17 Clinical Trial Registries, Results Databases, and Research Data Repositories

Karmela Krleža-Jerić, Mersiha Mahmić-Kaknjo, and Khaled El Emam

Abstract

Trial registration, results disclosure, and sharing of analyzable individual participant data (IPD) are considered powerful tools for achieving higher levels of transparency and accountability for clinical trials. The emphasis on disseminating knowledge and growing demands for transparency in clinical research are contributing to a major paradigm shift in health research. In this new paradigm, knowledge will be generated from the *culmination* of all existing knowledge—not just from bits and parts of previous knowledge, as has been largely the case until now. Fully transparent clinical research diminishes publication bias, increases accountability, avoids unnecessary duplication of research (and thus avoid research waste), efficiently advances research, provides more reliable evidence for diagnostic and therapeutic interventions, regains public trust, and contributes to research integrity. Transparency of clinical trials, at a minimum, means sharing information about the trial design, conduct, and results, as well as the analyzable data. Not only must the information itself be explicitly documented, but an access location or medium for distribution also must be provided. Thus, transparency is realized by making research protocols, results, and cleaned and anonymized IPDs publicly available using well-defined, freely accessible electronic tools. Many electronic tools enabling sharing clinical trial information have emerged. These tools include registries hosting protocol data, results databases hosting aggregate data, and research data repositories hosting reusable and analyzable data sets and other research-related information. These tools are at different levels of development and are plagued with heterogeneity as international standards for trial registration do not yet address the sharing of individual patient data.

K. Krleža-Jerić (✉)
IMPACT Observatory, Mediterranean Institute for Life Sciences—MedILS, Split, Croatia

Croatian Association for Sciences and Arts-HDZU, Sarajevo, Bosnia and Herzegovina Croatian, Cochrane, London, UK

Vivli Cambridge Croatian, Cochrane, Split, Croatia

ISRCTN Registry Advisory Group, London, UK

M. Mahmić-Kaknjo
Zenica Cantonal Hospital,
Zenica, Bosnia and Herzegovina

Sarajevo Medical School, Sarajevo School of Science and Technology, Sarajevo, Bosnia and Herzegovina

K. El Emam
School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada
e-mail: kelemam@ehealthinformation.ca

R. L. Richesson et al. (eds.), *Clinical Research Informatics*, Health Informatics,
https://doi.org/10.1007/978-3-031-27173-1_17

Additionally, the need to measure and improve clinical trial transparency has led to development of specific electronic tools. This chapter is relevant for any professional involved in clinical trials and the use of the knowledge generated from them, including clinical and biomedical researchers, clinical trialists, systematic reviewers, information technology and informatics specialists, patients, journal editors, and public and private research funders and sponsors. Suggested competencies and learning activities for specific roles are presented at the end of the chapter.

Keywords

Clinical research transparency · Trial registration, public disclosure of results and analysable Individual Participant Data, IPDs · Trial registries · International standards · Results databases · Protocol-Results-Data · Research data repositories · Reuse of data- Open data User perspectives · Research integrity Anonymisation · Future developments Synthetic data

Learning Objectives

After reading this chapter, readers should be able to:

1. Understand the importance and benefits of transparency in clinical research.
2. Understand the importance of different types of research players and stakeholders, and articulate their particular roles, responsibilities, and contributions toward advancing the generation and dissemination of knowledge from clinical trials.
3. Find relevant information about clinical trial transparency and methods, based on research roles, to achieve it.
4. Identify and execute requirements for each stage of clinical study (from design to data sharing) to meet the transparency and open science requirements and to respect and ensure research integrity.
5. Find and use relevant standards for registration and reporting of clinical trials.
6. Recognize the on-going evolution of regulations for research transparency (including updates of the WHO standards) and the subsequent need to search for standards for results and data sharing as they change.

Introduction

We provide definitions for basic terms used in this chapter below; other relevant terms will be defined throughout the chapter.

Clinical trial. WHO defines a clinical trial as "any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes" [1].

Research integrity. Clinical research integrity is part of clinical integrity which in turn is a part of health care integrity. Research integrity should be present in all stages of the research process. In short, the research integrity means that research is carried out with a high level of integrity, upholds values of honesty, rigour, transparency, and open communication, as well as care and respect for those involved in research and accountability for a positive research environment (adapted from WCRI [2]).

Transparency of clinical trials means sharing information about the trial design, conduct, and results, as well as the analyzable individual participant data (IPD), so that the trial can be evaluated, interpreted and reproduced, and so that, above all, further research can be conducted by through in IPD meta-analyses to speed the creation of new knowledge.

Systematic review of clinical trials is based on re-analysing of aggregate data from several trials published in peer review journals.

IPD (Individual Participant Data) **meta-analysis of clinical trials** is a type of secondary analysis which involves the re-analysis of cleaned and anonymized individual participant data from several trials. The IPD meta-analysis and systematic reviews are considered the most reliable

source of evidence for decision making, as illustrated on Fig. 17.1.

Largely enabled by informatics advancements, the movement toward open science and open data (i.e., making raw data from research available for analysis) has penetrated clinical trials, as observed by Vickers in 2016 [3]. **Clinical trial raw data** refers specifically to the cleaned and anonymized individual participant data (IPD). However, consumers of these data ultimately need analyzable data sets, which include the IPD, metadata, and adjacent (or supporting) documents.

There are three broad types of clinical trial information that can be shared publicly or openly on the Internet: protocol, results and findings, and raw data sets [4]. More precisely, these include:

(a) The registration of selected protocol elements in trial registries, often complemented by publication of full protocols in journals.
(b) The public disclosure of summary results (aggregate data) in databases; these are, usually developed by clinical trial registries and are beyond publications in peer-reviewed journals.
(c) The public availability of analyzable data sets; these data sets are based on cleaned, anonymized individual participant data (IPD) and adjacent trial documentation.

There are several modes or mechanisms of finding and accessing IPD-based analyzable data sets for secondary analysis (often called pooled or meta-analysis of IPDs). These include (a) direct researcher-to-researcher contact (i.e. reviewer contacting initial data producers), (b)

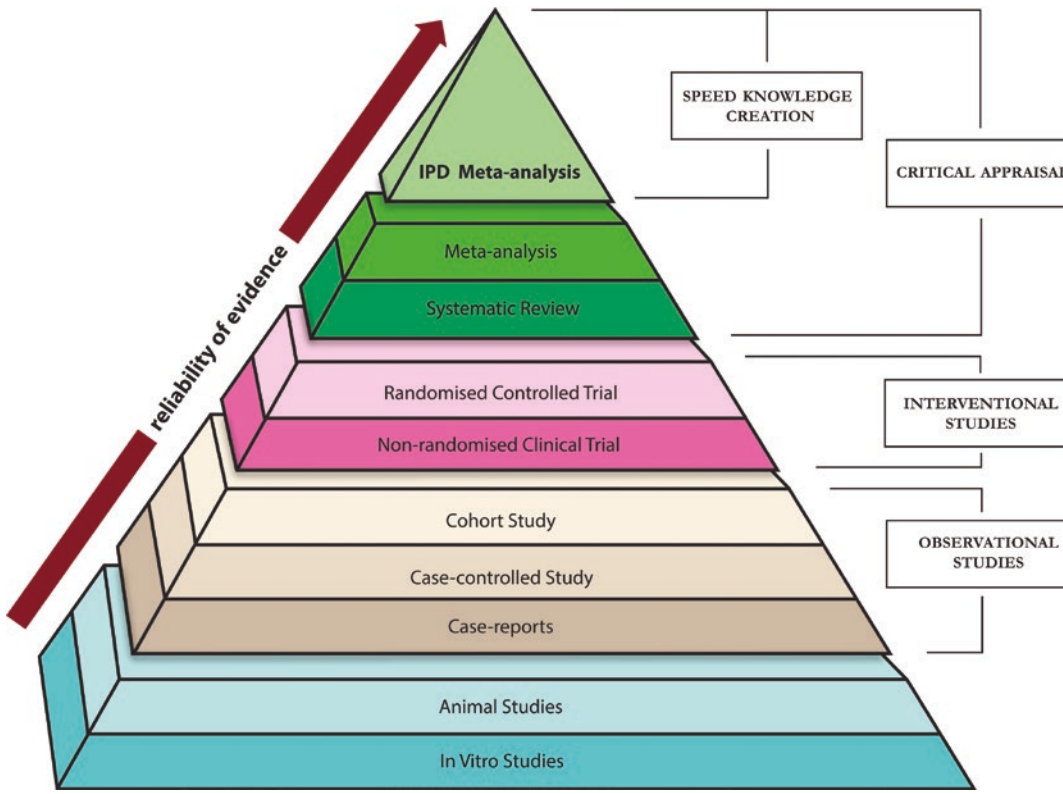

Fig. 17.1 Evidence pyramid—reliability of evidence that can be used for decision-making in health

initiatives and projects that play an intermediary role, and (c) publicly accessible research data repositories, databases, and platforms.

(a) *Direct researcher-to-researcher contact*. The reviewer gets the data directly from the original data creator by contacting him or her. The reviewer identifies studies mainly by following the literature and/or by visiting trial registries.
(b) *Intermediary contact in which the researcher requests data from special* initiatives, projects or platforms including Clinical Study Data Request [5] Project Datasphere [6], Yoda [7], or Vivli [8] which developed a clinical trial data sharing and analytics platform. The reviewer applies for data to an independent panel, a sort of peer-reviewed panel that is formed by a group of data providers or producers (currently mainly academia and the pharmaceutical industry). The independent panel is usually an international panel. Increasingly, government agencies, such as the European Medicine Agency (EMA) are moving in this direction [9].
(c) *Open-access, publicly accessible research data repositories*. They might be either domain repositories that specialize in hosting clinical trial data or general repositories that host clinical trial data in addition to hosting raw data from several or all research areas. There are currently several such open-access general research data repositories in public domain that host CT data.

In this chapter, we focus on registration and registries, results disclosure and databases, and research data repositories as tools of clinical trial transparency, with emphasis on research integrity and the impact of informatics on quality of the research process, its reporting, and data sharing, including anonymization of data.

Rationale for Registration and Reporting

Trial registration, results disclosure, and making analyzable IPD-based data publicly available all share the same underlying rationale and principles of maximizing the outputs of clinical research, diminishing research waste, and enhancing knowledge creation, all while respecting research integrity. Trial registration, results disclosure, and data sharing are considered powerful tools for achieving higher levels of transparency and accountability in clinical trials, as well as for improving the quality and integrity of research. Increasing emphasis on knowledge sharing and growing demands for transparency in clinical research are contributing to a major paradigm shift in health research that is well underway. In this new paradigm, knowledge will be generated from the *culmination* of all existing knowledge—not just from bits and parts of previous knowledge, as has been largely the case so far [10].

A stepwise process of opening clinical trial data began with the registration of protocol elements, but it was clear from the very beginning that without results disclosure, clinical trial registration would be an empty promise. Later, it became well understood that transparency would not be achieved without results *and* data disclosure. We are firmly in the era of evidence-informed decision-making in health for both individuals and populations at all levels—local, regional, national, and global. This decision-making is multifaceted, from the individual patient via physician to health administrators and policymakers [11]. Registration of protocol items, publication of the complete protocol, and public disclosure of trial findings in peer-reviewed journals—complemented with public (Internet-based) disclosure of results including aggregate data and IPD-based analyzable data sets—collectively represent a *totality* of evidence and knowledge for a given subject area and are integral to supporting efforts toward evidence-informed decision-making. Given the new appreciation of sharing data along with results, an operational definition of clinical trials results disclosure should now include at least three components: publication in a journal, posting summary results in open-access Internet-based database or registry, and publishing analyzable data sets in research data repository.

