Tabs of the generic dashboard template

TAB	Basic pathways structure (modules, mod) provided in the template
Recruitment	Mod_recruitment provides a recruitment plot together with two information boxes
Recruitment	Mod_retention provides details on number and reason of loss to follow up
Data quality	Mod_completeness provides an example of how data completeness might be shown Mod_timeliness provides an example of how time between events and their entry into the database might be shown Mod_queries provides an example of how number of queries and query status might be shown
Follow-up visits	Mod_fup provides an example of how tracking of participant progress through a trial might be shown
Safety management	Mod_sae shows counts of SAE and characteristics of reported SAEs

2.3 Manuscript III: Towards full clinical trial registration and results publication: longitudinal meta-research study in Northwestern Switzerland

Katharina Klatte[1], Constantin Sluka[1], Viktoria Gloy[1], Ala Taji Heravi[1], Christof Schönenberger[1], Nienke Jones[2], Elena Brunnschweiler[2], Christiane Pauli-Magnus[1]*, Matthias Briel[1,3]*

[1] Department of Clinical Research, Meta-Research Centre, University Hospital Basel and University of Basel, Basel, Switzerland

[2] Ethics Committee Northwest and Central Switzerland, Basel, Switzerland

[3] Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

* shared senior authorship

Corresponding author

Katharina Klatte

Department of Clinical Research, University Hospital Basel and University of Basel, Spitalstrasse 12, CH- 4031 Basel, Switzerland

Telephone: +41 61 328 50 92

Email: Katharina.Klatte@usb.ch

Abstract

Objective: The registration of clinical trials is required by law in Switzerland. We investigated (1) the proportion of pro- and retrospectively registered clinical trials, (2) the availability of results for ethically approved trial protocols, (3) factors associated with increased registration, and (4) reasons for non-registration.

Design and Setting: We included all clinical trials with mandatory prospective registration, which were approved by the ethics committee of Northwestern and Central Switzerland between January 1, 2016, and December 31, 2020.

Methods: We extracted relevant trial characteristics from the Swiss Business Administration System for Ethics Committees and systematically searched the International Clinical Trials Registry Platform and primary trial registries for corresponding registry entries. We used multivariable logistic regression to examine the association between trial characteristics and registration. We qualitatively assessed reasons for non-registration of trials through an email questionnaire for trial investigators.

Results: Of 473 included clinical trials, 432 (91 %) were registered at all and 371 (78%) were prospectively registered. While the percentages of registration and prospective registration of investigator-sponsored trials increased from 85% to 93% and from 70% to 81% over five years, respectively, industry-sponsored trials consistently remained at a high level of prospective registration (92% to 100%). Trials with multiple centres, higher risk category, or methodological support from the local clinical trials unit were independently associated with increased registration rates. Of 103 clinical trials completed before August 2020, results were available for 70% of industry-sponsored trials and 45% of investigator-sponsored trials as peer-reviewed journal publications or in trial registries. Most common reasons for non-registration provided by investigators were lack of time or resources (53%), lack of knowledge (22%), and lack of reminders by the ethics committee (36%).

Conclusions: In Northwestern and Central Switzerland about 10% of clinical trials remained unregistered despite the obligation by law. More support for investigators and stricter enforcement by regulators are needed to improve the transparency of investigator-sponsored trials in particular.

Background

Trial registries create a public record of all planned, ongoing, and completed clinical trials. Hereby, clinical trial registries help to detect unnecessary duplication of research and publication bias.[1] Through prospective documentation of important trial characteristics such

as the primary outcome, eligibility criteria, or planned sample size trial registration further helps to minimize selective outcome reporting, 'spin', or other bad research practices.[2-5] Registration of all clinical trials as well as their timely publication is an important aspect addressing the need for transparency in clinical research [6] and constitutes a big step towards "Open Science".[7-9] In 2004, the International Committee of Medical Journal Editors (ICMJE) recommended publishing trial reports only if the trial was registered. [10] The World Medical Association included a statement in the Declaration of Helsinki that "every research study involving human subjects must be registered".[11] Further, the Federal Drug Administration (FDA) expanded their "Final Rule" upon the requirement with additional data elements for both registration and results submission records in 2017.[12] In Switzerland, prospective registration of a clinical trial in a primary trial registry has been made mandatory by law in 2014 (Art 56 Human Research Act).[13]

Various studies have already examined trial registration and, in particular, prospective trial registration based on published randomized controlled trials (RCTs)[3] [14-18], with prospective registration rates of RCTs ranging from 61 % in 2007 to 83% in 2012.[18] In agreement with these findings a recent systematic review of clinical trials published in major respiratory journals between 2010 and 2018 found a positive trend for prospective trial registration - from 75% in 2010 up to 100% in 2018.[19] However, the group of published trials does not comprise all trials approved by an ethics committee and, therefore, the generalizability of these findings is still limited. An international meta-research study of 326 RCT protocols approved in 2012 found that one in five trials (70/326) remained unpublished at 10 years follow-up, and 21% of those unpublished trials (15/70) were not registered, i.e. they remain undetectable for the research community and the public. [20] Furthermore, an analysis of trials, required to register under the Food and Drug Administration Amendments Act (FDAAA) of 2007, by the "Trials Tracker" initiative revealed in 2020 that only 41 % of trials from all sponsors have reported their results at clinicaltrials.gov one year after trial completion.[21] However, the sensitivity of such automated search processes for trial results has not been examined yet in a local context beyond specific registries such as clinicaltrials.gov [22-24] or European Union Drug Regulating Authorities Clinical Trials (EudraCT) [25] .

In view of these findings, further action is needed to increase compliance with registration and publication requirements to improve clinical research transparency and, hereby, promote public trust. Having a national law in place that mandates prospective trial registration is an important step, however, it needs to be implemented and enforced in local research environments to achieve its intended purpose. We, therefore, investigated in close collaboration with the local Ethics Committee of Northwestern and Central Switzerland (EKNZ) (1) the proportion of registered and prospectively registered clinical trials and (2) the availability of trial results for protocols approved between 2016 and 2020, (3) factors

associated with trial registration rates including the use of methodological support provided by the Clinical Trials Unit (CTU) at the University Hospital Basel, (4) the sensitivity of automated publication tracking through the “Trials Tracker” approach in Northwestern and Central Switzerland, and (5) reasons for non-registration.

Methods

Study sample

Since January 1, 2016, it is mandatory to submit all study protocols for approval to a research ethics committee centrally via the Business Administration System for Ethics Committees (BASEC) in Switzerland. In the present study, we included all studies that were (1) classified as clinical trials (ClinV, clinical intervention studies) in BASEC and (2) approved by the EKNZ between January 1, 2016, and December 31, 2020.

Data collection

For all included trials we extracted relevant characteristics such as number of intervention arms, sponsorship, and target sample size from BASEC. Using a provided registry number, the study title, patient population, intervention, or specific outcomes we systematically searched the ICTRP of the World Health Organization (WHO) for corresponding registry entries of all included studies. We used a cloud-based database for data collection (squiekero). Two trained researchers performed the registry search and data extraction for each included study independently and in duplicate. Disagreements were resolved by discussion and consensus. If a registry entry could not be found for a trial on ICTRP, we consecutively searched primary registries such as clinicaltrials.gov and EudraCT, and finally conducted a google search. The last search was carried out on April 21, 2021. For all studies for which we could not identify a registration entry through electronic searches, we surveyed corresponding trial investigators (documented in BASEC) for further information about trial registration. If investigators provided a registration number until Sept 1, 2021, corresponding trials were classified as registered in our data set. If contacted investigators did not provide a valid registration number for a trial, we eventually considered that trial not registered. From identified registry records we extracted further trial information such as date of registration, date of first patient enrolled, actual sample size, status of the trial, and sponsorship. If registration occurred before or within 30 days of enrolment of the first patient, we classified the trial as prospectively registered. All trials registered after 30 days from enrolment of the first patient provided in the registry were classified as retrospectively registered. To inquire about reasons for non-registration we sent a questionnaire to all principal investigators of

studies not registered at the time of data extraction via email. In addition, the questionnaire aimed to assess investigators' awareness of trial registration obligations, and to explore obstacles for trial registration (see supplementary material for full questionnaire). Responses of investigators providing a registration number in the questionnaire and considered registered in the qualitative analysis (n=19) were still included in the analysis of quantitative outcomes (41 non-registered trials, 19 registered trials). Clinical trials making use of CTU services were identified by systematically searching internal CTU files containing meta-information of all CTU-supported studies and checking for BASEC ID numbers.

Data analysis

Quantitative data about trial registration, prospective trial registration, results publication, and reasons for non-registration were summarized as frequencies and percentages, stratified by sponsorship (industry- vs. investigator-sponsored). We conducted univariable and multivariable logistic regression analyses with trial registration as dependent variable and sponsorship (industry- vs. investigator-sponsored), multicenter vs. single center trials, risk category of trial (low, medium, high), and use of CTU services (yes vs. no) as independent variables. We hypothesized that industry-sponsorship, multicenter trials, higher risk category, and use of CTU services, were associated with higher prevalence of trial registration and prospective trial registration. For all regression models, we calculated unadjusted and adjusted odds ratios (ORs) with 95% confidence intervals (CIs), and p-values. We evaluated the sensitivity of automated processes as used with the "Trials Tracker" by comparing our findings on results publication with an automated process based on a registry such as Clinicaltrials.gov. All quantitative analyses were conducted using R version 3.5.3. We qualitatively analysed open-ended questions about reasons for non-registration using thematic analysis.[26]

Results

Study sample characteristics

Of 473 clinical trials approved by the EKNZ between 2016 and 2020, 342 (72.3%) were investigator-sponsored and 323 (68.3%) used a randomized design (**Table 1**). The median planned sample size for Switzerland was 32 participants (interquartile range [IQR] 16 to 75). 218 studies (46.1%) were multicentre of which most were international (78.9%; 172/218). Approximately half of the trials were classified as low risk according to the Swiss Human Research Act.

Table 1: Characteristics of included clinical trials

Characteristics	Categories	All trials (n=473)	Investigator-sponsored trials (n=342)	Industry-sponsored trials (n=131)
Target sample size in Switzerland (median, IQR)		32 (16-75)	45 (22-100)	15 (8-28)
Trial intervention, n (%)	Drugs	215 (45.5)	116 (33.9)	99 (75.6)
	Medical devices	96 (20.3)	69 (20.2)	27 (20.6)
	Behavioral	33 (7.0)	33 (9.6)	0 (0.0)
	Diagnostic	28 (5.9)	26 (7.6)	2 (1.5)
	Rehabilitation	23 (4.9)	23 (6.7)	0 (0.0)
	Dietary supplements	18 (3.8)	18 (5.3)	0 (0.0)
	Surgical	16 (3.4)	15 (4.4)	1 (0.8)
	Other**	44 (9.3)	42 (12.3)	2 (1.5)
Trial design, n (%)	Single arm	121 (25.6)	78 (22.8)	43 (32.8)
	Multiple arms***	352 (74.4)	264 (77.2)	88 (67.2)
	Randomized	323 (68.3)	239 (69.9)	84 (64.1)
	Non randomized	29 (6.1)	25 (7.3)	4 (3.1)
Risk category*, n (%)	Low risk	238 (50.3)	220 (64.3)	18 (13.7)
	Intermediate risk	83 (17.6)	71 (20.8)	12 (9.2)
	High risk	152 (32.1)	51 (14.9)	101 (77.1)
Trial sites, n (%)	Single Center	255 (53.9)	243 (71.1)	12 (9.2)
	Multicenter	218 (46.1)	99 (28.9)	119 (90.8)
	National	46 (9.7)	43 (12.6)	3 (2.3)
	International	172 (36.4)	56 (16.4)	116 (88.5)
Use of CTU service, n (%)		104 (22.0)	104 (30.4)	0 (0.0)

Abbreviation: IQR - interquartile range 25% percentile - 75% percentile; CTU – Clinical Trials Unit

* Classification of studies in the Human Research Act: Category A – low risk for trials with products authorized in Switzerland, and used according to Swiss Summary of Product Characteristics; Category B - intermediate risk for trials with products authorized in Switzerland, not used according to Swiss Summary of Product Characteristics; Category C - high risk for trials with products not authorized in Switzerland. Intermediate and high risk categories require additional authorization by federal authority (Swissmedic) [27]

** Includes: exercise trials, physiotherapy, transplant products, PK/PD safety trials, radiation therapy, palliation, other diet trials

*** Includes cross-over (n=71), parallel group (n=278), factorial (n=3)

Of all 473 clinical trials, 432 (91.3 %) could be identified in a primary registry either via our sensitive search strategy or by contacting the investigators directly (**Table 2**). Of the 427 registered trials for which a registration date and a start date for participant recruitment were available, 371 (78.4 %) trials were registered prospectively. Prospective registration was more prevalent in industry-sponsored trials than in investigator-sponsored trials (93.1 % vs. 72.8 %, Table 2). Over the observation time of five years, there was a trend of increasing registration in investigator-sponsored trials with an increase in prospective trial registration from 69.7 % in 2016 to 81.2 % in 2020 **(Figure 1, Panel A)**, while industry-sponsored trials remained a high level of registration throughout (**Figure 1, Panel B**; without stratification see Supplementary Figure 1).