Evidence is needed to support many personal and policy decisions in health and in research.

Randomized controlled trials (RCT), systematic reviews, and increasingly IPD-based meta-analyses of RCTs are considered gold standards for evidence creation, illustrated by their positions at the top of the pyramid of evidence (Fig. 17.1). There has been quite an evolution from the time (not so long ago) when we considered first RCTs, then their systematic reviews (i.e., re-analysing the *aggregate or summarized data*, usually obtained from publications) as a gold standard to the growing notion that the gold standard requires the meta-analysis using the *raw data* from the trial [12]. This position of clinical trials on the evidence pyramid implies that the reliability of results generated by clinical trials is indeed very important. As the evidence gained from clinical trials, their systematic reviews and above all, their meta-analysis might be directly implemented into clinical decision-making, it follows that the quality of these results should be continually scrutinized. Unfortunately, the reliability of trial-based evidence is questionable due to publication and outcome reporting bias, as well as still insufficient data sharing—which means that others cannot replicate or verify results and conduct the pooled analysis. Consequently, incomplete evidence can lead to biased clinical decisions, with often harmful consequences, and can damage public trust in research and in medical interventions. Following medical deontology, doctors' prescription habits are supposed to be judiciary, which requires complete and total knowledge of the benefits and potential harms of prescribed medications. This is difficult at best and impossible if the information about the given diagnostic tools, medications, or devices is not available or is incomplete [13] and thus biased [10, 11, 13, 14].

The full transparency of clinical research is a powerful strategy to diminish publication bias, increase accountability, avoid unnecessary duplication of research, avoid waste, advance research more efficiently [4], provide more reliable evidence for diagnostic and therapeutic prescriptions, speed knowledge creation, and regain public trust [11]. Transparency of clinical trials, at a minimum, means sharing information about design, conduct, and results. The information itself must be explicitly documented, but then an access location or medium for distribution must be provided. Until recently, the public disclosure of clinical trial data was realized by posting them in well-defined, freely accessible clinical trial registries and results databases. This is certainly a dynamic field and since the previous versions of this chapter in 2012 and 2019 [12, 15], much has changed. As we described in 2019, open-access research data repositories have been developing, and the anonymized, analyzable data sets (i.e. IPDs and adjacent documentation needed to make data analyzable) could be made publicly available by publishing them in such repositories. This was followed by elaboration and improvements, supported by recommendations, policies, and increased understanding of importance of it by all stakeholders.

The clinical trial enterprise is international, and therefore the standards and development of clinical trial registries, results databases, and research data repositories should be at an international level and with open access. Standards and guidelines are essential for every step of the process—from protocol development to posting IPD in repositories. Such internationally defined standards should be flexible to allow elaboration of required fields and addition of more fields as needed. While there are international standards for trial registration and registries, the standards for results disclosure and, most importantly, standards for preparing clinical trial data for public sharing, including anonymization and the definition of the requirements for repositories that host them, have yet to be developed. The lack of standards related to preparing and publishing data in research data repositories makes it rather difficult for researchers to decide where to publish and where to search for data from completed studies for further analysis. Fortunately, since early 2000 when several of us called for registration and results reporting in the Ottawa statement [16], there has been a substantial evolution in this field. As defining standards and procedures of sharing results and IPDs is ongoing it is expected to lead to more efficient data sharing and reuse and thus increase the transparency in clinical research, driving innovations and consequently benefit the health of population.

Considering that trials take place internationally and that the knowledge gained by them may be used by anyone anywhere in the world, their quality is constantly and internationally scrutinized. This has led to development of several tools to measure and influence the improvement of transparency by pharma and academic researchers, including Till Bruckner, Transparimed [14], IMPACT Observatory that evolved from the Ottawa statement [17], the Research Data Alliance (RDA) [18], the Good Pharma Scorecard (GPS) developed by the Bioethics International in 2009 [19] and refined in 2019 to measure IPD sharing present annual ranking of companies [20], Trials Tracker developed by Goldacre, DeVito et al. in 2018 [21]. Additionally, the WHO has been publishing lists of registered trials and those that submitted results to the registries that are regularly updated [22]. Public funders and regulators are increasingly fostering all three [23] aspects of trial transparency from registration via results to data sharing.

Trial Registration

Development of Trial Registration

Although the need for trial registration (i.e., publishing protocol information) has been discussed since 1980s [24], only at the beginning of this millennium did trial registration garner widespread attention from stakeholders representing diverse and varied perspectives. The practical development of trial registration began around 2000 with two critical boosts in 2004 and in 2006. The 2004 New York State Attorney General vs. Glaxo case [25, 26] inspired the International Council of Medical Journal Editors (ICMJE) [27] and Ottawa statement [16] as well as the recommendations of the Mexico Ministerial Summit organized by the World Health Organization (WHO) in 2004 [28]. These led to the development of international standards for trial registration by the WHO, which were launched in 2006 and changed the landscape of trial registration worldwide [29].

As we learned by the IMPACT Observatory scoping review [30], a number of circumstances had coincided by the year 2000 (earlier than initially thought) which enabled the development of data sharing, beginning with trial registration. These include:

- Development of the use of informatics in research, in this case including Internet-enabled storage and retrieval of large data sets
- The definition of data, metadata, and evidence-informed (initially called evidence-based) medicine
- The use of evidence gained by systematic reviews and initial IPD-based meta-analysis in decision-making
- The appreciation of the impact of trial registration on knowledge creation, sharing, and Knowledge Translation-KT
- The existence and experience of two major registries: the International Standard Randomized Clinical Trials Number (ISRCTN) [31] http://www.isrctn.com, based in the UK, and ClinicalTrials.gov [32], based in the USA
- Growing awareness of the need to enhance transparency
- The willingness of the international research community to embark on this undertaking
- The awareness of the harmful consequences of decision-making in the context of partial evidence
- The powerful arguments from oncology, paediatrics, rare diseases, AIDS, pregnancy, perinatal medicine, and media reporting trial -related scandals
- The awareness of unnecessary duplication of research and subsequent waste of precious resources

The initial international trial registration standards that were launched by WHO in 2006 provided essential contribution toward achieving the evidence-informed decision-making. These standards clearly define registries and trials that need to be registered, define the minimum data set, designate the timing of registration, assign unique numbers to trials, and set international standards

to facilitate the development of new national or regional registries as well as the comparability of data across registries. It is important to note that as of 2022, there are no international standards for results disclosure or public sharing of analyzable data. As such standards are needed, we strongly believe that they will likely be developed in the near future and thus create numerous opportunities for informatics and information technology experts to leverage and apply to new applications. Additionally, further evolution of trial registration standards and uses has been taking place, again leading to new applications and resources that will undoubtedly impact the development of new research and our subsequent understanding of health, disease, and effective therapies. For example, registration information has been increasingly used to analyse types of trials in a given country or area of research that might be used in planning of future research and meeting the health needs [33].

Research transparency includes having protocol documents electronically available. For example, if the protocol is posted on the registry website, all trial-related data from them ideally could be cross-referenced to results and findings. However, a trial protocol can be very complex and lengthy, which can make finding the needed information difficult. At the same time, it might be missing essential information potentially leading to improvisation during a trial and consequent bias. To overcome this, an international group defined the set of Standard Protocol Items for Randomized Trials (SPIRIT), developed SPIRIT guidelines, and made them publicly available [34–36]. SPIRIT is expected to increase the clarity of clinical research protocols and ensure that the collection of necessary items is indeed specified in the protocol, thus contributing to the overall quality of the protocol and the study and results it generates. The use of SPIRIT guidelines in development of protocols might also facilitate public disclosure, especially in combination with the growing use of electronic data management [37].

Since the previous edition of this book in 2019, SPIRIT has been increasingly used in protocol developments. It has been endorsed and recommended by numerous organizations including journals, regulators, funders, trial research groups and patient organisations. Furthermore, adaptations and extensions of SPIRIT have been developed to meet specific needs, including SPIRIT PRO, for patient reported outcomes [38], SPIRIT extension for Chinese traditional medicine [39], SPIRIT AI extension [40], SPIRIT for N-of 1 trials [41]. SPIRIT is included in the Enhancing the QUAlity and Transparency Of health Research, EQUATOR Network, [42] along with the CONSORT guidelines for trial reporting and other reporting guidelines.

Clinical Trial Registries

A *clinical trial registry* is an open-access, Internet-based repository of defined protocol information. Different kinds of clinical trial registries exist in the public and private domains, such as international-, country-, and region-specific registries, as well as corporate (sponsor-driven) registries. The presence of multiple registries might be seen as a natural consequence of increased pressure and interest and as a positive development; however, a proliferation of registries could potentially lead to information overload and confusion for patients, clinicians, policymakers, and research sponsors. For example, an inexperienced user may not have known which clinical trial registries to trust. Fortunately, the international health and research community has identified the need for international standards that define required features of registries as well as the content and supporting information that they must provide.

Standards, Policies, and Principles of Trial Registration

As clinical trials are conducted throughout the world, trial registration standards must be defined on the international level. WHO developed international standards for trial registration, which were endorsed by the ICMJE, most medical journal editors, the Ottawa group, some public

funders, organizations, and countries. It is important to note that individual countries often implement international standards by adopting and extending them with additional fields to host more information in their registries.

WHO international standards have helped shape many, if not all, trial registries and have been contributing to the quality and the completeness of data for registered trials. Also, it is expected that the WHO standards will play a major role in the evolution of trial registration. They are sometimes referred to as WHO/ICMJE standards, or even cited only as ICMJE requirements, because the journal editors endorsed the WHO international standards in their instructions to authors and in related FAQs. It is important to note that the ICMJE clearly indicates on its website that it is "*no longer the entity that reviews registries for acceptability. Registries should consult the WHO International Clinical Trials Registry Platform. Registries that the WHO designates as primary registries will be acceptable to the ICMJE"* [43].

WHO international standards define the scope, i.e., *all* clinical trials need to be registered; the registries that meet the well-defined criteria; the timing i.e., prospective nature of the registration prior to the recruitment of the first trial participant; the content i.e. a minimum data set that needs to be provided to the registry, and the assignment of the unique identifier (ID). These international standards also define the criteria that the registry must meet, which includes level (nationwide or regional), ownership and governance (public or private nonprofit), trial acceptance, open access, and structure. Structurally, the registry must have enough fields to host at least WHO Trial Registration Data Set (TRDS).

The WHO/International trial registration standards are meant to be revisited frequently as methodology evolves, demands for transparency increase, and with ongoing evaluation and analysis four more items were added to the initial list of 20, each with precise definition and description thus forming the current version 1.3.1 of the WHO Trial Registration Data Set, TRDS, consisting of 24 items [44]:

1. Unique trial number and the name of registry
2. Trial registration date
3. Secondary ID
4. Funding source(s)
5. Primary sponsors
6. Secondary sponsors
7. Responsible contact person
8. Research contact person
9. Public title
10. Scientific title
11. Countries of recruitment
12. Health condition or problem studied
13. Interventions (name, dose, duration of the intervention studied, and comparator)
14. Inclusion/exclusion criteria
15. Study type (randomized or not, how many arms, who is blinded)
16. Anticipated start date (and later on the actual start date)
17. Target sample size
18. Recruitment status (not yet recruiting, recruiting, temporarily stopped recruiting, or closed for recruitment)
19. Primary outcome(s) (name, prespecified time point of measurement)
20. Key secondary outcomes
21. Ethics review
22. Completion date
23. Summary results (includes Data Sharing Plan (Yes, No) and description)
24. IPD sharing statement

In order to foster the implementation of standards, to facilitate creation of new registries, to identify the best practice, and to help develop trial registration policies, WHO formed a freely accessible search portal in 2007, followed in 2008 by the formation of a network of registries and of the Working Group on Best Practice for Clinical Trial Registries. The WHO International Clinical Trials Registry Platform (ICTRP) [1] is a unique global portal to the trials in registries that

meet criteria as data providers i.e., WHO primary registries and ClinicalTrials.gov.

ICRTP initially displayed basic information including the WHO trial registration data set (in English), the criteria for registries and a list and links to registries-data providers. It also created the unique identifier for each registered trial which is to be used in any communication about a trial, including in the ethics committees/boards' communications, consent forms, reports, publications, amendments, and press releases. The ICRTP has been constantly evolving and it now provides more information, including the access to a central database containing datasets provided by the registries and links to the full original records. Its *trial search portal* offers access to the information about trials by phases trials with results, and trials for rare diseases/orphan drugs and genome editing, enabling users to view the full picture of a given trial, from start to finish.

WHO ICRTP is also supporting a development of policies and regulations and posts the list of organisations with policies on its website. While some countries only recommend trial registration (Canada, Australia), others (such as the USA and the EU) make it a compulsory prerequisite in the drug marketing authorization process (i.e. approving new drug for the market). So far only few countries have also developed regulations making trial registration compulsory. Some of these countries e.g., EU, Brazil, India, South Africa also have registries, while Argentina, Israel, and Switzerland have regulations but do not have a registry. Increasingly there is a call to make prospective registration of observational and other studies compulsory [45, 46].

Following the addition of Summary Results, item 23, to the Trial Registration Dataset, (TRDS), trial registries have been increasingly including summary results and cross reference to publications in peer reviews journals. It can be expected that they will gradually also include the name and links to a chosen data repository in near future thus completing the item 24 of the trial registration dataset regarding the intended sharing of deidentified individual clinical trial participant-level data (IPD). It now requires information about whether or not IPD will be shared, what IPD will be shared, when, by what mechanism, with whom and for what types of analyses. It consists of: data sharing plan (Yes, No) and plan description.

Patient Versus Trial Registries

The distinction between patient and trial registries might be confusing since they both capture certain disease-related information and also use Internet-based repositories. However, these two types of registries are quite different. Patient registries (discussed in Chap. 13) contain records and data on *individuals*, whereas trial registries focus on the descriptive aspects of a *research study* at various stages of its implementation and often provide a link to *study results*. While trial registries can be accessed via the WHO ICTRP global search portal, at present there is no single global search portal that can be used to identify or access patient registries.

Clinical trial registries contain information about ongoing and completed clinical trials, regardless of the disease or condition addressed. Patient registries contain the disease-specific information of individual patients. In a clinical trial registry, each entry represents one trial and contains defined protocol information of the trial. Clinical trials are prospective interventional studies, and they may recruit either healthy volunteers or patients with various diseases. Each trial may include any number from a few to thousands of participants. In a patient registry, each entry is an individual patient with the same disease or a condition of the same group, often chronic diseases (e.g., cancer, psychosis, and rare disease patient registries).

The most important difference between trial and patient registries is the purpose. The main goal of trial registries is to provide various stakeholders with information about ongoing and completed trials, in order to enhance transparency and accountability as well as to reduce the publication bias, increase the quality of published results, prevent harmful health consequences, and most importantly, provide knowledge that will ultimately enhance patient care. Patient registries, on the other hand, are developed in order to answer epidemiological questions such as incidence and prevalence and better understand the natural course of disease including morbidity or mortality.

Some trial registries also aim to inform potential trial participants about open or upcoming trials in order to enhance recruitment. Besides being tools for transparency, registries can also function as learning tools, as they might help improve the quality of the protocol and consequently the quality of the trials as they are completed. For example, while entering data in predefined fields, the researcher might realize that he or she is lacking some information (i.e., elements he or she forgot to define and include in the protocol) and will address the missing element(s) by editing and enhancing the protocol.

The first version of the protocol is the initial protocol that has been approved by the local ethics committee and at least its minimum dataset is submitted to the trial registry. Updates are expected and consist of providing information about the protocol in various stages of the trial: prior to recruitment, during the implementation (recruitment, interventions, follow-up), and upon completion. During trial implementation, changes of protocol, called *amendments*, often take place for various reasons. Amendments lead to new protocol versions, which are dated and numbered sequentially as version 2, 3, 4, etc. Annual updates of registry data enable posting of such amendments after approval by the ethics committees. The ability to manage multiple versions of protocol documents is an important feature for a trial registry. The basic rule for the registry is to preserve all of the descriptive data of a protocol that is ever received. Once registered, trials are never removed from the registry, but rather a *status* field indicates the stage of a trial (e.g., prior to recruitment, recruiting, do not recruit any more, completed). Earlier versions of protocol-related data are kept, are not overwritten, and should still be easily accessible by trial registry users.

WHO endorses trial registries that meet international standards and calls them *primary registries.* Registries that do not meet all the criteria of international standards are considered *partner registries*, and they provide data to the WHO search portal via one or more primary registries. The need for international access and utilization of registries implies the need for a common language. While some of these registries initially collect data in the language of the country or region, they provide data to the WHO portal in English because the WHO ICTRP currently accepts and displays protocol data in English only.

It is important to note that registries that adhere to international standards tend to add additional data fields to meet their registry-specific, often country-specific, needs. Regardless of these additional fields, the essential 24 items should always be included and well-defined. Although they are bound by the international standards, the presentation of a registry's website (i.e., the web-based access and query interface) is not the same across primary registries. Some registries collect and display protocol descriptive data beyond the basic predefined 24-item fields. Those registries that collect more data typically have more extensive and detailed data for each trial record and are potentially more useful for consumers. Some registries have free-text entry fields with instructions about which data need to be provided in the fields targeted to those registering their trials, while other registries employ self-explanatory and structured fields, such as drop-down lists. UMIN (University hospital

Medical Informatics Network), which is part of the Japanese primary registry, is the only primary registry that has fields for IPDs of trials it has registered [47, 48].

The WHO formed the Working Group on Best Practice for Clinical Trial Registries in 2008 to identify best practices, improve systems for entering new trial protocol records, and support the development of new registries [1]. The working group includes primary and some partner registries. Since the first edition of this book in 2012, 4 additional primary registries were developed, and as of February 2022, there are 18 registries that directly provide data to the WHO portal, specifically 17 WHO primary registries and the ClinicalTrials.gov registry which is not a part of the ICRTP primary registry network but it is accepted as a data provider. As can be seen from the geographic distribution shown in Fig. 17.2, the network includes at least one registry per continent.

Clinical trial registries can cross-reference a registered trial to its website if one exists; many large trials establish their own websites. Also, registries have been developing results databases and links and cross-references to publications in peer-reviewed journals, and it is expected that they will provide links to research data repository in which a given trial posted the anonymized individual participant data-IPD. The number and type of these links will likely increase as results databases and repositories continue to be developed.

Each registry displays the number of trials registered. This information is regularly updated, and WHO has been publishing the clinical trial statistics on its Global Observatory website [22]. Some registries accept trials from their country or region while some are open to register trials conducted in any country. We counted them by opening registry by registry and on the WHO website ClinicalTrials.gov is holding more than half of all

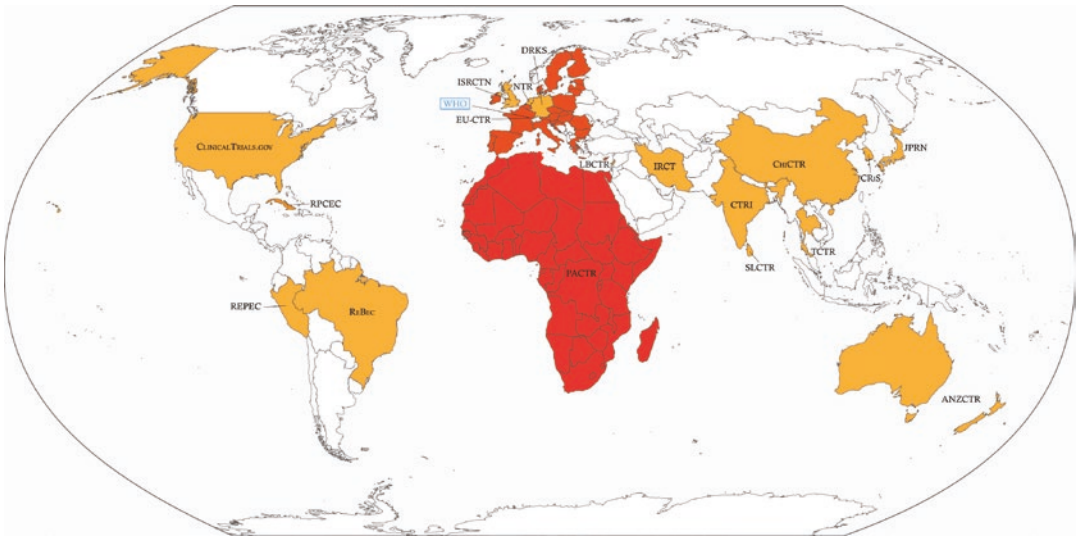

Fig. 17.2 Network of registries providing data to WHO ICRTP and its Search Portal. This map provides the worldwide distribution of registries that directly provided data to WHO as of July 2018. *ANZCTR* Australian New Zealand Clinical Trials Registry, *ReBec* Brazilian Clinical Trial Registry, *ChiCTR* Chinese Clinical Trial Registry, *CRiS* Clinical Research Information Service, Republic of Korea, ClinialTrials.gov (USA), *CTRI* Clinical Trials Registry, India, *EU-CTR EU* Clinical Trials Register, *RPCEC* Cuban Public Registry of Clinical Trials, *DRKS* German Clinical Trials Register, *IRCT* Iranian Registry of Clinical Trials, ISRCTN.org (UK), *JPRN* Japan Primary Registries Network, LBCTR The Lebanese Clinical Trials Registry, *NTR* The Netherlands National Trial Register, *PACTR* Pan African Clinical Trial Registry, *REPEC* Peruvian Clinical Trial Registry, *SLCTR* Sri Lanka Clinical Trials Registry, *TCTR* Thai Clinical Trials Registry, WHO Search Portal, Geneva. Note: The source of information: WHO ICRTP [1]. Since 2012 four registries, EU-CRT, TCTR, REPEC and most recently, in 2019 LBCTR joined the WHO primary registry network that directly provide data to WHO

registered trials. This phenomenon is partially due to the fact that the ClinicalTrials.gov registers trials from any country, and it is apparently often the secondary registry for trials registered in other primary registries. Also, the FDA requires a registration of a trial if the sponsor intends to get on the USA market based on the findings of given trial(s).

Timing

A responsible registrant, usually a specially delegated individual from the trial team or sponsoring organization, provides protocol-related data to the trial registry. Because all research protocols must be reviewed and approved by the ethics committee or board of the local institution, the descriptive protocol data set is usually submitted to the trial registry after institutional ethics approval. Otherwise, registration in the trial registry is considered conditional until the ethics approval is obtained.