Table 2: Registration status of EKNZ approved clinical trials 2016-2020

Registration status	**All trials (n=473)**	**Investigator-sponsored trials (n=342)**	**Industry-sponsored trials (n=131)**
Registered (n, %)	432 (91.3 %)	306 (89.5 %)	126 (96.2 %)
Prospectively*	371 (78.4 %)**	249 (72.8 %)**	122 (93.1 %)
Retrospectively	56 (11.18%)**	52 (15.2 %)**	4 (3.0 %)

* Before or within one month (-30 days) of first patient enrolled

** 5 studies without date of first patient enrolled

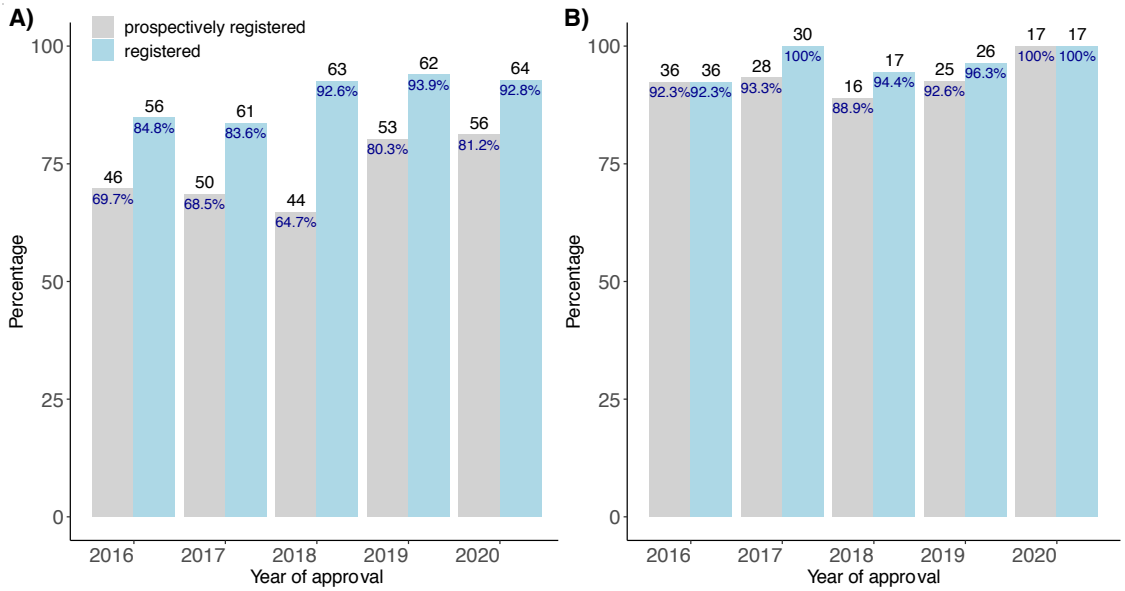


Figure 1: Percentage of clinical intervention studies registered and prospectively registered from 2016 to 2020 stratified by sponsorship. Panel A: Investigator-sponsored studies, Panel B: Industry-sponsored studies.

Trial characteristics associated with registration

Table 3: Associations between trial characteristics and registration status in logistic regression

Characteristics*	Registered n= 432	Non-registered n=41	Univariable			Multivariable		
			OR	95% CI	p-value	OR	95% CI	*p*-value
Single center (vs. multicenter)	219 (85.9%)	36 (14.1%)	0.14	0.05-0.34	<0.001	0.20	0.064-0.60	0.003
Investigator (vs. industry) sponsorship	306 (89.5%)	36 (10.5 %)	0.34	0.11-0.81	0.026	1.66	0.64-7.59	0.42
Risk category low	201 (84.5%)	37 (15.5%)	Reference			Reference		
Risk category intermediate	81 (97.6%)	2 (2.4%)	7.45	2.21-46.52	0.006	5.26	1.74-37.54	0.026
Risk category high	150 (98.7%)	2 (1.3%)	13.81	4.14-85.74	<0.001	9.00	2.56-71.32	0.008
Use of CTU service (vs. no service)	103 (99.0%)	1 (1.0%)	12.52	2.67-223.52	0.013	15.63	3.24-281.23	0.007
*Reference values: sample size <100, multi-center trials, investigator-initiated trials and drug trials.								

Abbreviations: OR, odds ratio; CI, confidence; CTU, Clinical Trials Unit.

We found that higher risk categories (intermediate and high), multicenter studies, and use of CTU services were independently associated with increased study registration (**Table 3**). We found similar results for prospective registration (**Supplementary Table 1**).

Availability of trial results

Of 103 registered clinical trials with a completion date before August 2020, 58 (56.3%) had publicly available results until September 2021; in 51 trials (49.5%) results were published in a peer-reviewed journal, 16 (15.5%) trials provided results via a trial registry, and 7 trials did both. Of the 51 journal publications, 29 (56.9%) explicitly reported the registration number (**Table 4**). The percentage of reported trial results at 12 months after study completion was 69.2% for industry-sponsored trials, and 45.4% for investigator-sponsored trials; 53.8% of industry-sponsored trials reported results in a trial registry versus 2.6% of completed investigator-sponsored trials.

Table 4: Availability of results in completed clinical trials

Completed registered studies*(n)	**All trials (n=103)**	**Investigator-sponsored trials (n=77)**	**Industry-sponsored trials (n=26)**
Publicly available trial results, 12 month after study completion	58 (56.3 %)	40 (45.4 %)	18 (69.2 %)
Publication of results in registry	16 (15.5 %)	2 (2.6 %)	14 (53.8 %)
Publication in peer-reviewed journal	51 (49.5 %)	38 (49.4 %)	13 (50 %)
Journal publication mentioned Registration Number**	29 (56.9 %)	18 (47.4 %)	11 (84.6 %)

*Completed by August 2020 according to status of study provided in the registry

** Percentage of journal publications

With respect to the sensitivity of an automated approach searching for trial publications such as the “Trials Tracker”, we noted that an automated approach searching primary registries does not consider non-registered studies; in our sample, 41 of 473 studies (8.7 %) approved by the EKNZ could not be identified in any primary registry. If the automated approach considers clinicaltrials.gov only (Trials Tracker), studies registered in other primary registries are missed. In our sample, 72.9 % (345/432) of registered studies were registered in clinicaltrials.gov, and 18.4 % (87/432) were exclusively registered in another primary registry (**Supplementary Figure 2**). Thus, 27.1 % (128/473) of studies would be missed through an

automated export from clinicaltrials.gov. Considering additionally that only 56.9 % of identified results publications explicitly mentioned the registration number, automated searching for the study registration number via PubMed likely misses a substantial number of study publications.

Reasons for non-registration and investigators' awareness of registration facts

Table 5: Survey of trial investigators with non-registered studies as of April 2020

Topic	**36 of 60 investigators filled out the questionnaire, 24 did not respond**			
General awareness of Researchers	Prospective registration is required by law	Registration is required before first participant enters the study	Swiss National Clinical Trials Portal (SNCTP)* is not a primary registry	Registration is reasonable
	27 (45.0 %)	26 (43.3 %)	14 (23.3 %)	32 (53.3 %)
Study support by service team	Clinical Trials Unit	Contract Research Organisation	Others	No support service
	7 (11.6 %)	3 (5.0 %)	2 (3.3 %)	24 (40.0 %)
Knowledge of primary registries	ClinicalTrials.gov	German Clinical Trials Register (DRKS)	EU Clinical Trial Register (EudraCT)	ISRCTN-Register
	32 (53.3 %)	13 (21.7 %)	14 (23.3 %)	7 (11.6 %)
Perceived Barriers to study registration	Insufficient knowledge of primary registries/ registration processes:	Limited time/ resources for registration process	Missing reminder of obligation to register the study	Others**
	8 (13.3 %)	18 (30.0 %)	13 (21.7 %)	6 (10.0 %)
Stated reasons for non-registration	- Study postponed/ unclear study start date (n=2) - Missing local SOPs for registration (n=1) - Unclear interpretation of regulations for Phase I studies (n=1) - Study not considered as clinical trial by investigator (n =4) - Unaware of the obligation to register (n=2) - Short study, retrospective registration considered as unnecessary/confusing (n=1) - One researcher responsible for all registrations in the research institute (n=1) - No reason specified (n=24)			

* In Switzerland every study approved by an ethics committee and registered in a primary registry will be listed on the Swiss National Clinical Trials Portal (SNCTP).

**Others included unclear definition of the study, unclear responsibilities for registration within institution, COVID-19 induced delay

Abbreviations: EudraCT, EU Clinical Trial Register; ISRCTN, International Standard Randomized Controlled Trial Number; SOP, Standard Operating Procedure; SNCTP, Swiss National Trials Portal.

In total, 36 out of 60 contacted investigators returned a filled questionnaire (60% response rate). 19 of the corresponding trials were eventually identified as registered through the questionnaire, while 41 remained in the non-registered group. Overall, 27 (45.0%) of contacted investigators were aware of the obligation to register a clinical trial, and 14 (23.3%) were aware that the Swiss National Trials Portal (SNCTP-KOFAM) is not a primary registry (**Table 5**). Most researchers stated to know one of the common primary registries. Of the suggested barriers in the registration process listed in the questionnaire, the most commonly stated barrier was lack of "Time and Resources" (30.0%), followed by "Missing reminder of obligation to register the study" (36.1 %). Most respondents did not take advantage of any CTU or CRO support services. Individual reasons for non-registration included researchers' view that their study was not a clinical trial and un-awareness of the obligation among others.

Discussion

Our empirical study of 473 clinical trials with mandatory registration found that registration and prospective registration increased for investigator-sponsored trials over time but still needing further improvement, while industry-sponsored trials had high registration levels throughout the five years of observation. Multicenter studies and studies in a higher risk category were associated with increased registration, probably reflecting more intense supervision / control of those studies. In addition, 99% of investigator-sponsored trials with CTU support were registered suggesting an effective process at the CTU to ensure trial registration. Overall, results were made available for 70% of completed industry-sponsored trials and 45% of investigator-sponsored trials. Only about 3% of completed investigator-sponsored trials had results published in a registry, whereas 54% of industry-sponsored trial results were available in registries. Automated tracking of results publications of approved clinical trials proofed challenging in our regional context due to a considerable proportion of unregistered trials, an appreciable distribution of trials registered in a number of different registries, and insufficient reporting of the registration number in trial publications. Reasons for non-registration provided by investigators included lack of time/resources, lack of knowledge, and lack of enforcement by ethics committees.

Strengths and limitations

The strengths of this study include a comprehensive sample of all clinical trials approved between 2016 and end of 2020 in the jurisdiction of the EKNZ and full access to all study information in BASEC. We conducted a sensitive search for registry entries supplemented by

a survey of investigators. We limited the number of variables in our regression models to reduce the probability of spurious associations. Finally, we complemented our quantitative analyses by a qualitative investigation of registration barriers.

Our study has the following limitations: First, our sample size was modest limiting the precision of stratified analyses over time. In some categories, for example industry-sponsored single center trials or industry-sponsored low risk trials, the sample size was very low. Second, only 36 of 60 contacted investigators of non-registered trials returned a filled questionnaire compromising our qualitative analysis and leaving the completion status for 24 trials unclear. Researchers responding to the survey may have a more positive view towards trial registration. Third, our sample was limited to trials approved by one Swiss ethics committee; therefore, our findings cannot be automatically extrapolated to other Swiss ethics committees or other countries.

Comparison with other studies

A recently published meta-research study found that 6% of RCTs from a sample of 326 RCT protocols approved in 2012 by research ethics committees in Switzerland, UK, Germany, and Canada were not registered, with non-registration being more common among non-published RCTs. [20] The proportion of prospectively registered RCTs was 84%, which is slightly higher than the proportion in our study sample (78.4 %). In our sample around 9 % of trials were not registered. A systematic review and meta-analysis published in 2018 found that in different medical specialties, 2-79% of RCTs were not registered[28], which shows a large variation depending on medical specialty. In addition, proportions of study registration may dependent also on the study sample (published, approved) and the countries involved. A recently published editorial by DeVito and Goldacre summarized the current trial reporting in the EU [29]; while progress has been observed in terms of trial results published [25], it is mainly driven by a few countries [30 311]. The different timeframes of the assessment also provide an explanation for the wide range of proportions found in different studies [20 28 32 33].