Although international standards require registration prior to recruitment of trial participants, this has not been yet fully implemented [49–54]. Such prospective registration is important as it not only guarantees that all trials are registered but also that the initial protocol is made publicly available. For various reasons, the protocol might be changed early on, and/or a trial might be stopped within the first few weeks. Information about early protocol changes or stopped trials is lost unless trials are prospectively registered. Full data sharing is essential for the advancement of science and helps to avoid repeating such trials. Registries record the date of initial registration and date all subsequent updates. Additionally, the assignment and subsequent use of a unique ID for each trial upon registration enables any stakeholder to easily find what interests them.

Some countries hesitate to simply "import" the international standards or policies out of fear that these might change and put the country (regulator, or funding agency) in an odd position. One can debate the justification of such positions, but they are a reality. Implicit application of international standards occurs more often, with or without referencing them. Such is the case with the Declaration of Helsinki (DoH) [55], which obliges physicians via their national medical associations and is thus implicitly implemented. The DoH gradually addressed clinical trial registration and results disclosure, and the latest amended version from 2013, explicitly calls for the registration and results disclosure of trials that can be considered the good news for transparency of clinical trials with long term positive impact on knowledge development and health [56, 57].

Quality of Clinical Trial Registries

The quality of trial registries can be judged by the extent to which they meet the predefined goal of achieving high transparency of trials. Considering that meeting international standards is a prerequisite to qualify as a WHO primary registry, the quality and utility of trial registries mainly depend upon the quality and accuracy of data and the timing of reporting [1]. To realize research transparency, clinical trials need to be registered prior to the recruitment of trial participants. As already mentioned, the has not been fully implemented.

Registries constantly work on ensuring and improving the quality of data. The aim is to have data that are meaningful and precise. Accuracy of data requires regular updates in case of any changes and keeping track of previous versions. Registries impose some logical structure onto submitted data, but the quality is largely in the hands of data providers (i.e., principal investigators or sponsors). Many researchers and some registries perform analysis and evaluation of registry data [49, 58, 59]. IT experts might contribute by developing new, system-based solutions for quality control of entered trial data. Quality of data is a particularly sensitive issue as trial registries are based upon self-reporting by researchers, their teams, or sponsors. Following international standards and national requirements are prerequisites for attaining an acceptable level of data quality. Note that the practical and theoretical aspects of data quality are described in Chap. 10).

The numerous analyses and evaluations of implementation of standards and the quality of registries will enable revisions and updates, thereby improving trial registries at large. Furthermore, trial registries should reflect the reality of clinical trials methodology, which is constantly developing. Understandably, this presents a continuing challenge to those involved with the IT aspects of the data collection.

As registries that meet the WHO international standards might accept trials from any number of countries with data in the country's native language, it is essential to ensure the high quality of the translation of terms from any other language to English. Criteria that define quality also include transfer-related issues including coding and the use of standard terms, such as those developed by the Clinical Data Interchange Standards Consortium (CDISC) [60]. For this reason, definitions of English terms used across registries created in different countries also require standardization, and there have been efforts to this end, notably those on the standard data interchange format developed by CDISC. Standardization of terms is an important issue, and solutions must balance the resources required for researchers and trial registry administrators to implement standard coding against the potential benefits for information retrieval, interoperability, and knowledge discovery. The ability of protocol data to be managed and exchanged electronically, including difficulties with computerized representation due to various coding standards for several elements such as eligibility criteria, is described in Chap. 6.

A major concern for trial registries is the issue of duplicate registration. Duplicate registration of trials, especially of multi-center and multi-country trials, has been observed from the very beginning and was discussed by the WHO Scientific Advisory Group (SAG) while developing the standards. There are several reasons, starting with the initial concern that duplicate registration in registries, data providers to WHO ICRTP, might lead to counting one trial as two, or even as several trials, and might skew conclusions of systematic reviews. Therefore, these registries perform intra-registry deduplication process, while the WHO search portal established mechanisms of overall deduplication called *bridging*. In that process, most registries have created a field for an identification number (ID) that a particular trial was given by another registry. They usually also have the field for the ID from the source, which is assigned by the funder and/or sponsor. Parallel registration in a hospital, sponsor-based, or WHO partner registry does not count as duplicate registration; because those other registries have to provide their data to one WHO primary registry or ClinicalTrials.gov to meet criteria of international standards.

It is important to note that clinical trials are sometimes justifiably registered in more than one primary registry. For example, international trials might be registered in more than one primary registry if regulators in different jurisdictions require registration in specific registries. In these cases, researchers need to cross-reference IDs assigned from one registry to another. For this reason, the creation of a field in the registry to host the ID(s) received by other registries is of utmost importance. Also, it is important that researchers provide the same trial title and the same version of protocol information in case of duplicate registration. The latter is particularly important in case of delayed registration in one of the registries and/or initial data entry from a protocol that was amended in the meantime. Primary registries usually date the e-data entry, but it would be very useful to also number and date the protocol versions.

In 2009, as a part of implementing international standards, WHO established the universal trial number, UTN, and registries developed a field to host it. This number is also meant to help control duplicate registrations. While designing a registry, it is thus necessary to anticipate the field to host the UTN. Likewise, nonprimary registries as well as eventual trial websites should create fields for UTN, and the IDs assigned by primary registries.

Evolution and Spin-Off

The trigger for trial registration was the lack of transparency and the subsequent and disastrous health consequences shown by the New York

State Attorney General vs. Glaxo trial [25, 26]. That case mobilized stakeholders and elicited consequent action from various interest groups, i.e., journals, research communities, consumer advocates, regulators, including the already mentioned Ottawa Statement developed by the Ottawa Group [16] etc. Nowadays, trial registries aim to inform research and clinical decisions as well as to control publication bias in response to scientific and ethical requirements of research. As a result of the international dialogue among various stakeholders, most registries now aim to meet the needs of all involved in order to elevate research to another level.

Mandates for registries determine their scope, substance, and consequent design. Ever since the beginning, trial registries have been experiencing constant and rapid evolution, and the learning curve is steep for registrants, registry staff, registry users, and of course, IT professionals. The major impetus for the progress of trial registries followed the development of the WHO international standards in 2006 that expanded their scope from randomized controlled trials (RCTs) to all trials, regardless of the scope and type, and from a few items that indicated the existence of a trial to a 20 and then 24 items summary of the protocol and to fields hosting summary of results. At the same time, registries expanded fields and some of them started accepting trials from other countries. Initially, registration included only RCTs that aimed at developing new drugs and collected only basic information. Nevertheless, there is still significant potential for improvement. For example, many trials are still registered retrospectively or with a delay, but as this has been carefully studied and the negative consequences of retrospective registration have been presented, and as registries increasingly foster the prospective registration, it can be expected that it continues improving.

Further evolution of the international trial registration standards is expected to respond to the evolution of trial methodology. For example, phases 0, I, and II might need different fields, while some fields designed for RCTs no longer apply. This should be kept in mind while designing a registry.

The purpose and consequent use and appreciation of registration and registries has also evolved. Over time the main purpose of registries has been gradually shifting from a recruitment tool to a transparency tool while still focusing on benefits to trial participants. Some registries, such as ClincalTrials.gov, primarily originated from a mandate to enable potential trial participants to find a particular trial and to enrol in it. While registries continue to facilitate patients and clinicians searching for ongoing studies, they are also becoming a source of data on various completed trials and analysis of specific research areas and conditions such as surgical, cancer, or paediatric trials. For example, the analysis of registered trials informs about the health conditions and areas studied as in surgery emergency [61], critical care [62], patient reported outcomes [63], pregnancy and childbirth [64], and have been inspiring and guiding future research [65] and conducting registry-based trials [66].

Furthermore, trial registration has been contributing to the quality of protocols and thus studies, and to preventing selective reporting of study results. It has also been contributing to quality of systematic reviews and meta-analysis. As advised by the PRIZMA statement, systematic reviewers should search for unpublished clinical trial data in trial registries [67]. However, many reviewers have not been using registries' data which creates concerns about the potential bias and the quality of evidence produced by such systematic reviews and meta-analysis as it presented in many studies. Such underuse of registries is reported for example in nursing, critical care, pregnancy and childbirth and orthopaedic surgery [52, 62, 64, 68]. However, as such underuse of registries has been identified and reported, it can be expected that this will improve.

As the compliance with international standards is rather weak and selective when registration is voluntary, it is gradually becoming compulsory and required by journals, funders, regulators. Still, even when regulated, compulsory registration does not necessarily meet all the requirements of the WHO international standards. For example, in the USA, registration of a trial in ClinicalTrials.gov is required by law. Investigators must comply

or risk a penalty; however, the law does not require registration of all trials, and it allows a delay of 21 days for registration of trials that are covered by the Food and Drug Administration Amendments Act (FDAAA) of 2007 [69].

The experience gained so far is expected to inspire the registration of other types of studies or the development of other research-type registries. It is expected that such registries will function based on similar principles as trial registries and would inspire development of standards and creation of specific fields. One "spin-off" is already taking place and includes registration of observational studies in trial registries. Another example of a spin-off is a registry of systematic reviews of clinical trials and corresponding standards. The registry PROSPERO, international prospective register of systematic reviews [70], was launched in February 2011. PROSPERO is prospectively registering a systematic review (i.e., its design and conduct, protocol, or equivalent) and is displaying a link to eventual publication of the completed review. All the information is provided by the researcher and publicly displayed on PROSPERO's website. The registration and the usage are free of charge and freely accessible. Individual studies are the unit (record) of entry in such registries, and a mechanism for cross-referencing of study entries across various registries is established. It might be expected that a cross-referencing will be established btw systematic review registries and trial registries. Registries have been providing fields to capture results or link to various levels of reporting trial results and findings, such as links to publications, capturing aggregate results data in results fields, and linking to a database with microlevel data and registry of systematic reviews.

In addition to the WHO international trial registration standards, some countries develop their own specific standards, which may meet and expand or somewhat differ from the existing standards. For example, FDAAA differs by exempting the so-called phase I and some device trials from compulsory registration. Consequently, ClinicalTrials.gov offers fields for such trials, but their registration is voluntary. There are also initiatives to develop regional registries and software that will facilitate development of individual country registries in given regions such as in the Americas [71].

Creation and Management of a Trial Registry: The User Perspective

Design of Trial Registries

As mentioned earlier, every WHO primary trial registry now contains fields for a 24-item minimum data set as defined by the international standards and usually a few additional ones. These includes sthe fields for the ID assigned by any other registry, the unique trial registration number (UTRN) assigned by WHO, trial website URL, publications, etc. The required items are often expanded in several fields. For example, there may be special fields to indicate whether healthy volunteers are being recruited or to specify which participants are blinded. In parallel with registration of a minimum data set, arguments have been built for publishing the full protocol, and some journals have already started doing so. It will be particularly useful to have publicly available electronic versions of structured protocols, following SPIRIT guidelines. However, even if and when that happens, the data provided in trial registries will be very useful as a summary of the protocol. These two major tools of protocol transparency (trial registry and publicly available SPIRIT-based protocol) each attract different users but undoubtedly will provide a foundation for a number of navigation and analytic tools directed toward all users.