In agreement with our results a systematic review on clinical registration in major respiratory journals reported that single center studies were more likely to be retrospectively registered or not registered. [19] An analysis of clinical trials approved in Switzerland from 2016 -2020 showed that more than half of the trials were monocentric trials.[34] Since awareness and regulatory control might also be less in monocentric trials, education and support of the registration and dissemination processes for all research facilities in Switzerland should be aspired. In a survey of 149 researchers who had retrospectively registered a trial on ANZCTR between 2010 and 2015, the majority (56%) of survey respondents cited lack of awareness as a reason for not registering their study prospectively.[35] Seventy-four per cent

stated that linking registration to ethics approval would facilitate prospective registration. In a survey conducted by Mayo-Wilson et al. in the United States before the “The Final Rule” mandating trial registration, revealed that only a minority of academic organizations had policies and resources that facilitate clinical trial registration and reporting. They strongly suggest allocating resources to trial registration and reporting.[36] The medical university of South Carolina identified issues affecting their own compliance rate with FDAAA801 and evaluated newly implemented processes such as hiring a designated full time trial registration and reporting coordinator and a workflow that identifies trials early in the approval process requiring registration. Evaluation after 12 months demonstrated a marked increase to 98% over all compliance with the US federal regulations [37]. This is in agreement with our finding that the proportion of registration in trials with CTU service was 99%.

Besides the general obligation to register clinical trials and update registry information, a reliable linkage of publications to the registration number would increase the accuracy of automated processes that continuously provide information on trial result publication. Huser et al. also evaluated automated checking of trial registration ID in publications of five ICMJE founding journals, which revealed a registration in 88 % of cases [38]. We only found a registration ID in 57 % of trials published in journals. This difference is most likely explained by the sample of journals enforcing stricter rules for registration and including registration IDs. However, only looking at publication and their linkage to a registration number is not sufficient to identify trials where results are not available. Considering this limitation, the “Trials Tracker” initiative is now focusing on trials, which are required to report results on ClinicalTrials.gov or EudraCT and thus allow conclusive results to compliance of reporting on these platforms. In order to provide the complete content of research results of the scientific community, publication of results within the registries would reduce the likelihood of publication bias and spin [39 40]

Implications

Our study revealed encouraging results in terms of the development of registration rates over the last years, but further efforts are still needed. DeVito et al. propose that academic institutions should educate researchers about their responsibilities in terms of reporting and also ethics committees and funders should consider their responsibilities.[29] From our qualitative evaluation, a strong need for support in the registration process was identified and suggests that missing resources available for trial registration are often the reason for retrospective or non-registration. Education of investigators and support in the registration and publication processes would constitute important steps to more complete transparency of medical research. CTUs could catalyze these steps. Ethics committees may send email

reminders to trial investigators informing them about their legal obligations and prospective trial registration should be stricter enforced by publishing journals.[41]

Patient and citizen involvement in clinical trials has been shown to improve participation rates. [42] Also improving the reputation of clinical research in society is important. Stakeholders are requesting that all clinical trials should be routinely registered and lay-language summaries should be provided, as requested by the Swiss main public funding agency (Swiss National Science Foundation), to lower the barriers to patients and citizens being better informed and participating in clinical research. [43] During the Covid-19 pandemic, the importance of clinical research suddenly became publicly visible stressing the need for research transparency and availability of results. [44]

Conclusion

We have observed that rates of registration and prospective registration have increased in investigator-sponsored trials over the past years in Northwestern and Central Switzerland. Making study registration mandatory by law is an important step but not sufficient to achieve a 100% prospective registration rate. Further efforts in terms of law enforcement, education and local support of clinical researchers are needed. Monitoring of study registration and results publication is necessary to detect problems and further improve transparency in clinical research. Automated approaches have to consider local settings in order to achieve sufficient sensitivity.

References

1. Sim I, Chan AW, Gulmezoglu AM, et al. Clinical trial registration: transparency is the watchword. *Lancet* 2006;367(9523):1631-3. doi: 10.1016/S0140-6736(06)68708-4 [published Online First: 2006/05/23]
2. Mayo-Wilson E, Li T, Fusco N, et al. Cherry-picking by trialists and meta-analysts can drive conclusions about intervention efficacy. *J Clin Epidemiol* 2017;91:95-110. doi: 10.1016/j.jclinepi.2017.07.014 [published Online First: 2017/08/27]
3. Chan AW, Pello A, Kitchen J, et al. Association of Trial Registration With Reporting of Primary Outcomes in Protocols and Publications. *JAMA* 2017;318(17):1709-11. doi: 10.1001/jama.2017.13001 [published Online First: 2017/09/12]
4. Goldacre B, Drysdale H, Dale A, et al. COMPare: a prospective cohort study correcting and monitoring 58 misreported trials in real time. *Trials* 2019;20(1):118. doi: 10.1186/s13063-019-3173-2 [published Online First: 2019/02/15]
5. Chan AW, Hrobjartsson A, Haahr MT, et al. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291(20):2457-65. doi: 10.1001/jama.291.20.2457 [published Online First: 2004/05/27]
6. Chalmers I, Glasziou P, Godlee F. All trials must be registered and the results published. *BMJ* 2013;346:f105. doi: 10.1136/bmj.f105 [published Online First: 2013/01/11]
7. UNESCO. Open Science https://en.unesco.org/science-sustainable-future/open-science/recommendation, 2021 [accessed April 21 2022.
8. European Commission. Open Science https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science_en2021 [accessed April 21st 2022.
9. Swiss National Science Foundation. Open Science https://www.snf.ch/en/dah3uC2QX95tfPNd/topic/open-science2021 [accessed April 21st 2022.
10. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *CMAJ* 2004;171(6):606-7. doi: 10.1503/cmaj.1041281 [published Online First: 2004/09/16]
11. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013;310(20):2191-4. doi: 10.1001/jama.2013.281053 [published Online First: 2013/10/22]
12. Food and Drug Agency. Food and Drug Administration Modernization Act of 1997 (FDAMA). *Available from:https://wwwgpogov/fdsys/pkg/PLAW-*

105publ115/pdf/PLAW-105publ115pdf, 1997, Section 801 of the Food and Drug Administration Amendments Act of 2007 (FDAAA 801)), 1997 (2007)

13. Art. 56 of Human Research Act. *English version accessible under https://fedlexdataadminch/filestore/fedlexdataadminch/eli/cc/2013/617/20200101/en/pdf-a/fedlex-data-admin-ch-eli-cc-2013-617-20200101-en-pdf-apdf* 30 September 2011 (Status as of 1 January 2020)
14. Farquhar CM, Showell MG, Showell EAE, et al. Clinical trial registration was not an indicator for low risk of bias. *J Clin Epidemiol* 2017;84:47-53. doi: 10.1016/j.jclinepi.2016.11.011 [published Online First: 2017/01/16]
15. Harriman SL, Patel J. When are clinical trials registered? An analysis of prospective versus retrospective registration. *Trials* 2016;17:187. doi: 10.1186/s13063-016-1310-8 [published Online First: 2016/04/16]
16. Mathieu S, Boutron I, Moher D, et al. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;302(9):977-84. doi: 10.1001/jama.2009.1242 [published Online First: 2009/09/03]
17. Scott A, Rucklidge JJ, Mulder RT. Is Mandatory Prospective Trial Registration Working to Prevent Publication of Unregistered Trials and Selective Outcome Reporting? An Observational Study of Five Psychiatry Journals That Mandate Prospective Clinical Trial Registration. *PLoS One* 2015;10(8):e0133718. doi: 10.1371/journal.pone.0133718 [published Online First: 2015/08/20]
18. Zarin DA, Tse T, Williams RJ, et al. Update on Trial Registration 11 Years after the ICMJE Policy Was Established. *N Engl J Med* 2017;376(4):383-91. doi: 10.1056/NEJMsr1601330 [published Online First: 2017/01/26]
19. Suntila NSM, Smith A, Al-Dabbagh Y, et al. A systematic review of clinical trial registration in major respiratory journals 2010-2018. *Eur Respir J* 2018;52(6) doi: 10.1183/13993003.01491-2018 [published Online First: 2018/09/29]
20. Speich B GD, Briel M Non-registration, discontinuation, and non-publication of randomized trials: A repeated meta-research analysis *accepted by PLoS Medicine* 2022
21. DeVito NJ, Bacon S, Goldacre B. Compliance with legal requirement to report clinical trial results on ClinicalTrials.gov: a cohort study. *Lancet* 2020;395(10221):361-69. doi: 10.1016/S0140-6736(19)33220-9 [published Online First: 2020/01/21]
22. Powell-Smith A, Goldacre B. The TrialsTracker: Automated ongoing monitoring of failure to share clinical trial results by all major companies and research institutions. *F1000Res* 2016;5:2629. doi: 10.12688/f1000research.10010.1 [published Online First: 2017/01/21]

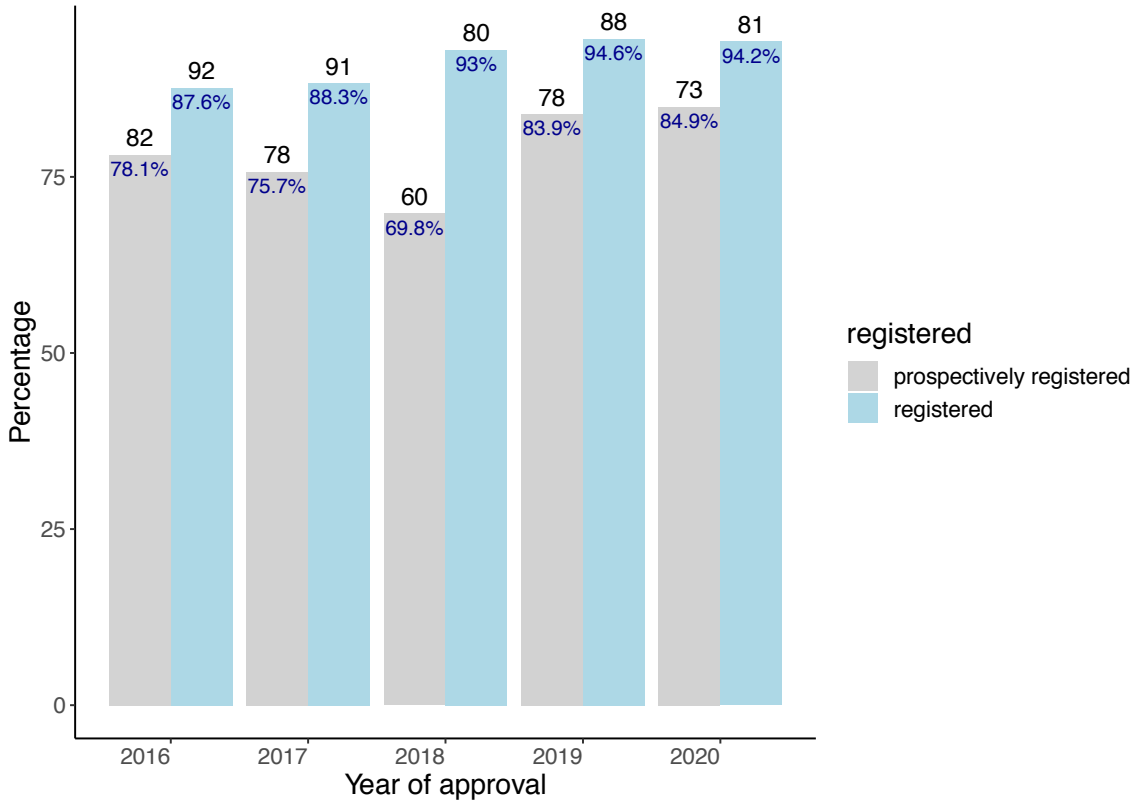
23. DataLab E. Trial Tracker Available from: https://trialstracker.ebmdatalab.net/#/2016 [accessed April 28th 2022.
24. DataLab E. FDAAA TrialsTracker Available from: https://fdaaa.trialstracker.net/2018 [accessed April 28th 2022.
25. DataLab E. EU Trialstracker Available from: https://eu.trialstracker.net2018 [accessed April 28th 2022.
26. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. *BMJ* 2000;320(7227):114-6. doi: 10.1136/bmj.320.7227.114 [published Online First: 2000/01/22]
27. Ethics S. https://swissethics.ch/assets/basec/Basec_assets/210705_basec_list-of-all-documents_clinical-trials.pdf2021 [accessed April 28th 2022.
28. Trinquart L, Dunn AG, Bourgeois FT. Registration of published randomized trials: a systematic review and meta-analysis. *BMC Med* 2018;16(1):173. doi: 10.1186/s12916-018-1168-6 [published Online First: 2018/10/17]
29. DeVito NJ, Goldacre B. Trends and variation in data quality and availability on the European Union Clinical Trials Register: A cross-sectional study. *Clin Trials* 2022;19(2):172-83. doi: 10.1177/17407745211073483 [published Online First: 2022/02/12]
30. Dal-Re R, Goldacre B, Mahillo-Fernandez I, et al. European non-commercial sponsors showed substantial variation in results reporting to the EU trial registry. *J Clin Epidemiol* 2022;142:161-70. doi: 10.1016/j.jclinepi.2021.11.005 [published Online First: 2021/11/13]
31. Viergever RF, Li K. Trends in global clinical trial registration: an analysis of numbers of registered clinical trials in different parts of the world from 2004 to 2013. *BMJ Open* 2015;5(9):e008932. doi: 10.1136/bmjopen-2015-008932 [published Online First: 2015/09/27]
32. Al-Durra M, Nolan RP, Seto E, et al. Prospective registration and reporting of trial number in randomised clinical trials: global cross sectional study of the adoption of ICMJE and Declaration of Helsinki recommendations. *BMJ* 2020;369:m982. doi: 10.1136/bmj.m982 [published Online First: 2020/04/16]
33. Dal-Re R, Marusic A. Prevention of selective outcome reporting: let us start from the beginning. *Eur J Clin Pharmacol* 2016;72(10):1283-88. doi: 10.1007/s00228-016-2112-3 [published Online First: 2016/08/04]
34. Clinical Trial Unit Basel. Descriptive statistics on research covered by the Swiss Federal Act on Research involving Human Beings, 2016-2021.
35. Hunter KE, Seidler AL, Askie LM. Prospective registration trends, reasons for retrospective registration and mechanisms to increase prospective registration

compliance: descriptive analysis and survey. *BMJ Open* 2018;8(3):e019983. doi: 10.1136/bmjopen-2017-019983 [published Online First: 2018/03/03]