International Standards

WHO International standards have been the major impetus for the development of trial registries. Among other advantages, standards ensure the trustworthiness of data and comparability among registries. It is important that data provided is precise and meaningful, which depends on the precision of instructions for registration and also on the fields [49]. These instructions,

inspired by the WHO standards, might be developed by regulators in combination with the registry and/or journal editors as for example the Australian Clinical Trial Toolkit [72], Pan African registry explanations developed by Pienaar [73] and few others as shown on the Table 17.1. Registries usually have levels of compulsory completion of fields. Furthermore, they might

Table 17.1 Basic characteristics of WHO primary registries and Clinicaltrial.gov

	Instructions	Results	Data sharing plan	Link to publication
ANZCTR - Australia & NZ	✓	✓	✓	✗
ReBecv - Brazil	✗	✗	✗	✗
ChiCTR - China	✓	✓	✓	✗
CRiS - Republic of Korea	✓	✓	✓	✗
CTRI - India	✓	✓	✓	✓
RPCEC - Cuba	✓	✗	?	?
EU - CTR	✓	✓	?	✓
DRKS Germany	✓	✓	✓	✓
IRCT - Iran	✗	✗	✓	✗
ISRCTN	✓	✓	✓	✓
jRCT - Japan	✓	✓	✓	?
LBCTR - Lebanon	✗	✓	✓	✓
TCTR - Thailand	✗	✓	✓	✓
NTR - The Netherlands	✓	✗	✓	✓
PACTR - Pan African	✓	✓	✓	✗
REPEC - Peru	✓	✓	✓	✗
SLCTR - Sri Lanka	✗	✓	✓	?
ClinicalTrials.gov - USA	✓	✓	✓	✓

indicate which fields or items are required by the WHO standards and/or by the appropriate national regulator, or journals. It is important to note that at this time, there are no standards for registration of observational studies, so currently the trial fields are used to register observational studies and registries allow other descriptive data to be added.

Data Fields

The design of trial registries' fields is extremely important. Possibilities include free-text, drop-down, or predefined entries. It is advisable to define which data is needed and develop a drop-down list whenever possible. Such a drop-down list should include all known possibilities and the category "other" with text field to elaborate. Considering the rapidly developing field of clinical trials, it is necessary to anticipate additional items in a drop-down list.

Well-defined fields are prerequisite to obtain high-quality protocol data in trial registries. For example, if a registry field is free text and the data entry prompt reads *type of trial*, the answer will likely be simply "randomized controlled trial" or "randomized clinical trial" or even just the acronym "RCT." However, the registry might prespecify in a drop-down list whether the trial is controlled or uncontrolled and whether it is an RCT and whether its design is parallel, crossover, etc.

Although phases I–IV are still in use as descriptive terms, they will probably be replaced with more specific descriptions of studies in the future. Elaboration of those numbered phases is already taking place: the phase 0 has been added, and existing phases are subdivided into a, b, and c (e.g., phase II a, b, etc.). In some cases, two phases are streamlined into one study (e.g., I/II or II/III).

Other examples of terminology issues arise within the *Study Design* field, which might include allocation concealment (non-randomized or randomized) control, endpoint classification, intervention model, masking, or blinding, and who is blinded. Thus, in the case of RCTs, the trial registry data will not simply classify a study as an RCT but will also indicate if it is a parallel or crossover trial, which participants are blinded, whether the trial is one center or multi-center, and if the latter plans to recruit in one or several countries.

First-Level Fields

First-level fields cannot be skipped as they are aligned by the WHO International standards, and some might also be required by regulators. For example, ClinicalTrials.gov has fields that cannot be skipped because the FDAAA requires them. While designing a registry, one should keep in mind the possibility of expansion and provide a few fields for such unexpected information.

Second-Level Fields

Second-level fields are not made compulsory by some registries but are required by others. For example, as public funders or journal editors may require additional information beyond the international standards, there is an expectation that the relevant information will be provided by registrants; however, registries themselves cannot necessarily make these fields compulsory on their end, and consequently, some registries might not have these fields. Because adding fields to registries can sometimes be difficult, posting such additionally required information elsewhere in the registry is allowed. It may be placed along with or below other information or in the *Other* or *Additional information* field. For this reason, it is necessary to anticipate creation of such fields. For example, Canadian Institutes of Health Research (CIHR) requires the explicit reporting and public visibility of the ethics approval and confirmation of the systematic review justifying the trial [74].

Third-Level Fields

Third-level fields are optional and contain information that might be suggested by the registry, research groups, or offered by the researcher as important for a given trial. Such third-level data are usually entered in the *Additional information* field. This variation in fields means that, although there are international standards, there are differences among registries, specifically in the number of fields and their elaboration. There is still a

lot to learn in this process of ongoing evolution of registration and registries and the constant analysis and evaluation of current practices will point to better policies and practices in the future.

Trial Registry Features and Data Quality

In April 2022, we visited each WHO data provider registry and presented their basic features in Table 17.1. This includes the information whether the registry provides guidelines or instructions how to register a trial, data sharing plan, results, and a link to publications. As can be seen in the Table 17.1, several registries would merit further development in that respect.

As can be seen in Table 17.1, any form of results can be uploaded on 14 registries. However, we could not find any notion of sharing results in any form—not even a field in which applicants can share any information on the results in 4 primary registries (ReBEC, Brazil, RPCEC Cuba, ICTR Iran, NTR Netherland). However, it seems that evolution in results sharing is nevertheless moving forward, since only 3 (ReBEC, RPCEC, and EU CT) out of 17 primary registries lack data fields for clinical trials' applicants to provide information on their data sharing plan.

Instructions in the form of guidelines or learning modules are needed to ensure the quality of data entered. Consequently, many registries are developing such instructions to help researchers achieve better quality of data submitted as indicated in the Table 17.1. For example, the Australian New Zealand Clinical Trial Registry developed "data item definition and explanation" [72, 75]. International standards, the two countries' regulations, funders, and registries' policies all inform the content of this tool. Initial analysis of data entry in existing acceptable registries showed that a substantial amount of meaningless information was entered in open-ended text fields [76] but later studies have shown improvement in this area over time [58, 77, 78]. It is important to find a balance between general versus specific information. For example, indicating that the trial is blinded or double-blinded is much less informative than specifying who is blinded.

It may be expected that many registrants will do only what is required, which is often determined by regulations, policies of funders, World Medical Association's Declaration of Helsinki (DOH) or simply recommended by WHO international standards and by ICMJE instructions. Although SPIRIT has been helping researchers in protocol development and the WHO International Standards defined 24 protocol items that need to be posted in the registry, it has been observed that there are quality issues of registered data especially in case of multiple registration. Some of it might be explained by different timing of such multiple registration, reflecting a dynamic of trials protocols which often experience adaptations/ changes. The quality of data created a concern of research community and led to analysis and suggestions for improvement. For example, in the recent analysis of trial data in case of multiple registration, Speich and co-authors pointed to issues of inconsistency of key trial information for various reasons and expressed concerns about the reliability of information in registries [79]. Quality issues might also appear unrelated with multiple registration, as shown by the analysis of the European trial Registry by De Vito and Goldacre and Palludan-Muller et al. [80, 81].

Maintenance of Trial Registries

The researcher or sponsor of a trial provides annual updates of the trial record, all of which should be displayed in the registry. These updates should capture all amendments (i.e., changes of the protocol, the stage of trial implementation, eventual early stopping, etc.). It is important that registries have dedicated fields for updates and thus prevent overwriting previous information. Such an approach enables the identification of changes and tracks the flow of the trial implementation. The registry can be designed so that a reminder is sent automatically to registrants to send the annual update, and many have been already doing it. As mentioned earlier, registries develop special mechanisms of deduplication within the registry and with other registries.

Clinical Trial Results Databases/ Results Databases

Traditionally the main vehicle to disseminate trial results and findings in a trustworthy way has been via publication in a peer-reviewed journal. However, due to publication and outcome reporting bias, the availability of the Internet based informatics, and the ongoing initiative to avoid waste and get the most out of the research done, there is a growing international push to get findings and data publicly available for any future use.

Results databases in public domain are being developed predominantly by trial registries. Their development was a logical step on transparency road that starts with trial registration as the trial registration without publicly available results and IPD would be the empty promise. No wonder that the WHO started discussion and consultations about results data bases as early as 2008 [82].

Results databases are complex, and they might include aggregate data, metadata, and analysable data sets. Public disclosure of results in such databases is expected to complement the registration and publication in peer-reviewed journals, and it is an integral part of the transparency tool set. As noted by Zarin et al., results data base developed by trial registry is even providing information not available otherwise (even in peer review publication) [83]. It also contributes to avoiding the research waste as the only way clinical trial findings and results are disseminated if for any reason, they are not published in peer reviewed journal. Namely, although publication of trial results has been improved, the *publication bias*, i.e., the non-publication of clinical trial results is still taking place, as shown by Speich et al. [84], who reported that 21% of analysed trials were not published. It is important to note that even when trials are published in the peer review journals, the trial registries results databases are very useful for various stakeholders.

Primary registries have fields at least for summary results (participant flow; baseline characteristics; outcome measures and statistical analyses; adverse events) which are part of 24 items trial registration data set (TRDS) required by WHO standards. Results fields vary greatly between registries: tabular results, informal reports with little structure or a link to published works in journal or elsewhere—all of the mentioned meet the current TRDS criteria. Results dissemination is more and more consistently being considered as a non-negotiable issue—not different than disclosing ethics approval or informed consent [85]. WHO continuously calls all the stakeholders (ethics committees, regulatory agencies, professional bodies, sponsors, investigators, and funding agencies) to act in their jurisdictions to ensure that results from all interventional clinical trials are reported and publicly disclosed [86].

Results repositories are less developed than trial registries, and, as registries that host results databases are operated by national authorities, they differ in spite of trying to follow the WHO standards. As identified by the international meeting of the Public Reporting Of Clinical Trials Outcomes and Results (PROCTOR) group in 2008 [87] and discussed later on by us [11] especially in the IMPACT Observatory [88, 89], and by others [90, 91], numerous issues need to be resolved in order speed up getting the results data, especially having analysable datasets.

According to the WHO [23, 92–94] and many other stakeholders, including EU and USA legislators, the main findings of clinical trials are to be publicly available within 12 months of study completion by posting to the results section of the primary clinical trial registry and the publication in the peer review journals 24 months. If a primary registry has no results database available, the results should be posted on a free-to-access, publicly available, searchable institutional website of the regulatory sponsor, funder or principal investigator.

DeVito et al. report that the progress has been too slow for too long: clinicians, patients and the public cannot make informed decisions when the results of clinical trials are routinely withheld or incompletely reported [95]. Apparently, the compliance with the European Commission Requirement to post results 12 months after completion of the trial has been poor with half of

the trials non-compliant. Omissions, mislabelling and inconsistencies are common. The same group also indicated that commercially sponsored trials reported results and thus complied with the EU regulations better than non-commercial sponsored trials, sponsored by universities/academic organizations [21]. Such poor reporting was confirmed by several researchers [85, 96, 97]. The proportion of timely reported trials, even for trials funded by public funders, remains low [85, 96] and that's where funding agencies can propel real progress: beside adopting policies that require trial registration and reporting, provide budgetary [78] and other resources (e.g. staff, technical resources, software) to support registration and reporting [97].

Since the previous edition of this book, numerous efforts have been taking place aiming at improving the reporting of trials, by WHO ICRTP, registries, journals, researchers, and regulators. This led to major shifts in results' sharing landscape and new activities have been taking place. For example, in January 2022 European medicine Agency (EMA) launched Clinical Trials Information System (CTIS) [98]. It is a regulatory portal, trial registry, and results repository and it is supposed to help improve transparency as well as making trial authorisation more efficient and reducing administrative burden of sponsors.