36. Mayo-Wilson E, Heyward J, Keyes A, et al. Clinical trial registration and reporting: a survey of academic organizations in the United States. *BMC Med* 2018;16(1):60. doi: 10.1186/s12916-018-1042-6 [published Online First: 2018/05/03]
37. Snider SH, Flume PA, Gentilin SL, et al. Overcoming non-compliance with clinical trial registration and results reporting: One Institution's approach. *Contemp Clin Trials Commun* 2020;18:100557. doi: 10.1016/j.conctc.2020.100557 [published Online First: 2020/04/08]
38. Huser V, Cimino JJ. Evaluating adherence to the International Committee of Medical Journal Editors' policy of mandatory, timely clinical trial registration. *J Am Med Inform Assoc* 2013;20(e1):e169-74. doi: 10.1136/amiajnl-2012-001501 [published Online First: 2013/02/12]
39. Boutron I, Dutton S, Ravaud P, et al. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA* 2010;303(20):2058-64. doi: 10.1001/jama.2010.651 [published Online First: 2010/05/27]
40. Snyder RJ. Lack of Transparency in Publishing Negative Clinical Trial Results. *Clin Podiatr Med Surg* 2020;37(2):385-89. doi: 10.1016/j.cpm.2019.12.013 [published Online First: 2020/03/10]
41. Speich B, Gloy VL, Klatte K, et al. Reliability of Trial Information Across Registries for Trials With Multiple Registrations: A Systematic Review. *JAMA Netw Open* 2021;4(11):e2128898. doi: 10.1001/jamanetworkopen.2021.28898 [published Online First: 2021/11/02]
42. Scientists, keep an open line of communication with the public. *Nat Med* 2020;26(10):1495. doi: 10.1038/s41591-020-1111-1 [published Online First: 2020/10/09]
43. Swiss Academy of Medical Science(SAMS). White Paper: Clinical Research. *Swiss Academies Communications 16 (4)* 2021
44. Park JJH, Mogg R, Smith GE, et al. How COVID-19 has fundamentally changed clinical research in global health. *Lancet Glob Health* 2021;9(5):e711-e20. doi: 10.1016/S2214-109X(20)30542-8 [published Online First: 2021/04/19]

Supporting Information

Supplementary Figure 1: Prospective Registration 2016-2020 all studies



Supplementary Table 1: Associations between trial characteristics and prospective trial registration

Trial characteristics*	Prospectively registered trials n= 371	Not prospectively registered trials n=97	Univariable			Multivariable		
			OR	95% CI	p-value	OR	95% CI	*p*-value
Singlecentre (vs. multicentre)	167 (85.9%)	85 (14.1%)	0.12	0.06-0.21	<0.001	0.18	0.08-0.35	<0.001
Investigator (vs. Industry) sponsorship	249 (89.5%)	88 (10.5 %)	0.21	0.10-0.41	<0.001	1.10	0.41-2.89	0.85
Risk category low	152 (84.5%)	83 (%)	Reference			Reference		
Risk category medium	74 (97.6%)	8 (2.4%)	5.05	2.45-11.82	<0.001	3.81	1.79-9.14	0.0011
Risk category high	145 (98.7%)	6 (1.3%)	13.10	6.01-34.50	<0.001	6.69	2.68-19.55	<0.001
CTU service	90 (86.5%)	14 (13.5%)	2.08	1.14-4.06	0.023	3.04	1.58-6.21	0.0014
*Reference values: sample size <100, multicentre trials, investigator-initiated trials and drug trials.								

Abbreviations: OR, odds ratio; CI, confidence; CTU, Clinical Trials Unit.

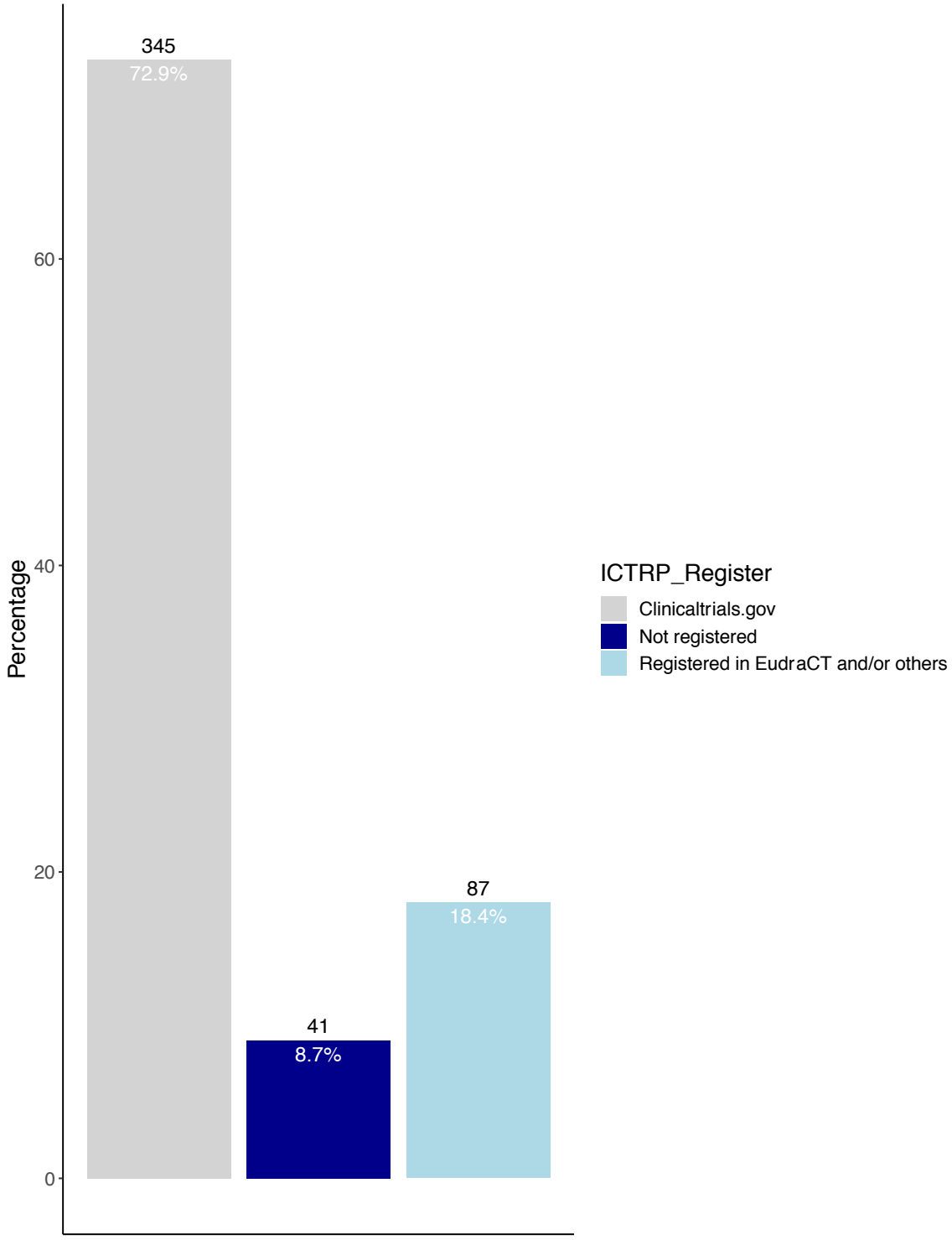
Supplementary Table 2: Association between the use of DKF Services and registration status

Registration status	Investigator-sponsored trials with CTU Services (n=104)*	Investigator-sponsored trials without CTU Services (n=238)	All investigator-sponsored trials (n=342)
Registered (n, %)	103 (99.0 %)	203 (85.3 %)	306 (89.5 %)
Prospectively	90 (86.5 %)	157 (66.0 %)	247 (72.2 %)

*104 (30.4 %) of investigator-sponsored trials made use of CTU services

Abbreviations: CTU, Clinical Trials Unit.

Supplementary Figure 2: Distribution of used trial registries for trials approved by the EKNZ



Abbreviations: ICTRP, International Clinical Trials Registry Platform; EudraCT, EU Clinical Trial Register.

Supplementary material: Questionnaire

Applicant:

Title of Project:

Trial Registration

1. HAVE PATIENTS ALREADY BEEN ENROLLED SINCE THE ETHICS APPROVAL OF THE STUDY?

Yes ☐ **No** ☐

- If **No**, (← for international studies):
  - Has the study not started yet ☐
  - Has the study not started in Switzerland ☐
- If **No**, when is the start of the study planned in Switzerland? _____

2. WAS THE STUDY REGISTERED IN ONE OF THE PRIMARY REGISTRIES?

Yes ☐ **No** ☐

- If **Yes**, please provide registration number: _____
- **If No**, would you like the CTU to register the study **Yes** ☐ **No** ☐

3. ARE YOU AWARE THAT IT IS MANDATORY TO REGISTER PROSPECTIVE CLINICAL TRIALS?

Yes ☐ **No** ☐

4. ARE YOU AWARE THAT CLINICAL TRIALS HAVE TO BE REGISTERED BEFORE THE FIRST PATIENT IS ENROLLED?

Yes ☐ **No** ☐

5. ARE YOU AWARE OF THE PRIMARY REGISTRIES OF THE WORLD HEALTH ORGANISATION?

- ClinicalTrials.gov ☐
- German Clinical Trials Register (DRKS) ☐
- EU Clinical Trial Register (EU-CTR/EudraCT) ☐
- ISRCTN-Register (International Standard Registered Clinical/soCial sTudy Number) ☐

- Others _____ ☐
- Not aware of any of the Primary registries ☐

6. ARE YOU AWARE THAT THE SWISS NATIONAL TRIALS PORTAL (SNCTP- KOFAM) IS NOT A PRIMARY REGISTRY?

Yes ☐ **No** ☐

7. DID YOU TAKE ATVANTAGE OF A SUPPORT SERVICE FOR THE STUDY?

- Clinical Trial Unit (CTU) ☐
- Contract Research Organisation (CRO) ☐
- Other: _____ ☐
- No support service ☐

8. WHAT BARRIERS OF TRIAL REGISTRATION DO YOU PERCEIVE?

- Insufficient knowledge of primary registries ☐
- Process of registration is unclear or unknown ☐
- Limited time/ resources for registration process ☐
- Missing support in the registration process ☐
- Missing reminder of obligation to register the study ☐
- Other ☐

_____

_____

9. DO YOU CONSIDER TRIAL REGISTRATION AS REASONABLE?

Yes ☐ **No** ☐

Comments:

_____

_____

_____

_____

FURTHER PUBLICATIONS

3.1 Rationale and design of repeated cross-sectional studies to evaluate the reporting quality of trial protocols: the Adherence to SPIrit REcommendations (ASPIRE) study and associated projects.

Gryaznov D, Odutayo A, von Niederhäusern B, Speich B, Kasenda B, Ojeda-Ruiz E, Blümle A, Schandelmaier S, Mertz D, Tomonaga Y, Amstutz A, Pauli-Magnus C, Gloy V, Bischoff K, Wollmann K, Rehner L, Lohner S, Meerpohl JJ, Nordmann A, **Klatte K**, Ghosh N, Heravi AT, Wong J, Chow N, Hong PJ, Cord KM, Sricharoenchai S, Busse JW, Agarwal A, Saccilotto R, Schwenkglenks M, Moffa G, Hemkens LG, Hopewell S, von Elm E, Briel M.

Trials. 2020 Oct 28;21(1):896

Abstract

Background

Clearly structured and comprehensive protocols are an essential component to ensure safety of participants, data validity, successful conduct, and credibility of results of randomized clinical trials (RCTs). Funding agencies, research ethics committees (RECs), regulatory agencies, medical journals, systematic reviewers, and other stakeholders rely on protocols to appraise the conduct and reporting of RCTs. In response to evidence of poor protocol quality, the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guideline was published in 2013 to improve the accuracy and completeness of clinical trial protocols. The impact of these recommendations on protocol completeness and associations between protocol completeness and successful RCT conduct and publication remain uncertain.