Furthermore, the COVID-19 pandemic has emphasized the importance of building and maintaining trust in medicine and biomedical research, and that failure to fully, accurately, and rapidly report trial results could make needless harm to the credibility of medicine, industry and academia. This includes TranspariMed activities [14], the Scorecard, [19, 20], TrialTracker [99] and related research, all of which caused a substantial increase in just few years regarding results' reporting. For instance, following the Trial Tracker analysis nearly 80% of eligible results have been posted in the EU trial registry [95]. However, although huge improvement was made in total numbers of results appearing in the EU trial registry between 2018 and 2021, major gaps remain among some large sponsors and countries, leading Dal-Re et al. to conclude that enforcement of reporting regulations should be prioritized [100]. Several studies showed that despite the ongoing improvements reporting of results remains low and delayed [21, 84, 95, 101, 102], and that there is a gap between registration and results sharing as trial results are often withheld or not completely reported [95].

There are many advantages of sharing summary results via results databases hosted by trial registries. The prespecified form of clinical trials results for sharing through registers is helpful and it is in probably simpler and at the same time more efficient way to search for results in registries' databases than in publications. Furthermore, sharing of the results of clinical trials through databases can be a significant way to circumvent paywalls and thus promote open science. Besides, sharing clinical trials results via registries is not supposed to be as demanding and time consuming as publishing a peer-reviewed journal article [96] and funders could emphasize the publication of preprints as another way to increase timeliness of results dissemination [85]. Some registries are already providing summary results and a link(s) to publication(s), regardless whether it is open access or not. Coordinated action of all stakeholders will be crucial to reach the ultimate transparency grail to share all trial results in a timely manner.

Standards

There are no international standards for public disclosure of trial results, and there are no standards for preparing and use of the analyzable data sets, based on cleaned, anonymized individual participant data (IPD) and adjacent needed documentation (metadata, dictionary, etc.). However, there is much discussion on how these should be designed, and some initiatives have been contributing to accumulation of experience [56, 87, 103]. In 2010, the journal *Trials* started posting them on the Internet as the series "Sharing clinical research data," edited by Andrew Vickers. The topic of results disclosure actually includes a spectrum of information from aggregate (summary) data to fully analyzable, i.e., IPD-based data sets. In 2017, following several years of consensus building process that involved partici-

pants from various areas and backgrounds, the ECRIN leg of the CORBEL project developed a set of recommendations regarding clinical trial data sharing [91]. It is important to note that with the exception of UMIN registry, clinical trial registries enable the public disclosure of summary data and findings of clinical trials many of which are also published in peer-reviewed journals, while the IPD-based analyzable data sets are published in research data repositories which are increasingly in platforms.

Some of the outstanding challenges and disclosure issues regarding summary results and analyzable data are comparable to those of trial registration. These include the need to develop international standards, quality and completeness of data, timing of reporting, standardization of terms and following the research integrity rules. Other issues are more specific to the practical details of public disclosure of analyzable data sets, as discussed below.

Many of these issues suggest a need to develop levels of detail related to levels of access. In the era of electronic data management, some of these steps, such as cleaning of raw data, are becoming less of an issue as they take place simultaneously with the data collection. Furthermore, much can be learned from other areas especially from the experience of genome data sharing, which has boosted the development of the field [104, 105].

A lot has changed since the first version of this chapter published in 2012 [12], when these data were either protected in the hands of regulators or might have been shared with systematic reviewers only upon request and only under certain conditions. Meanwhile many constituencies engaged in making data available, especially in order to facilitate systematic reviews and meta-analysis that include IPD data sets. For example, journal editors have been increasingly encouraging data sharing upon publication of trial findings in their respective journals [106]. As of 2018, ICMJE requires the authors to provide a data sharing statement while submitting a manuscript with trial results http://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html.

Sharing of Clinical Trial Data, Research Data Repositories and Platforms

Sharing of clinical trial data is becoming more and more appealing to all stakeholders. Earlier hesitation has been gradually lightening, and we are witnessing increased transparency and a consecutive change of the research paradigm. Although many issues have yet to be resolved, this area is constantly and rapidly evolving [78, 107–109].

Repositories, i.e., research data repositories, are electronic databases hosting research raw data and facilitating their reuse. They are the newest research transparency tool complementing trial registries and results databases and they are increasingly included in platforms which facilitates the use and analysis of data.

When talking about data sharing from clinical trials, we are talking about the cleaned anonymized individual participant data (IPD) sets and adjacent documentation forming the analyzable data. In this chapter we are focusing on sharing of analysable IPDs and adjacent info and repositories or platforms hosting it. One can say that it has been a step wise approach to clinical trial transparency starting with trial registration via results disclosure and publications to sharing of analysable datasets. Their relationship and interplay is illustrated in the Fig. 17.4.

As abundance of data is produced during clinical trials, there is a growing need to provide expert service in organizing and curating these data, which emphasize the need for medical information specialists as for example librarians specialised in managing medical data.

However, several dilemmas and issues are still present and will require research and resolution. These include the lack of standards on how to prepare analyzable datasets for public sharing, heterogeneity of repositories, and finding the balance of privacy versus transparency [88] and the willingness to enforce the requirements. These include the cleaning of data, quality of data, accountability, defining which adjacent documentation is needed, who is the guarantor of truth, intellectual property rights, privacy issues/

anonymization, and issues related to anonymization efforts. These elements create specific challenges, require interdisciplinary work, and at the same time present an opportunity for clinical research informatics and information technology experts.

General research data platforms and repositories can be classified by the scientific area they cover or the level (university, region, country, international) at which they are organized. Re3data [110] classifies them into disciplinary, institutional, and other. Some of repositories hosting clinical trial data are based at universities and accept data only from researchers from a given university or consortium, such as Edinburgh Data Share [111] or DRUM (Data Repository for the University of Minnesota) [112]. Most general open-access repositories in public domain host data from any research.

There are also disease-specific repositories and research data repositories organized by public funders, such as several repositories run by the NIH institutes- including BioLINCC. However, only a small portion of them host clinical trial data and only Vivli [8] and the Japanese trial registry UMIN [47] host exclusively clinical trial data.

Various groups and entities have been analysing and listing repositories aiming at informing researchers where to publish their data including journals PLoS, Nature, Science, etc. In our ongoing study we have been searching for open access repositories in public domain that also host clinical trial data and analysed their basic features [113–115]. It is interesting to note that repositories we initially identified (in 2016) were hosting data from industry which reflects the important role of clinical trials in development of new medicines. For example, Datasphere was organized by pharmaceutical industry with the aim to share cancer trial data, Yoda was a partnership btw the Yale university and 3 companies, while the Wellcome Trust coordinated the application of 13 companies that shared their data via Clinical Study Data Request, CSDR.

Public funders have an important role in furthering transparency of clinical trials. One of the recent initiatives took place in January 2022, when NIH Office of Data Science Strategy (ODSS) launched the Generalist Repository Ecosystem Initiative (GREI) [116] with the aim to further data sharing and improve access to data from its funded studies and it will supplement the NIH Domain repositories. The six established repositories that are part of GREI are expected to work together to establish consistent metadata, develop use cases for data sharing, train and educate researchers on FAIR data and the importance of data sharing, and more. Following six repositories are included in the GREI: Dryad, Dataverse, Figshare, Mendeley Data, Open Science Framework, and Vivli.

Vivli is the only repository of GREI group hosting exclusively clinical trials data while few other repositories host clinical trial as well as data from other types of research- including Dryad that accepts data if the research is published, Figshare that accepts data from anywhere, and Dataverse, which is an open-source web application to share, preserve, cite, explore, and analyse research data [117]. There are 33 Dataverse repositories (installations) around the word, and one of them, Harvard Dataverse, also hosts clinical trial data [118]. Mendeley data is special as it has been indexing several open research repositories including Zenodo.

It is interesting to note that Vivli was also identified as the only clinical trial domain repository hosting data from any clinical trial and from anywhere by the 2019 study that analysed clinical trial data repositories [48].

There has been an ongoing effort to further the clinical trial data sharing and over time many entities have been contributing to increasing of clinical trial transparency in various ways. WHO added summary results and IPD sharing statement to trial registration dataset and Declaration of Helsinki (DoH) calls for registration and results sharing [56]. Journal editors play a very important role in transparency of clinical trials. Following the WHO requirement for trial registration, in 2018 ICMJE developed policy requesting that submitted manuscripts of trial starting on or after January 2019 contain data sharing plan and precise what that plan should include. http://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html#one.

As data management should begin at data collection, public funders are increasingly demanding that the data management plan be developed up front. This leads to the understanding that the data preservation and storage of academic trials starts at the academia, that the institution—academia conducting a trial should anticipate data sharing and act accordingly—preferably develop a database and then might send data to established repositories. Indeed, several universities have been doing this. One of the first was the Edinburgh University that established Edinburgh DataShare repository which also hosts clinical trial data. It started with a JISK project led by Edinburgh University in partnership with two other UK universities (Oxford and Southampton). While it initially hosted data from the international stroke trial, it is now hosting data from other studies conducted at the Edinburgh University [111]. The key role in setting and running of this repository has been played by research librarians. Actually, management and storage of research data have become a field of interest of research librarians, and it can be expected that they will be increasingly engaged in this field.

Some repositories hosting clinical trial data might limit the uploading of data to members of a given university or consortium, but all of them enable open access to data for secondary use. There is usually a limited control of data quality at entry and no curatorship of data already in the repository. As a general rule, repositories rely on the clinician trialist—data provider to clean, anonymize, and organize data for publication.

Several specific projects and software have been influencing developments of this field in various ways. For example, Research Data Alliance (RDA) aims at building the social and technical infrastructure to enable open sharing of data. It functions through interest and working groups that elaborate specific topics and provide recommendations for the community [119]. Related tools to data sharing by repositories include persistent identifiers/PID, DataCite, re3data, and the CoreTrustSeal of certification organization [120].

re3data is a registry of research data repositories from various academic disciplines. In 2014 it merged with another similar tool, Databib, and it is now managed by DataCite. Re3data registers repositories from various disciplines and describes basic features of each of them. "It presents repositories for the permanent storage and access of data sets to researchers, funding bodies, publishers, and scholarly institutions. re3data.org promotes a culture of sharing, increased access, and better visibility of research data. The registry went live in autumn 2012 and it is funded by the German Research Foundation, DFG [110, 121].

Citability and findability of published data are very important. Among other benefits, they stimulate public data sharing. Citability and to certain extend findability are achieved by assigning the *persistent identifier (PI or PID)* to published data sets. PID is a long-lasting reference to a document, file, web page, or other object. The term "persistent identifier" is usually used in the context of digital objects that are accessible over the Internet. Once plugged in the web browser, it will link to related data sets which enables citation of given data sets [122]. Persistent identifiers help the research community locate, identify, and cite research data with confidence.

DataCite is a leading global non-profit organization that provides persistent identifiers (DOIs) for research data [123]. DataCite assigns DOI persistent identifier to each repository registered in re3data. Repositories in turn assign persistent identifier to hosted data sets, i.e., data sets published in them. In our ongoing scanning of general repositories within the IMPACT Observatory we noticed that most of the open access general repositories in public domain that host clinical trial data assign DOI, or some other PID [114].

The research community realized the importance of ensuring the quality of repositories, and in 2017, the *CoreTrustSeal* certification organization was established, developed by *the ICSU World Data System (WDS) and the Data Seal of Approval (DSA)* under the umbrella of RDA. The CoreTrustSeal has a set of criteria that a given repository has to meet [120]. *The* re3data indicates for each indexed repository whether it is certified or whether it supports repository standards.