Objectives and methods

Aims of the Adherence to SPIrit REcommendations (ASPIRE) study are to investigate adherence to SPIRIT checklist items of RCT protocols approved by RECs in the UK, Switzerland, Germany, and Canada before (2012) and after (2016) the publication of the SPIRIT guidelines; determine protocol features associated with non-adherence to SPIRIT checklist items; and assess potential differences in adherence across countries.

We assembled an international cohort of RCTs based on 450 protocols approved in 2012 and 402 protocols approved in 2016 by RECs in Switzerland, the UK, Germany, and

Canada. We will extract data on RCT characteristics and adherence to SPIRIT for all included protocols. We will use multivariable regression models to investigate temporal changes in SPIRIT adherence, differences across countries, and associations between SPIRIT adherence of protocols with RCT registration, completion, and publication of results. We plan substudies to examine the registration, premature discontinuation, and non-publication of RCTs; the use of patient-reported outcomes in RCT protocols; SPIRIT adherence of RCT protocols with non-regulated interventions; the planning of RCT subgroup analyses; and the use of routinely collected data for RCTs.

Discussion

The ASPIRE study and associated substudies will provide important information on the impact of measures to improve the reporting of RCT protocols and on multiple aspects of RCT design, trial registration, premature discontinuation, and non-publication of RCTs observing potential changes over time.

3.2 Evaluation of Planned Subgroup Analysis in Protocols of Randomized Clinical Trials.

Taji Heravi A, Gryaznov D, Schandelmaier S, Kasenda B, Briel M; Adherence to SPIRIT Recommendations (ASPIRE) Study Group (including Klatte K).

JAMA Netw Open. 2021 Oct 1;4(10)

Abstract

This repeated cross-sectional study compared randomized clinical trial protocols to assess the prevalence and reporting quality of planned subgroup analyses over time.

3.3 Reporting quality of trial protocols improved for non-regulated interventions but not regulated interventions: A repeated cross-sectional study

Szimonetta Lohner , Dmitry Gryaznov, Belinda von Niederhäusern, Benjamin Speich, Benjamin Kasenda, Elena Ojeda-Ruiz, Stefan Schandelmaier, Dominik Mertz, Ayodele Odutayo, Yuki Tomonaga, Alain Amstutz, Christiane Pauli-Magnus, Viktoria Gloy, Karin Bischoff, Katharina Wollmann, Laura Rehner, Joerg J Meerpohl, Alain Nordmann**, Katharina Klatte**, Nilabh Ghosh, Ala Taji Heravi, Jacqueline Wong, Ngai Chow, Patrick Jiho Hong, Kimberly McCord, Sirintip Sricharoenchai, Jason W Busse, Arnav Agarwal, Ramon Saccilotto, Matthias Schwenkglenks, Giusi Moffa, Lars G Hemkens, Sally Hopewell, Erik von Elm, Anette Blümle, Matthias Briel

J Clin Epidemiol. 2021 Nov;139:340-349.

Abstract

Objectives: To investigate the adherence of randomised controlled trial (RCT) protocols evaluating non-regulated interventions (including dietary interventions, surgical procedures, behavioural and lifestyle interventions, and exercise programmes) in comparison with regulated interventions to the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2013 Statement.

Methods: We conducted a repeated cross-sectional investigation in a random sample of RCT protocols approved in 2012 (n = 257) or 2016 (n = 292) by research ethics committees in Switzerland, Germany, or Canada. We investigated the proportion of accurately reported SPIRIT checklist items in protocols of trials with non-regulated as compared to regulated interventions.

Results: Overall, 131 (24%) of trial protocols tested non-regulated interventions. In 2012, the median proportion of SPIRIT items reported in these protocols (59%, interquartile range [IQR], 53%-69%) was lower than in protocols with regulated interventions (median, 74%, IQR, 66%-80%). In 2016, the reporting quality of protocols with non-regulated interventions (median, 75%, IQR, 62%-83%) improved to the level of regulated intervention protocols, which had not changed on average.

Conclusions: Reporting of RCT protocols evaluating non-regulated interventions improved between 2012 and 2016, although remained suboptimal. SPIRIT recommendations need to

be further endorsed by researchers, ethics committees, funding agencies, and journals to optimize reporting of RCT protocols.

3.4 Reliability of Trial Information Across Registries for Trials with Multiple Registrations: A Systematic Review

Benjamin Speich, Viktoria L Gloy, **Katharina Klatte**, Dmitry Gryaznov, Ala Taji Heravi, Nilabh Ghosh, Ioana R Marian, Hopin Lee, Anita Mansouri, Szimonetta Lohner, Ramon Saccilotto, Edris Nury, An-Wen Chan, Anette Blümle, Ayodele Odutayo, Sally Hopewell, Matthias Briel, Adherence to Spirit Recommendations (ASPIRE) Study Group

JAMA Netw Open. 2021 Nov 1;4(11)

Abstract

Importance: Clinical trial registries are important for gaining an overview of ongoing research efforts and for deterring and identifying publication bias and selective outcome reporting. The reliability of the information in trial registries is uncertain.

Objective: To assess the reliability of information across registries for trials with multiple registrations.

Evidence review: For this systematic review, 360 protocols of randomized clinical trials (RCTs) approved by research ethics committees in Switzerland, the UK, Canada, and Germany in 2012 were evaluated. Clinical trial registries were searched from March to September 2019 for corresponding registrations of these RCTs. For RCTS that were recorded in more than 1 clinical trial registry, key trial characteristics that should be identical among all trial registries (ie, sponsor, funding source, primary outcome, target sample size, trial status, date of first patient enrollment, results available, and main publication indexed) were extracted in duplicate. Agreement between the different trial registries for these key characteristics was analyzed descriptively. Data analyses were conducted from May 1 to November 30, 2020. Representatives from clinical trial registries were interviewed to discuss the study findings between February 1 and March 31, 2021.

Findings: The analysis included 197 RCTs registered in more than 1 trial registry (151 in 2 registries and 46 in 3 registries), with 188 trials in ClinicalTrials.gov, 185 in the European Union Drug Regulating Authorities Clinical Trials Database (EudraCT), 20 in ISRCTN, and 47 in other registries. The agreement of key information across all registries was as follows: 178 of 197 RCTs (90%; 95% CI, 85%-94%) for sponsor, 18 of 20 (90%; 95% CI, 68%-99%) for

funding source (funding was not reported on ClinicalTrials.gov), 154 of 197 (78%; 95% CI, 72%-84%) for primary outcome, 90 of 197 (46%; 95% CI, 39%-53%) for trial status, 122 of 194 (63%; 95% CI, 56%-70%) for target sample size, and 43 of 57 (75%; 95% CI, 62%-86%) for the date of first patient enrollment when the comparison time was increased to 30 days (date of first patient enrollment was not reported on EudraCT). For results availability in trial registries, agreement was 122 of 197 RCTs (62%; 95% CI, 55%-69%) for summary data reported in the registry and 91 of 197 (46%; 95% CI, 39%-53%) for whether a published article with the main results was indexed. Different legal requirements were stated as the main reason for inconsistencies by representatives of clinical trial registries.

Conclusions and relevance: In this systematic review, for a substantial proportion of registered RCTs, information about key trial characteristics was inconsistent across trial registries, raising concerns about the reliability of the information provided in these registries. Further harmonization across clinical trial registries may be necessary to increase their usefulness.

3.5 Nonregistration, discontinuation, and nonpublication of randomized trials: A repeated metaresearch analysis.

Speich B, Gryaznov D, Busse JW, Gloy VL, Lohner S, **Klatte K**, Taji Heravi A, Ghosh N, Lee H, Mansouri A, Marian IR, Saccilotto R, Nury E, Kasenda B, Ojeda-Ruiz E, Schandelmaier S, Tomonaga Y, Amstutz A, Pauli-Magnus C, Bischoff K, Wollmann K, Rehner L, Meerpohl JJ, Nordmann A, Wong J, Chow N, Hong PJ, Mc Cord-De Iaco K, Sricharoenchai S, Agarwal A, Schwenkglenks M, Hemkens LG, von Elm E, Copsey B, Griessbach AN, Schönenberger C, Mertz D, Blümle A, von Niederhäusern B, Hopewell S, Odutayo A, Briel M.

PLoS Med. 2022 Apr 27;19(4)

Abstract

Background

We previously found that 25% of 1,017 randomized clinical trials (RCTs) approved between 2000 and 2003 were discontinued prematurely, and 44% remained unpublished at a median of 12 years follow-up. We aimed to assess a decade later (1) whether rates of completion and publication have increased; (2) the extent to which nonpublished RCTs can be identified in trial registries; and (3) the association between reporting quality of protocols and premature discontinuation or nonpublication of RCTs.

Methods and findings

We included 326 RCT protocols approved in 2012 by research ethics committees in Switzerland, the United Kingdom, Germany, and Canada in this metaresearch study. Pilot, feasibility, and phase 1 studies were excluded. We extracted trial characteristics from each study protocol and systematically searched for corresponding trial registration (if not reported in the protocol) and full text publications until February 2022. For trial registrations, we searched the (i) World Health Organization: International Clinical Trial Registry Platform (ICTRP); (ii) US National Library of Medicine (ClinicalTrials.gov); (iii) European Union Drug Regulating Authorities Clinical Trials Database (EUCTR); (iv) ISRCTN registry; and (v) Google. For full text publications, we searched PubMed, Google Scholar, and Scopus. We recorded whether RCTs were registered, discontinued (including reason for discontinuation), and published. The reporting quality of RCT protocols was assessed with the 33-item SPIRIT checklist. We used multivariable logistic regression to examine the association between the independent variables protocol reporting quality, planned sample size, type of control (placebo versus other), reporting of any recruitment projection, single-center versus multicenter trials, and industry versus investigator sponsoring, with the 2 dependent variables: (1) publication of RCT results; and (2) trial discontinuation due to poor recruitment.

Of the 326 included trials, 19 (6%) were unregistered. Ninety-eight trials (30%) were discontinued prematurely, most often due to poor recruitment (37%; 36/98). One in 5 trials (21%; 70/326) remained unpublished at 10 years follow-up, and 21% of unpublished trials (15/70) were unregistered. Twenty-three of 147 investigator-sponsored trials (16%) reported their results in a trial registry in contrast to 150 of 179 industry-sponsored trials (84%).

The median proportion of reported SPIRIT items in included RCT protocols was 69% (interquartile range 61% to 77%). We found no variables associated with trial discontinuation; however, lower reporting quality of trial protocols was associated with nonpublication (odds ratio, 0.71 for each 10% increment in the proportion of SPIRIT items met; 95% confidence interval, 0.55 to 0.92; $p = 0.009$). Study limitations include that the moderate sample size may have limited the ability of our regression models to identify significant associations.

Conclusions

We have observed that rates of premature trial discontinuation have not changed in the past decade. Nonpublication of RCTs has declined but remains common; 21% of unpublished trials could not be identified in registries. Only 16% of investigator-sponsored trials reported results in a trial registry. Higher reporting quality of RCT protocols was associated with

publication of results. Further efforts from all stakeholders are needed to improve efficiency and transparency of clinical research.

3.6 Reporting quality of clinical trial protocols: a repeated cross-sectional study about the Adherence to SPIrit Recommendations in Switzerland, CAnada and GErmany (ASPIRE-SCAGE)

Gryaznov D, von Niederhäusern B, Speich B, Kasenda B, Ojeda-Ruiz E, Blümle A, Schandelmaier S, Mertz D, Odutayo A, Tomonaga Y, Amstutz A, Pauli-Magnus C, Gloy V, Lohner SI, Bischoff K, Wollmann K, Rehner L, Meerpohl JJ, Nordmann AJ, Klatte K, Ghosh N, Taji Heravi A, Wong J, Chow N, Hong PJ, McCord K, Sricharoenchai S, Busse JW, Agarwal A, Saccilotto R, Schwenkglenks M, Hemkens LG, Moffa G, Hopewell S, von Elm E, Briel M.

BMJ Open 2022 May 24;12(5)

Abstract

Objectives: Comprehensive protocols are key for the planning and conduct of randomised clinical trials (RCTs). Evidence of low reporting quality of RCT protocols led to the publication of the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) checklist in 2013. We aimed to examine the quality of reporting of RCT protocols from three countries before and after the publication of the SPIRIT checklist.

Design: Repeated cross sectional study.

Setting: Swiss, German and Canadian research ethics committees (RECs).

Participants: RCT protocols approved by RECs in 2012 (n=257) and 2016 (n=292).

Primary and secondary outcome measures: The primary outcomes were the proportion of reported SPIRIT items per protocol and the proportion of trial protocols reporting individual SPIRIT items. We compared these outcomes in protocols approved in 2012 and 2016, and built regression models to explore factors associated with adherence to SPIRIT. For each protocol, we also extracted information on general trial characteristics and assessed whether individual SPIRIT items were reported RESULTS: The median proportion of reported SPIRIT items among RCT protocols showed a non-significant increase from 72% (IQR, 63%-79%) in 2012 to 77% (IQR, 68%-82%) in 2016. However, in a preplanned subgroup analysis, we detected a significant improvement in investigator-sponsored protocols: the median

proportion increased from 64% (IQR, 55%-72%) in 2012 to 76% (IQR, 64%-83%) in 2016, while for industry-sponsored protocols median adherence was 77% (IQR 72%-80%) for both years. The following trial characteristics were independently associated with lower adherence to SPIRIT: single-centre trial, no support from a clinical trials unit or contract research organisation, and investigator-sponsorship.