Anonymization Methods of Clinical Research Data

The sharing of clinical research data for secondary purposes will require that the data be anonymized. A secondary purpose if one for which consent was not obtained from the patient. For example, real world data that was collected in the course of providing care, and that is then used for research purposes, the research would typically be considered a secondary purpose. As another example, if data is collected under patient consent for a specific clinical trial, and then that data is used to answer a different research question, then that other research question would be a secondary purpose.

There is no legislative requirement to obtain patient consent for the secondary purpose for using the data if the data is deemed to be anonymized. This is the case in many jurisdictions today. The act of converting original personally identifying information into anonymized information would also not require explicit consent from the patients in most jurisdictions.

When a sponsor or academic institution wishes to share their data with other external parties or wishes to reuse the data internally for a secondary purpose, then the data should be anonymized. Anonymization can be applied to clinical reports (such as clinical study reports from clinical trials) or to microdata (which are individual level data i.e. IPD datasets). The techniques will be different between the two. Anonymization techniques for documents have been discussed elsewhere [124] and will not be covered here as they involve additional topics related to information extraction.

The anonymization process encompasses multiple technologies and here we will describe two of them: risk-based anonymization and synthetic data generation. These are examples of privacy enhancing technologies (PETs). There are other PETs that can be applied to share microdata or to enable the analysis of microdata in a privacy preserving manner, such as secure-multi-party computation and different forms of federated analysis. However, these are at an earlier stage of adoption and the use case assumed here is of sharing microdata with an analyst.

Managing Identity Disclosure Risk in Microdata

One of the basic types of disclosure risks is identity disclosure. This is defined as the risk of correctly assigning a person's identity to a record in a dataset. It is also referred to as the identifiability of a record.

The identifiability of a record in the dataset falls on a spectrum as shown in Fig. 17.3. This

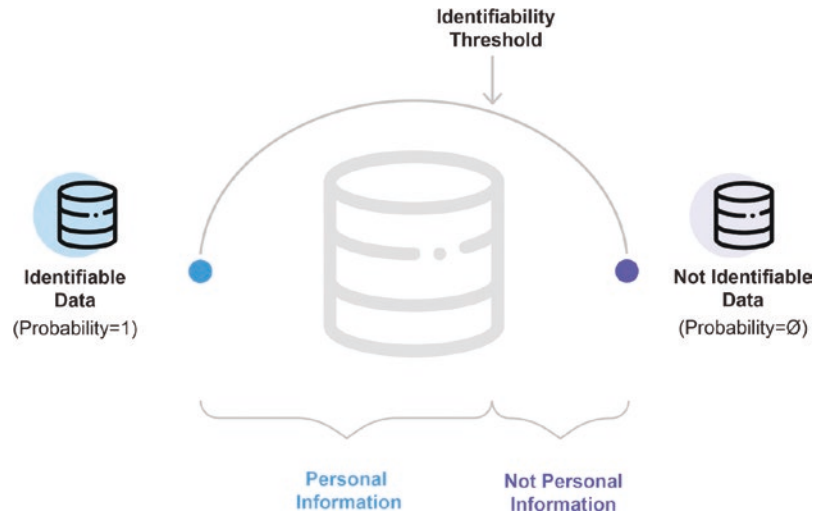

Fig. 17.3 The spectrum of identifiability

spectrum can be thought of as a probability that varies from zero to one. If the probability is one then that record is definitely identifiable (e.g., it has a name and address associated with it). If the probability is zero, then that record is not identifiable whatsoever.

There is a large body of literature that has developed over the last 50 years or so focused on methods for quantitatively assessing where on this spectrum a record, or a whole dataset, falls [125].Therefore, the ability to measure or estimate this probability exists.

Because privacy regulations conceptualize identifiability as a binary concept (a record or a dataset is personally identifiable or not; it is personal information or not), this spectrum can be split by a threshold value. If the measured identifiability is above the threshold, then a record is considered to be identifiable. If the measured identifiability is below the threshold, then a record is considered not to be identifiable.

The threshold will not be zero. This means that zero identifiability is not a realistic standard that can be achieved. If one wants zero identifiability, then no data can be used or disclosed. When data is used or disclosed, there will be some risk of re-identification. Even random matching of individuals to records in a dataset has a non-zero probability of being correct. The goal is to ensure that the risk is low enough to be acceptable. The threshold represents the level of risk that is acceptable for the given dataset.

Because data custodians have been sharing data for many decades, there are many precedents for the choice of a threshold. Therefore, that is not a controversial point since the precedents come from reputable organizations globally and they have worked well to protect datasets. A commonly used risk threshold value is 0.09, especially in the context of health data [126, 127].

Other Risks in Microdata

When PETs such as synthetic data generation (SDG) are used [126, 128] the risk of identity disclosure will tend to be low. This is because with SDG the dataset is generated from a model and therefore there is no one-to-one mapping between the synthetic records and real records. There are other types of risks that need to be managed, such as attribution disclosure and membership disclosure.

Attribution disclosure is when there is a digital twin for a real person in the synthetic data and we can learn something new about the real person from that digital twin.

Membership disclosure is when we can learn that an individual is in the real data using the synthetic data. Both of these are types of inferences about individuals. We are not identifying their records because that concept does not fit well with synthetic data, but we are learning something new about them from the synthetic data.

This highlights the point that the relevant types of disclosures will be dependant on the PET that is being used, and the appropriate type of disclosure risk should be assessed. Furthermore, because most privacy laws only consider identity disclosure risks, methods like SDG will generally be deemed to be anonymized information.

In deciding which PETs to apply to create anonymized datasets, it is also important to consider the perspectives of regulators. With multiple known re-identification attacks [129], there is an expectation that advanced PETs will be used to reduce the overall privacy risks when data are used and disclosed for secondary purposes.

The User Perspective of Registration-Results-Data Sharing Process

Several repositories that host clinical trial data are open for hosting data from certain groups of researchers, usually those linked to a given university, or area, but all of them allow open access to data they host. The lack of standards and heterogeneity of repositories makes the analysis of hosted data across several repositories very difficult if not impossible, without contacting the original data provider. It can be expected that the interest and the need for reanalysis will trigger development of needed standards. Such standards

should be developed by the research community, not by repository. Ideally, internationally renowned organizations, such as WHO, will lead standards' development and include key stakeholders in the consensus building process, as was the case with development of the trial registration standards.

Evolution and Future Directions of Sharing of Trials Results

The future of clinical research and informatics is closely interwoven, and it can be expected that these evolving fields will mutually inform and influence each other. Methods for achieving clinical trial transparency—including the sharing of analyzable data sets—are still lagging behind other research areas. Technical, social, and political barriers—some of which are specific to clinical trials—remain and will present exciting challenges for researchers, information technology experts, and other stakeholders to advance existing tools and develop sustainable strategies for public disclosure of trial information—from protocol via results to data. These challenges will include the stewardship and reuse of such data for the creation of new knowledge, which will in turn speed development of new and more powerful diagnostics and therapeutics. It can be expected that this will further the interest of librarians to engage in health research informatics.

One indicator of dynamics are increased activities and new initiatives since the previous edition of this book, including GREI, the Clinical Trials Information System (CTIS), the European new registry that we discussed above, and especially a development of the Vivli platform for sharing clinical trial data.

In 2019, building on the previous activities of CORBEL and IMPACT, the group of researchers including one of us, searched for repositories that host clinical trial data and analysed 25 including a sample of NIH disease specific repositories [48]. In that study we identified only two repositories exclusively hosting clinical trial data: UMIN and Vivli. UMIN [47] is the clinical trial registry, part of Japanese primary registry that also hosts IPDs of trials that were registered in it thus being the only primary registry hosting CT data.

Vivli [8] on the other hand is a data sharing and analytics platform, open to host data from any clinical trial from anywhere and facilitating their re-analysis. It started in 2018 and at the time of the above mentioned 2019 study it was in the early stage of development, but by 2022 it established partnership with 40 member organizations, including academia, public funders and industry and has been hosting data from more than 6000 trials from various countries. As already mentioned, Vivli is one of six generalist repositories invited by the NIH to join the GREI (Generalist Repository Ecosystem Initiative).

It is anticipated that data flow from trials to the public domain and the linking and cross-referencing of related data will create a more efficient system of information sharing and knowledge creation as presented in Fig. 17.4. Although it has not yet been completely accomplished, there is a clear tendency to move in that direction, which will ensure a high level of transparency, getting closer to open data and open science.

The Fig. 17.4 illustrates actions that need to be taken from registering trial elements, posting results in results databases and IPDs in repositories. It is expected that existing systematic reviews will be updated with the meta-analysis of IPD-based analyzable data to inform various levels of decision-making with the updated evidence. Finally, in an ongoing effort to increase transparency of research and to build on the experience of trial registries, other types of studies are being registered in trial registries, and other types of research registries are being developed. However, although there are no standards and guidelines for the preparation of clinical trial data for public release and although repositories are heterogeneous, the existence of platforms hosting open-access repositories is a big step forward toward opening of clinical trial data.

Trial registries host defined protocol items, and they are in constant evolution, from the elaboration of fields to the establishment of

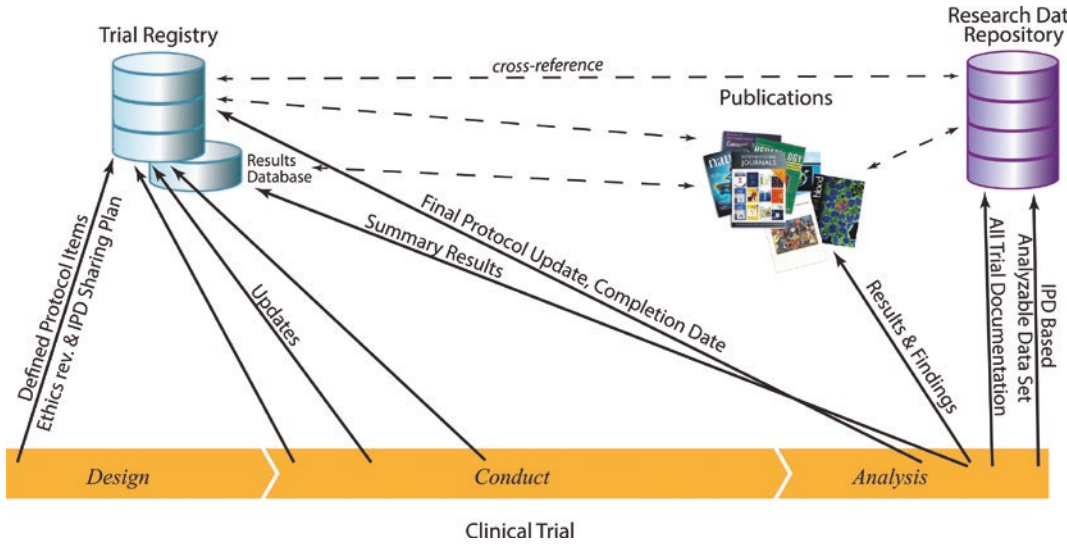

Fig. 17.4 Anticipated flow of data from clinical trial to public domain. *Please note* that the major change of this flow of data took place by the establishment of open-access research data repositories in public domain

hyperlinks. It can be expected that the analysis and evaluation of the existing primary registries' experience will inform the best practice and potential expansion of the data included, like adding fields to host more data than required by the current 24-item international standards [44].