Conclusions: In 2012, industry-sponsored RCT protocols were reported more comprehensively than investigator-sponsored protocols. After publication of the SPIRIT checklist, investigator-sponsored protocols improved to the level of industry-sponsored protocols, which did not improve.

Keywords: Clinical trials; EPIDEMIOLOGY; Protocols & guidelines

DISCUSSION

4.1 Improving clinical research in Switzerland

In the last years several initiatives have been launched to increase the quality of clinical research in Switzerland. This process encompasses the improvement of clinical research infrastructure by the foundation of Clinical Trial Units (CTUs) and the foundation of the Swiss Clinical Trial Organization (SCTO) as the central cooperation platform for patient-oriented clinical research on the national level. Special programs dedicated to clinical research supported by the Swiss National Science Foundation (SNSF) have been initiated, and the funding and promotion of methodological research and Implementation Science has increased.[6] The White paper: Clinical Research mandated by the State Secretariat for Education, Research and Innovation (SERI) and recently published by a working group of various stakeholders set up by the Swiss Academy of Medical Sciences (SAMS) summarizes these efforts and outlines goals and recommendations.[6] The authors recommend that efforts on common overarching priorities should be more aligned in Switzerland and resources should be used more efficiently. Challenges identified include premature discontinuation of a significant proportion of SNSF-funded clinical trials and recruitment inefficiency inflating costs and delaying study timelines.[6] Within the Swiss initiative to reduce waste, the consensus-based framework INQUIRE (INcreasing QUality In patient-oriented clinical REserach) was developed. The aim of this framework is to support academia in developing quality enhancement initiatives and setting research agendas for patient-oriented research at all study stages [59]. Patient and citizen involvement in clinical trials has been shown to improve participation rates [60], and improving the reputation of clinical research in society is important. It is requested by the SNSF that all clinical trials are routinely registered and lay-language summaries provided, to lower the barriers to patients and citizens being better informed and to participate in clinical research. By making trial registration mandatory by law, Switzerland has taken an important step towards increased transparency in clinical research.

In-line with these efforts, this PhD work focused on the improvement of clinical research in terms of efficient trial conduct and increased transparency through trial registration. Providing evidence on the effectiveness of current approaches to optimize and support clinical research practices, identifying barriers in current practices as well as proposing new approaches to improve the quality of clinical research is important.

In Implementation science research methods critical for understanding the process, context, and outcomes of implementation are systematically applied, enabling scale-up and population-level benefits. We want to achieve permanent change, which is largely dependent

on the successful implementation and uptake of the new concept and comprised methods. In this context, involving multiple stakeholders to support the integration of new effective interventions into clinical research settings is key.

In our project addressing the problem of inefficient trial management we performed a contextual analysis and involved local, national and international stakeholders in the development of the new risk-tailored approach. The dashboard refinement was an iterative process involving all end-user groups to optimize the user friendliness and the integration of the developed tool into research practice.

In the project on trial registration we closely collaborated with the EKNZ and contacted researchers and support service providers to identify barriers in the registration process as well as options for improvement. By involving key stakeholders of clinical research in Switzerland, a comprehensive approach addressing the major needs of the research community has been developed and the basis for a successful implementation has been established. Barriers in the registration process of Switzerland-based trials have been identified and suggestions for improvement have been gathered with respective stakeholders. Thus, this work has the potential to increase feasibility and transparency of clinical research in Switzerland.

4.2. An evidence-based risk-tailored approach to support the monitoring and management of clinical trials in the academic setting

Basing new approaches on a summary of the evidence about the effectiveness of previous approaches is essential. When addressing the need for a more efficient and comprehensive approach to support the overall conduct of clinical trials, our first step was to summarize evidence on existing monitoring strategies. Since only a small number of prospectively evaluated monitoring approaches were available, the certainty of the evidence collected in our Cochrane review is limited. However, providing a summary of evaluated approaches and their results on effective measures can still guide the development of new efficient methods to support the conduct of clinical intervention studies. Updating the review on further evidence provided by prospective evaluation of newly developed approached will be important in this context.

Within the last years, on-site monitoring has been reduced to the most critical issues of patient rights and safety and the primary endpoint assessment, to save resources within a limited trial budget. [61] Since intensive on-site monitoring of source data verification has been shown to have limited impact on the overall study process [18-20 62], a reduction of the presence

of trial monitors on-site is reasonable. While reducing costs for on-site monitoring increases flexibility of study resources, redirecting resources to the benefit of the overall study management needs to be considered.

Multicentric trials provide the advantages of shorter patient recruitment periods, increased generalizability of results, and more robust statistical analysis, however effective trial management becomes even more important and challenging. [63 64] The pandemic underlined the need for a study management infrastructure supported by central data monitoring and remote communication. [61 65] New clinical trial management systems are developed [66 67], and have the potential to provide operational benefit to multisite trials. A survey recently published by Stansbury et al. evaluated the implementation of risk-based monitoring (RBM) and the larger risk-based quality management (RBQM) in clinical trials. Of 5987 trials ongoing at the end of 2020, 77% implemented at least one RBM/RBQM component, an increase from 47% for studies ongoing at the end of 2019. [68] Increasing the overall feasibility of a clinical trial should be the goal of all new study management and monitoring approaches. Since trial monitoring and trial management have overlapping elements ensuring adequate study processes, considering them as integrative parts is sensible. The awareness for a combined approach to increase efficiency of trial conduct support has increased in the last years [69] and commercial solutions that address the need for overall support approaches including management elements have been reported.[70 71] However, commercial solutions are often expensive and software-dependent and, thus, constitute a challenge to the limited budget of academically funded trials. Especially for the academic setting an efficient and resource-optimized approach is needed. The Clinical Trials Transformation Initiative (CTTI) recommendation supports the intention of our work to identify and focus on the most critical elements of the study conduct. What really matters needs to be continuously evaluated.[72 73] Combining this focus on critical elements of the overall trial conduct with a visualization tool for up-to-date evaluations, has the potential to support the feasibility of investigator-initiated trials.

With our study-specific risk assessment guide and the accompanying study dashboard we address the need for more efficient trial management integrating trial monitoring in the academic setting. The risk assessment guide can be routinely applied to all clinical trials and will help in the process of identifying most critical elements of trial conduct. By providing our generic template comprised of R modules, academic research institutions have the opportunity to efficiently develop their trial specific dashboards independent of the software used for trial data storage. Different modules can be combined to fit the needs of specific trials and new modules can be created based on the pool of existing modules. An integration of on-site monitoring, if required, is possible. Critical assets of the study conduct that are not necessarily covered by the on-site monitoring can be reflected in the study dashboard and

early interference and support can be initiated by the dashboard. Feedback from trial monitors suggest that an already “cleaned” database prior to on-site visits supports the focus on patient rights and safety concerns, e.g. verifying a correct inform consent process, and can promote the role of experienced study monitors as a site support team in terms of training and advice. In addition, the chance to collect missing data can be assessed instantaneously, possibly increasing the amount of analyzable data.

The strength of the projects addressing the problem of inefficient trial management is that we considered the needs from all different groups involved and the continuous feedback further optimized the usability and user-friendliness of the dashboard tool. The dashboard is structured into already established R module, allowing for a high flexibility, compatibility and resource efficiency when generating new trial-specific dashboards.

4.3 Ideal trial registration and dissemination of trial results

Trial registration is required by law in Switzerland. [47] We aimed to evaluate if this legal requirement is sufficient to ensure a complete registration of clinical intervention studies.

In collaboration with the EKNZ, we were able to provide a complete picture of trial registration and publication within a 5-year timeframe for the corresponding catchment area. The positive trend of trial registration observed is mediated by the increase of trial registration within the group of investigator-sponsored trials, while the industry-sponsored trials remained at their already very high proportion of trial registration. This indicates that educative and supportive measures within the academic setting are effective but need to be further increased and optimized to establish a complete registration of all clinical trials approved by an ethics committee. Taking advantage of supportive services provided by the DKF had a very clear positive effect on registration prevalence and suggests that providing an easily accessible support infrastructure is key. Since the analyzed sample of trials had an approval date ranging from five to one year prior to our analysis, the number of completed trials was limited. In a repeated analysis of our sample in 5 years it will be interesting to assess the percentage of published, unpublished and discontinued trials.

A recent meta-research of randomized trials revealed encouraging results in terms of registration rates and making trial results available. [5] An international initiative to tackle selective non-publication of clinical studies (publication bias) has developed 47 recommendations targeted at a variety of stakeholders [74], including the recommendation that legislators in all countries make trial registration mandatory, funding agencies request results dissemination of all funded projects, and that ethics committees require trial registration before the recruitment of the first patient and request annual reports describing the dissemination of study results. In addition, publishing results independent of the outcome within a trial registry would increase scientific neutrality of results and prevent

misinterpretation and spinning of results. [75] In our quantitative analysis we identified barriers of trials registration indicating a need for support and education at research facilities. Ethics committees could take the responsibility for sending out reminders for trial registration and updating results in the registry, but the enforcement of registration and dissemination by publishing journals should further support the goal to increase transparency in clinical research. [76] Publishing negative results should also become more common and all available information of discontinued trials has to be published within trial registries in order to maximize the output of data collected and lessons learned. Satalkar et al. have identified investigators' sense of failure and associated negative emotions as a key reason why investigators are not more transparent following trial discontinuation. This knowledge underlines the need for mandatory trial registration, facilitating the dissemination of results of all clinical trials. [77] During the Covid-19 pandemic, the importance of clinical research suddenly became publicly visible stressing the need for research transparency and availability of results. [7 78] The legal requirement in Switzerland constitute an important step toward full trial registration, but is not sufficient without the required enforcement and support of the registration process. The barriers identified in this work need to be addressed and the completeness of trial registry entries including results publication should be further pursued.

4.4 Future directions

In the future, the following objectives should be addressed: (1) to implement, user test, and revise the risk assessment and study dashboard in different medical fields and study types; (2) to investigate the cost-effectiveness of the study dashboard in order to inform decisions on actual value improvement; (3) to implement and evaluate measures to achieve full trial registration and results publication.

Here, at the University Hospital in Basel, we have the opportunity to address transparency and feasibility of clinical research in an academic setting, which is linked to the national network of clinical trial units (Swiss Clinical Trial Organization). Implementing an evidence-based monitoring and management approach at the University Hospital of Basel that has the potential to be transferred to the whole academic network for clinical research in Switzerland is aspired.

At local level, the DKF will have to promote the implementation of the risk assessment as a routine step in the joint planning of clinical trials with the respective study teams. A risk profile containing many critical and likely risks that need to be monitored throughout the trial will be most suitable for the use of a trial dashboard. Even if the cost-efficiency of the developed dashboard tool and the appropriate scope for smaller or mono centric trials is still uncertain,

a benefit for the conduct and feasibility of large international multicenter trials was perceived by all stakeholders and end-user groups. However, performing the specific risk assessment for all trials independent of size and setting will be beneficial for trial monitors as well as trial staff to increase the awareness of critical elements of the study that will require continuous oversight throughout the study conduct. Research on cost-effectiveness of the new concept would further inform the decision for which studies a dashboard would be feasible. In this context, the results of an empirical evaluation of the risk-tailored monitoring concept in Basel through a prospective Study within a Trial (SWAT) would provide valuable evidence on the effectiveness of the new approach. However, setting up a SWAT is challenging to plan and organize as trials would have to be randomized to dashboard use or no dashboard use. This would require the inclusion of a large set of trials. Collecting data on costs for setting up and maintaining a study-specific dashboard for investigator-sponsored clinical trials at the University Hospital in Basel should be the next step in the assessment of efficiency and feasibility. Dashboards can now be developed based on the provided generic structure and code, reducing the amount of resources needed to setup study dashboards for future trials. However, study-specific issues that need to be addressed in optional tabs might still be resource intensive. The vision is, that as more and more trials will use dashboards the pool of study-specific tabs will grow, increasing the amount of study-specific pathway structures convenient for trials with similar critical issues. Optimizing the timeline and resources needed for the development of a study specific dashboard will be important to advance the generation of affordable and efficient dashboards. The effectiveness of the generated dashboards on the outcome of several trials (e.g. analyzable data, timeline of trial, safety reports) has to be evaluated. In addition, the study dashboard should be tested in different research settings and study designs. Trials with a long duration and even cohorts will benefit from the comprehensive and continuous oversight of the study progress. In cohort studies, providing oversight on follow-up visits, recruitment, retention and data collection is most important. The data science team at the CTU is already in the process of developing a dashboard for a cohort study and an evaluation of perceived benefits of users will also be relevant in this context.

At Swiss national level, we shared and assist in the implementation of our concept of support for the conduct of clinical trials by providing a generic template and an accompanying description at the Github platform (https://github.com/CTU-Basel/viewTrial). A project for the joint development and revision of modules and tabs of the study dashboard within the SCTO has already been started. Implementation of our risk-tailored approach and testing of the study dashboard in other research institutions throughout Switzerland would help to better understand the actual value of the tool in clinical research practice. Continuing the collection

of feedback from users will further increase user-friendliness and help to refine structure and content.

At international level, our concept and accompanying dashboard tool are presented at international stakeholder meetings like the International Clinical Trials Methodology Conference to further disseminate this approach and gain feedback and ideas. Evidence on the effectiveness of trial monitoring approaches that address inefficient trial management and monitoring has to be gathered in the international research community. Evaluation of these methods, including our risk-tailored approach, in different settings will facilitate the uptake into clinical research practice to prevent discontinuation, ensure validity, legitimacy, and maximize knowledge gained throughout clinical trials.

Results and perceptions obtained from our analyses on trial registration in Northwestern and Central Switzerland should guide the implementation of measures to achieve full trial registration and, thereby, increase the transparency of conducted trials.
At local level, the EKNZ has to increase the awareness of the obligation to register a clinical trial. Organizing and programming an automated email reminder to applicants after the approval of a trial by the ethics committee could support this process. In addition, educating research institutes on the classification of clinical trials, and the concept, intention and importance of trial registration is essential. Further, referring researchers to the option of utilizing services offered by CTUs/DKF could foster the exchange of knowledge on trial registration processes. Finally, it will be important to evaluate the successful implementation of these measures and assess the value of improvement achieved.

At Swiss national level, sharing our analyses with the research community in Switzerland might initiate the implementation of proposed measures to support trial registration in other areas of Switzerland. A joint initiative of all ethics committees in Switzerland, and public research institutions, including universities, should focus on the adequate training and support of researchers to enhance trial registration. In addition, results reporting should be addressed, promoting the update of results in the trial registry independent of a journal publication. Besides these supportive measures, a consistent process for the enforcement of trial registration and results reporting should be obtained. For example, entering a trial registration number into the BASEC system in a set timeframe after the trial has been approved should be mandatory.

At international level, journals should extend and enforce their policy of not publishing results of trials not included in a primary trial register. [42]The importance of trial registries as the

evidence base for health care decisions needs to be underlined at international conferences in all fields of clinical research.

Finally, the impact of this work critically depends on its rigorous implementation, evaluation, and refinement. It constitutes an important step in terms of support of trial conduct and provides a basis for future projects tackling the problem of inefficient trial management and overall feasibility of clinical trials in the academic setting. Future projects guiding the implementation of measures to increase trial registration and results publication should consider the identified barriers and associated factors identified in our analyses.

4.5 Closing remarks

In this thesis work, we have initially gathered evidence on trial monitoring in clinical research and developed a new risk-tailored approach for trial monitoring and management supported by a visualization tool. Our comprehensive approach considered reduced monitoring resources on-site as well as trial management issues critical for the successful conduct of a clinical trial. Additionally, our methodological work has shed some light on the current trial registration and publication in Northwestern and Central Switzerland and offered possible options to achieve complete trial registration. The full dissemination of trial results and data via public registries are challenges that will need to be addressed in future research.

The impact of these research efforts is a comprehensive approach of trial monitoring and overall management, and a tool that enables continuous oversight of critical processes during trial conduct. An affordable solution for investigator-initiated trials that is based on a freely available risk assessment guide and a code repository available at github is provided. The vision for the future is that fewer studies have to be discontinued as problems in trial conduct are recognized early on and allow for corrective and supportive actions and that all information available from any clinical trial conducted will be publicly available in clinical trial registries.

REFERENCES

1. Amstutz A, Schandelmaier S, Frei R, et al. Discontinuation and non-publication of randomised clinical trials supported by the main public funding body in Switzerland: a retrospective cohort study. *BMJ Open* 2017;7(7):e016216. doi: 10.1136/bmjopen-2017-016216 [published Online First: 2017/08/03]
2. Sertkaya A, Wong HH, Jessup A, et al. Key cost drivers of pharmaceutical clinical trials in the United States. *Clin Trials* 2016;13(2):117-26. doi: 10.1177/1740774515625964 [published Online First: 2016/02/26]
3. Speich B, von Niederhausern B, Blum CA, et al. Retrospective assessment of resource use and costs in two investigator-initiated randomized trials exemplified a comprehensive cost item list. *J Clin Epidemiol* 2018;96:73-83. doi: 10.1016/j.jclinepi.2017.12.022 [published Online First: 2018/01/03]
4. Kasenda B, von Elm E, You J, et al. Prevalence, characteristics, and publication of discontinued randomized trials. *JAMA* 2014;311(10):1045-51. doi: 10.1001/jama.2014.1361 [published Online First: 2014/03/13]
5. Speich B GD, Briel M Non-registration, discontinuation, and non-publication of randomized trials: A repeated meta-research analysis *accepted by PLoS Medicine* 2022
6. White Paper: Clinical Research. *Swiss Academies Communications 16 (4)* 2021
7. Park JJH, Mogg R, Smith GE, et al. How COVID-19 has fundamentally changed clinical research in global health. *Lancet Glob Health* 2021;9(5):e711-e20. doi: 10.1016/S2214-109X(20)30542-8 [published Online First: 2021/04/19]
8. Janiaud P, Axfors C, Ioannidis JPA, et al. Recruitment and Results Reporting of COVID-19 Randomized Clinical Trials Registered in the First 100 Days of the Pandemic. *JAMA Netw Open* 2021;4(3):e210330. doi: 10.1001/jamanetworkopen.2021.0330 [published Online First: 2021/03/02]
9. Glasziou PP, Sanders S, Hoffmann T. Waste in covid-19 research. *BMJ* 2020;369:m1847. doi: 10.1136/bmj.m1847 [published Online First: 2020/05/14]
10. Chalmers I, Bracken MB, Djulbegovic B, et al. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;383(9912):156-65. doi: 10.1016/S0140-6736(13)62229-1 [published Online First: 2014/01/15]
11. Al-Shahi Salman R, Beller E, Kagan J, et al. Increasing value and reducing waste in biomedical research regulation and management. *Lancet* 2014;383(9912):176-85. doi: 10.1016/S0140-6736(13)62297-7 [published Online First: 2014/01/15]
12. Chan AW, Song F, Vickers A, et al. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;383(9913):257-66. doi: 10.1016/S0140-6736(13)62296-5 [published Online First: 2014/01/15]
13. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383(9913):267-76. doi: 10.1016/S0140-6736(13)62228-X [published Online First: 2014/01/15]
14. Ioannidis JP, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383(9912):166-75. doi: 10.1016/S0140-6736(13)62227-8 [published Online First: 2014/01/15]
15. Macleod MR, Michie S, Roberts I, et al. Biomedical research: increasing value, reducing waste. *Lancet* 2014;383(9912):101-4. doi: 10.1016/S0140-6736(13)62329-6 [published Online First: 2014/01/15]
16. Klatte K, Pauli-Magnus C, Love S, et al. Monitoring strategies for clinical intervention studies. *Cochrane Database of Systematic Reviews* 2019(12) doi: 10.1002/14651858.MR000051

17. Baigent C, Harrell FE, Buyse M, et al. Ensuring trial validity by data quality assurance and diversification of monitoring methods. *Clin Trials* 2008;5(1):49-55. doi: 10.1177/1740774507087554 [published Online First: 2008/02/20]
18. Duley L, Antman K, Arena J, et al. Specific barriers to the conduct of randomized trials. *Clin Trials* 2008;5(1):40-8. doi: 10.1177/1740774507087704 [published Online First: 2008/02/20]
19. Hearn J, Sullivan R. The impact of the 'Clinical Trials' directive on the cost and conduct of non-commercial cancer trials in the UK. *Eur J Cancer* 2007;43(1):8-13. doi: 10.1016/j.ejca.2006.09.016 [published Online First: 2006/11/23]
20. Tudur Smith C, Stocken DD, Dunn J, et al. The value of source data verification in a cancer clinical trial. *PLoS One* 2012;7(12):e51623. doi: 10.1371/journal.pone.0051623 [published Online First: 2012/12/20]
21. Tudur Smith C, Williamson P, Jones A, et al. Risk-proportionate clinical trial monitoring: an example approach from a non-commercial trials unit. *Trials* 2014;15:127. doi: 10.1186/1745-6215-15-127 [published Online First: 2014/04/18]
22. Treweek S, Littleford R. Trial management- building the evidence base for decision-making. *Trials* 2018;19(1):11. doi: 10.1186/s13063-017-2322-8 [published Online First: 2018/01/10]
23. Eisenstein EL, Collins R, Cracknell BS, et al. Sensible approaches for reducing clinical trial costs. *Clin Trials* 2008;5(1):75-84. doi: 10.1177/1740774507087551 [published Online First: 2008/02/20]
24. WHO. International Clinical Trials Registry Platform https://www.who.int/clinical-trials-registry-platform [accessed May 14th 2022.
25. ICH Harmonised Tripartite Guideline: Guideline for Good Clinical Practice E6 (R1) Available from: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf. 1996 [accessed May 5th 2022
26. Eisenstein EL, Lemons PW, 2nd, Tardiff BE, et al. Reducing the costs of phase III cardiovascular clinical trials. *Am Heart J* 2005;149(3):482-8. doi: 10.1016/j.ahj.2004.04.049 [published Online First: 2005/05/03]
27. Grignolo A. The Clinical Trials Transformation Initiative (CTTI). *Ann Ist Super Sanita* 2011;47(1):14-8. doi: 10.4415/ANN_11_01_04 [published Online First: 2011/03/25]
28. FDA, Food and Drug Agency. Guidance for Industry Oversight of Clinical Investigations —A Risk-Based Approach to Monitoring 2013 [accessed May 5th 2022
29. EMA, European Medical Agency. Reflection paper on risk based quality management in clinical trials. 2013 [accessed May 5th 2022.
30. EC, European Commission. Risk Proportionate Approaches in Clinical Trials- Recommendations of the Expert Group on Clinical Trials for the Implementation of Regulation (EU) 2014 [accessed May 5th 2022.
31. OECD, Organisation for Economic Co-operation and Development. OECD recommendation on the governance of clinical trials Available from: https://www.pharmamedtechbi.com/~/media/Supporting%20Documents/The%20Gray%20Sheet/39/9/oecdrecommendationgovernanceofclinicaltrials.pdf2013 [accessed May 5th 2022.
32. Hurley C, Shiely F, Power J, et al. Risk based monitoring (RBM) tools for clinical trials: A systematic review. *Contemp Clin Trials* 2016;51:15-27. doi: 10.1016/j.cct.2016.09.003 [published Online First: 2016/10/30]
33. Brosteanu O, Houben P, Ihrig K, et al. Risk analysis and risk adapted on-site monitoring in noncommercial clinical trials. *Clin Trials* 2009;6(6):585-96. doi: 10.1177/1740774509347398 [published Online First: 2009/11/10]
34. Journot V, Pignon JP, Gaultier C, et al. Validation of a risk-assessment scale and a risk-adapted monitoring plan for academic clinical research studies--the Pre-Optimon study. *Contemp Clin Trials* 2011;32(1):16-24. doi: 10.1016/j.cct.2010.10.001 [published Online First: 2010/10/19]

35. Venet D, Doffagne E, Burzykowski T, et al. A statistical approach to central monitoring of data quality in clinical trials. *Clinical Trials* 2012;9(6):705-13. doi: 10.1177/1740774512447898
36. Stenning SP, Cragg WJ, Joffe N, et al. Triggered or routine site monitoring visits for randomised controlled trials: results of TEMPER, a prospective, matched-pair study. *Clin Trials* 2018;15(6):600-09. doi: 10.1177/1740774518793379
37. Knott C, Valdes-Marquez E, Landray M, et al. Improving efficiency of on-site monitoring in multicentre clinical trials by targeting visits. *Trials* 2015;16
38. Basel CTUBDoCRUH. Descriptive statistics on research covered by the Swiss Federal Act on Research involving Human Beings, 2016-2021.
39. Dickersin K, Rennie D. Registering clinical trials. *JAMA* 2003;290(4):516-23. doi: 10.1001/jama.290.4.516
40. John LK, Loewenstein G, Prelec D. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychol Sci* 2012;23(5):524-32. doi: 10.1177/0956797611430953
41. Food and Drug Administration Modernization Act of 1997 (FDAMA) Available from: https://www.gpo.gov/fdsys/pkg/PLAW-105publ115/pdf/PLAW-105publ115.pdf

    1997 [accessed May 5th 2022.
42. ISRCTNregistry. The ISRCTN registry is a primary clinical trial registry recognised by WHO ans ICMJE that accepts all clinical research studies (proposed, ongoing or completed) Available from www.isrctn.com [accessed May 5th 2022.
43. EudraCT. European clinical Trials database is a database of all clinical trials which commenced in the community from 1 May 2004 Available from http://eudract.ema.europa.eu/eudract-web/ [accessed May 5th 2022.
44. ANZCTR. The Australian and New Zealand Clinical Trials Registry is an online registry of clinical trials being undertaken in Australia, New Zealand and elsewhere. Available from www.anzctr.org.au [accessed May 5th 2022.

45. De Angelis C, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *N Engl J Med* 2004;351(12):1250-1. doi: 10.1056/NEJMe048225
46. World Medical A. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013;310(20):2191-4. doi: 10.1001/jama.2013.281053 [published Online First: 2013/10/22]
47. Art. 56 of Human Research Act. *English version accessible under https://fedlexdataadminch/filestore/fedlexdataadminch/eli/cc/2013/617/20200101/en/pdf-a/fedlex-data-admin-ch-eli-cc-2013-617-20200101-en-pdf-apdf* 30 September 2011 (Status as of 1 January 2020)
48. Farquhar CM, Showell MG, Showell EAE, et al. Clinical trial registration was not an indicator for low risk of bias. *J Clin Epidemiol* 2017;84:47-53. doi: 10.1016/j.jclinepi.2016.11.011
49. Harriman SL, Patel J. When are clinical trials registered? An analysis of prospective versus retrospective registration. *Trials* 2016;17:187. doi: 10.1186/s13063-016-1310-8 [published Online First: 2016/04/16]
50. Mathieu S, Boutron I, Moher D, et al. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;302(9):977-84. doi: 10.1001/jama.2009.1242
51. Scott A, Rucklidge JJ, Mulder RT. Is Mandatory Prospective Trial Registration Working to Prevent Publication of Unregistered Trials and Selective Outcome Reporting? An Observational Study of Five Psychiatry Journals That Mandate Prospective Clinical Trial Registration. *PLoS One* 2015;10(8):e0133718. doi: 10.1371/journal.pone.0133718 [published Online First: 2015/08/20]
52. Zarin DA, Tse T, Williams RJ, et al. Update on Trial Registration 11 Years after the ICMJE Policy Was Established. *N Engl J Med* 2017;376(4):383-91. doi: 10.1056/NEJMsr1601330

53. Chan AW, Pello A, Kitchen J, et al. Association of Trial Registration With Reporting of Primary Outcomes in Protocols and Publications. *JAMA* 2017;318(17):1709-11. doi: 10.1001/jama.2017.13001 [published Online First: 2017/09/12]
54. DataLab E. Trial Tracker Available from: https://trialstracker.ebmdatalab.net/#/2016 [accessed April 28th 2022.
55. Powell-Smith A, Goldacre B. The TrialsTracker: Automated ongoing monitoring of failure to share clinical trial results by all major companies and research institutions. *F1000Res* 2016;5:2629. doi: 10.12688/f1000research.10010.1 [published Online First: 2017/01/21]
56. DataLab E. FDAAA TrialsTracker Available from: https://fdaaa.trialstracker.net/2018 [accessed April 28th 2022.
57. DataLab E. EU Trialstracker Available from: https://eu.trialstracker.net2018 [accessed April 28th 2022.
58. Thomas J BJ, Graziosi S EPPI-Reviewer 4.0: software for research synthesis. EPPI-Centre Software. *Social Science Research Unit, Institute of Education, University of London* 2010
59. von Niederhausern B, Guyatt GH, Briel M, et al. Academic response to improving value and reducing waste: A comprehensive framework for INcreasing QUality In patient-oriented academic clinical REsearch (INQUIRE). *PLoS Med* 2018;15(6):e1002580. doi: 10.1371/journal.pmed.1002580 [published Online First: 2018/06/08]
60. Scientists, keep an open line of communication with the public. *Nat Med* 2020;26(10):1495. doi: 10.1038/s41591-020-1111-1 [published Online First: 2020/10/09]
61. Barnes B, Stansbury N, Brown D, et al. Risk-Based Monitoring in Clinical Trials: Past, Present, and Future. *Ther Innov Regul Sci* 2021;55(4):899-906. doi: 10.1007/s43441-021-00295-8 [published Online First: 2021/04/30]
62. Embleton-Thirsk A, Deane E, Townsend S, et al. Impact of retrospective data verification to prepare the ICON6 trial for use in a marketing authorization application. *Clin Trials* 2019;16(5):502-11. doi: 10.1177/1740774519862528 [published Online First: 2019/07/28]
63. The pros and cons of multicentre studies; 2008.
64. Cheng A, Kessler D, Mackinnon R, et al. Conducting multicenter research in healthcare simulation: Lessons learned from the INSPIRE network. *Adv Simul (Lond)* 2017;2:6. doi: 10.1186/s41077-017-0039-0 [published Online First: 2018/02/17]
65. Shiely F, Foley, J., Stone, A. et al. Managing clinical trials during COVID-19: experience from a clinical research facility. *Trials* 2021;22:62.
66. Barlow C. Oncology Research: Clinical Trial Management Systems, Electronic Medical Record, and Artificial Intelligence. *Semin Oncol Nurs* 2020;36(2):151005. doi: 10.1016/j.soncn.2020.151005 [published Online First: 2020/04/08]
67. Mudaranthakam DP, Brown A, Kerling E, et al. The Successful Synchronized Orchestration of an Investigator-Initiated Multicenter Trial Using a Clinical Trial Management System and Team Approach: Design and Utility Study. *JMIR Form Res* 2021;5(12):e30368. doi: 10.2196/30368 [published Online First: 2021/12/24]
68. Stansbury N, Barnes B, Adams A, et al. Risk-Based Monitoring in Clinical Trials: Increased Adoption Throughout 2020. *Ther Innov Regul Sci* 2022;56(3):415-22. doi: 10.1007/s43441-022-00387-z [published Online First: 2022/03/03]
69. Agrafiotis DK, Lobanov VS, Farnum MA, et al. Risk-based Monitoring of Clinical Trials: An Integrative Approach. *Clin Ther* 2018;40(7):1204-12. doi: 10.1016/j.clinthera.2018.04.020 [published Online First: 2018/08/14]
70. CluePoints.Risk-basedmonitoring. http://cluepoints.com/risk-based-monitoring. 2016 [accessed 31st July 2022.
71. Medidata.Strategic monitoring https://www.medidata.com/wp-content/uploads/2018/12/Strategic-Monitoring_2018Medidata_Fact-Sheet.pdf2018 [accessed 31st July 2022.
72. CTTI. RECOMMENDATIONS: EFFECTIVE AND EFFICIENT MONITORING AS A COMPONENT OF QUALITY ASSURANCE IN THE CONDUCT OF CLINICAL

TRIALS Available at: https://ctti-clinicaltrials.org/wp-content/uploads/2021/06/CTTI_Monitoring_Recs.pdf2021 [accessed 21st July 2022.

73. CTTI. QUALITY BY DESIGN PROJECT - CRITICAL TO QUALITY (CTQ) FACTORS PRINCIPLES DOCUMENT Available at: https://ctti-clinicaltrials.org/wp-content/uploads/2021/07/CTTI_QbD_Workshop_Principles_Document.pdf 2021 [accessed 21st July 2022.
74. Meerpohl JJ, Schell LK, Bassler D, et al. Evidence-informed recommendations to reduce dissemination bias in clinical research: conclusions from the OPEN (Overcome failure to Publish nEgative fiNdings) project based on an international consensus meeting. *BMJ Open* 2015;5(5):e006666. doi: 10.1136/bmjopen-2014-006666 [published Online First: 2015/05/07]
75. Boutron I, Dutton S, Ravaud P, et al. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA* 2010;303(20):2058-64. doi: 10.1001/jama.2010.651 [published Online First: 2010/05/27]
76. Speich B, Gloy VL, Klatte K, et al. Reliability of Trial Information Across Registries for Trials With Multiple Registrations: A Systematic Review. *JAMA Netw Open* 2021;4(11):e2128898. doi: 10.1001/jamanetworkopen.2021.28898 [published Online First: 2021/11/02]
77. Satalkar P, McLennan S, Elger BS, et al. Investigators' sense of failure thwarted transparency in clinical trials discontinued for poor recruitment. *J Clin Epidemiol* 2022;145:136-43. doi: 10.1016/j.jclinepi.2022.01.024 [published Online First: 2022/02/07]
78. Gaba P, Bhatt, D.L. . The COVID-19 pandemic: a catalyst to improve clinical trials. *Nat Rev Cardiol* 2020;17 673–75

CURRICULUM VITAE

Katharina Klatte
Master of Science in Molecular Biotechnology
Date of birth: July 2nd 1986
Place of birth: Tokyo (Japan)
Married, 4 children
+49 151 61404110
katirondorf@me.com

Education and Research Experience

10/2018 -12/2022	**Doctoral student at the University of Basel in Clinical Research (Faculty of Medicine). Research project at the Department of Clinical Research (DKF) – University Hospital Basel** The PhD work focuses on creation of evidence on feasibility and transparency of clinical research and the development and implementation of concepts to improve the efficiency and transparency of clinical trials in Switzerland. • Meta-epidemiologic review • Leading a project group in the development of a comprehensive risk-based monitoring and management approach • Data collection and extraction using different software • Quantitative and qualitative data analysis • Data analysis using R programming • Introduction to R shiny programming
11/2017 – 05/2018	Online Courses Stanford University School of Medicine • Statistics for Medical Professionals • Thinking Critically: Interpreting Randomized Clinical Trials Programming courses • SAS online programming course • Statistics with R (Coursera, Duke University)
03/2011 – 01/2014	**Research position in the Department of Experimental Epileptology** Research interest: D-serine mediated co-transmission in Epilepsy; Global and local inhibition in the normal and epileptic hippocampus Graduate School *"Medical Neuroscience"*, University of Bonn. Supervisor: *Prof. Heinz Beck*. (12/2011 – 12/2012 Parental Leave) • D-serine mediated co-transmission in Epilepsy • Global and local inhibition in the normal and epileptic hippocampus • Data analysis using IGOR Pro and MATLAB programming • In-vitro electrophysiology combined with optogenetics • Laser stimulation of ChR2 expressing

	neurons, patch-recordings from different hippocampal neuron-classes
10/2008 – 03/2011	Master of Science in Molecular Biotechnology at the University of Bonn (1.2); Master thesis *"Properties of Synaptic and Extrasynaptic N-methyl-D-aspartate Receptors in Chronic Epilepsy"* in the Department of Experimental Epileptology at the *Life & Brain Research Center*, *Prof. Heinz Beck* • Data analysis and statistics using IGOR Pro • Combining electrical stimulation with patch-recordings in hippocampal slices • Drug application in control and disease model
10/2005 – 09/2008	Bachelor of Science in Molecular Life Science at the University of Lübeck; Bachelor thesis *"Monogenic Parkinson syndromes: Localization of a genetic cause and mutational analysis"* at the Clinic of Neurology at the University of Lübeck, Department of Neurogenetics, Prof. Christine Klein • Poster presentation at the 14th Annual Meeting of the German Society of Neurogenetics (September 25-27, 2008, Lübeck) • PCR-optimization • Sequencing • Genotyping (LICOR) • Coupling analysis • Mutation analysis
06/2005	University entrance diploma (Abitur), Carl-von-Ossietzky-Gymnasium, Bonn
08/2002 – 02/2003	Exchange student program, Tate, Mississippi

Scientific Publications outside of PhD and Awards

Klatte, K., Kirschstein, T., Otte, D., Pothmann, L., Müller, L., Tokay, T., Kober, M., Uebachs, M., Zimmer, A., Beck, H. Impaired D-serine mediated co-transmission mediates cognitive dysfunction in epilepsy. *J. Neurosci.* **33**, 13066-80 (2013).

Young Investigator Award of the Bernd and Ingeborg Hentschel Foundation for the presentations „Impaired D-serine mediated co-transmission mediates cognitive dysfunction in epilepsy", and "Basic mechanisms of cognitive symptoms in Epilepsy" at the conference of the German Liga Against Epilepsy (Deutsche Gesellschaft für Epileptologie e.V.), World Conference Center Bonn (Mai 2014).

Internships and Teaching Experience

03/2010 – 06/2010	Internship at the Neuroscience Department of the Einstein College of Medicine, Yeshiva University, New York, Prof. Pablo E. Castillo • Field potential recording techniques • Basic whole cell recordings skills in acute hippocampal brain slices
11/2009 – 12/2009	Internship at the Life & Brain Center, NeuroCognition Platform, NeuroPlasticity, Prof. Heinz Beck • Basic patch-clamp techniques on isolated single human cells
04/2007 – 07/2008	Student assistant at the Institute of Physics at the University of Lübeck
09/2007	Internship at the Life & Brain Center at the University of Bonn, Transgenics Platform, Prof. Andreas Zimmer • Methods of biochemical, surgical and behavioral experiments

Additional skills

Data Analysis	Basic programming skills in Java, Matlab, IGOR Pro, SAS, R, Microsoft Office
Foreign languages	German, mother tongue English, fluent (Overall 3 years of US residency) Latin, advanced proficiency exam
Computer literacy	Microsoft Office, Programming in R, Matlab, Igor Pro