Furthermore, there is a strong push for publication of the full protocol, either in the registry or elsewhere. It will certainly be particularly useful to have publicly available electronic versions of structured protocols, following SPIRIT guidelines. Even if this were to happen, the protocol data set that is available in registries will continue to provide valuable summaries of protocols with links to other trial related information including the full protocol, publications, trial website, systematic review, meta-analysis, results databases and research data repositories and thus continue to play an important role in achieving trial transparency.

Results databases have also evolved. They are being formed by trial registries and aim at providing summary/aggregate results data of registered trials. As shown on the Table 17.1, 14 of 18 open-access registries in public domain that are linked to the WHO, can host results although at different level of precision. As there are no international standards these registry-based results databases differ. Each of them follows the rules of their respective countries, and at the same time, they are trying to meet the WHO and ICMJE request to register and share summary results. Apparently, the need to synchronize has been understood, and it seems that ClinicalTrials.gov and EMA/European Clinical Trial Registry are working on developing comparable data fields which might inform future development of international standards of data sharing. It is important to note that even at this stage of development sharing summary results via open access results databases has numerous advantages: prespecified form contributes to clarity and makes posting and finding of results simpler and more efficient and enables access to basic information free of charge.

Coordinated action of all stakeholders will be crucial to reach the ultimate transparency grail—share all the results in a timely manner.

Open-access research data repositories in public domain are certainly the most important tool for data opening and can play a major role in enabling public availability of research data. However, they are heterogenous, and there are still no international standards to govern the public disclosure of analyzable data sets which include cleaned, anonymized IPDs (i.e., usually

numeric or encoded) and documentation sufficient to make the data reusable.

Development of such standards will require participation of all interested constituencies in thorough planning, analysis of quality control, resources, as well as dealing with specific issues, such as privacy, i.e., anonymization methods and practices.

It is important to note that although there are currently no standards and guidelines for the preparation and publication of clinical trial data for public release and although repositories are heterogenous, the existence of open-access repositories and a possibility to publish data in them are a big step forward toward opening of clinical trial data. Initiatives and projects addressing the needed standards development as mentioned CORBEL project by ECRIN [91] are encouraging. Clinical research Metadata Repository, MDR, is the most recent initiative of this group with the goal to facilitate searching and selection of trial data so that researchers could quicky identify data object of interest [130].

The progress achieved as well as the interest and expectations this data opening process has created so far is reassuring but still a lot needs to be done. As mentioned earlier, there are numerous initiatives contributing to increasing the transparency of clinical trials and opening of its data. As already discussed, various researchers' groups have been analysing the situation and suggesting improvements. One of the oldest is IMPACT Observatory [17] that evolved from the Ottawa statement [16]. It can be expected that this process will be observed and supported by key players at different levels, including regulators, public funders, clinicians, academia, pharmacists, journal editors, industry, patients, consumers, consumer advocates, and general public. Thus, researchers and IT experts will not be alone in this process as the clinical trials and their contribution to creation of the evidence needed for decisions in health are of paramount interests to numerous stakeholders.

The dynamics of the process are so immense and complex that they merit assessment of actions, initiatives, and practice of various players and their interactions. It is equally important to assess the impact of these dynamics on making analyzable data publicly available for reuse, on the consequent transformation of clinical trial research and on all adjacent issues. An observatory or natural experiment is the methodology of choice to collect, assess, and disseminate such data and thus inform the process and indicate trends. The IMPACT Observatory aims to do just that and serve as a tool, a hub, informing the process of opening of trial data [17, 115].

Various groups have been analysing research data repositories aiming at figuring which ones to advise the researcher to go to. Their lists often overlap, listing the same repositories. However, some of them are disease or country or sponsor specific. Overall, Vivli [8] is the only clinical trial domain repository accepting data from clinical trials conducted anywhere and it is included in most if not all of those lists. One might conclude that unless a trial is performed in Japan and registered in the UMIN registry, or it has to publish its data in one of NIH domain specific repositories, VIVLI is a repository of choice to publish clinical trial data and also to look for data for further analysis.

Transparency is realized by making research protocols, results, and cleaned and anonymized IPDs publicly available using well-defined, freely accessible electronic tools: registries, results data bases and research data repositories. These tools are at different levels of development and plagued with heterogeneity as international standards for trial registration do not yet address the sharing of results and individual patient data. One interesting possibility is that trial registration should develop the results and IPD sharing standards. These tools and related standards need to co-evolve and there are efforts in that direction starting with the ongoing evolution of existing standards and tools. Different types of research stakeholders have been developing electronic tools to measure transparency of clinical trial information—aiming at reducing waste, maintaining research integrity, and improving the quality and volume of research to address important and emerging health questions.

Conclusion

Clinical Trial (or clinical research) transparency is realized by making research protocols, results, and cleaned and anonymized IPDs publicly available using well-defined, freely accessible electronic tools that have emerged over time. It has unquestionable value for research, for evidence needed for decision-making, aiming at improving the health of people. Although data sharing tools and standards are still at different levels of development and plagued with heterogeneity as international standards exist only for trial registration, pushed by the many research stakeholders and evaluation tools they developed, trial transparency is in the process of further co-evolving toward the aims to reduce waste, maintain research integrity, and improve the efficiency and volume of research to address important and emerging health questions.

Although there are still hurdles to overcome, the ongoing evolution is encouraging, and most importantly there is *more than* sufficient knowledge about when, what, how, and where to register clinical trials, how to post the results and publish IPDs, to support clinical reporting moving forward. As illustrated in Fig. 17.4, a commitment and compliance with clinical trial reporting will ensure a future when all relevant trial information is publicly available for future use in electronic tools—clinical trial registries, results databases managed by these registries, and research data repositories or platforms.

Test Your Learning

After reading this chapter, you should be able to do the following.

Patient/study participants should be able to find

(a) the trials he/she might join
(b) the results of a study she/he participated in as well as of studies of interest

Researchers should

(a) have the ability to design, conduct and report the trial in timely manner, respecting the research integrity
(b) have the ability to collaborate with IT in creation, further development (expansion) of fields by providing/suggesting the information that needs to be included
(c) know when, where, how, to register clinical trial protocol, report the results, publish analysable datasets (IPDs) and provide cross referencing and links with peer review publications
(d) be aware of the potential multiple use of clinical trial registries' information especially for further research.

The researcher–clinical trialist should

(a) be able to choose the registry among the WHO primary registries and/or Clinicaltrial.gov, find and follow registry's instructions/guidelines how to register a trial and select protocol items that should be registered
(b) be able to select protocol items and register them in the chosen registry (one of the WHO primary registries or Clinicaltrial.gov) prior to recruiting the first trial participant to meet the requirements of the international standards and in some cases identify and add additional protocol information required by registry.
(c) be able to make results, findings and analysable IPDs publicly available in a timely manner
(d) know when and how to upload summary results of clinical trial in the registry in which the trial was registered
(e) be able to find instructions in the registry of choice to prepare and post/publish the summary results of their study in the registry/ies in which he/she registered a trial.
(f) be able to choose the appropriate data repository/platform for a given trial and upload analysable datasets.

Researcher–Systematic reviewer should

(a) look for and use data from trial registries and results database while preparing the systematic review

(b) look for trials in trial registries and look for and use IPDs from research data repositories and platforms preparing the secondary analysis/meta-analysis

IT/informatics expert should

(a) be able to develop data fields of the registry following international standards and if needed, following (to meet) regulators requirements
(b) create the registry/ results database/data repository webpage
- user friendly
- with instructions and guidelines visible, easy to acMcess, clear

(c) collaborate/exchange with other primary registries/ to have similar setup.

(Public) Funders:

Should be able to stimulate trials in under-researched area/areas of public health importance that have not been studied, based on the analysis of ongoing and recently performed studies.

Acknowledgments The authors would like to thank Nevena Jeric, Apropo Media for graphic design.

Disclaimer The views expressed here are the authors' and do not represent the views of any organization.

References

1. WHO. International Clinical Trials Registry Platform (ICTRP) [Internet]. [cited 2022 Mar 28]. https://www.who.int/clinical-trials-registry-platform.
2. WCRIF. Mission—World Conferences on Research Integrity [Internet]. [cited 2022 May 10]. https://wcrif.org/foundation/mission.
3. Vickers AJ. Sharing raw data from clinical trials: what progress since we first asked "Whose data set is it anyway?" Trials [Internet]. 2016;17:227. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4855346/.
4. Krleža-Jerić K. Clinical trial registration: the differing views of industry, the WHO, and the Ottawa Group. PLoS Med. 2005;2:1093–7.
5. Clinical Study Data Request [Internet]. [cited 2016 Aug 31]. https://www.clinicalstudydatarequest.com/.
6. Project Data Sphere [Internet]. [cited 2016 Jul 31]. https://www.projectdatasphere.org/.
7. YODA Project [Internet]. [cited 2016 Jul 18]. http://yoda.yale.edu/.
8. Vivli—A global Clinical trial data sharing platform: proposal, definition and scope background and objectives. 2016;(June). https://vivli.org.
9. European Medicines Agency—Clinical data publication—Documents from advisory groups on clinical-trial data [Internet]. [cited 2018 Aug 20]. http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/document_listing/document_listing_000368.jsp&mid=WC0b01ac05809f3f12#section1.
10. Krleža-Jerić K. Sharing of data from clinical trials and research integrity. In: Steneck N, Anderson M, Kleinert S, Mayer T, editor. Integrity in the Global Research Arena; Proceedings of the World Conference on Research Integrity; 3rd, Montreal 2013. World Scientific Publisher, Singapore. p 91.
11. Krleža-Jerić K. Sharing of clinical trial data and research integrity. Period Biol. 2014;116(4):337–9.
12. Krleža-Jerić K. Clinical trials registries and results databases. In: Richesson RL, Andrews JE, editors. Clinical research informatics. London: Springer; 2012. p. 389–408.
13. Chan AW, Hróbjartsson A. Promoting public access to clinical trial protocols: challenges and recommendations. Trials. 2018;19(1):116.
14. Clinical Trial Transparency | Transparimed [Internet]. 2019 [cited 2022 Mar 30]. https://www.transparimed.org/.
15. Krleza-Jeric K. Clinical trial registries, results databases, and research data repositories. In: Richesson RL, Andrews JE, editors. Clinical research informatics. 2nd ed. Springer International; 2019. p. 453–480.
16. Krleza-Jerić K, Chan AW, Dickersin K, Sim I, Grimshaw J, Gluud C. Principles for international registration of protocol information and results from human trials of health related interventions: Ottawa statement (part 1). BMJ [Internet]. 2005 [cited 2016 Jul 13];330(7497):956–8. http://www.ncbi.nlm.nih.gov/pubmed/15845980.
17. Krleza-Jeric K. IMPACT Observatory [Internet]. MedILS. [cited 2022 Jun 4]. https://www.medils.org/research/impact-observatory.
18. RDA. About RDA | RDA [Internet]. 2016 [cited 2022 Jun 3]. https://www.rd-alliance.org/about-rda.
19. Good Pharma Scorecard | Bioethics International [Internet]. 2009 [cited 2022 Mar 30]. https://bioethicsinternational.org/good-pharma-scorecard/.
20. Miller J, Ross JS, Wilenzick M, Mello MM. Sharing of clinical trial data and results reporting practices among large pharmaceutical companies: cross sectional descriptive study and pilot of a tool to improve company practices. BMJ. 2019;366:l4217.
21. Goldacre B, Devito NJ, Heneghan C, Irving F, Bacon S, Fleminger J, et al. Compliance with requirement to report results on the EU Clinical Trials Register: cohort study and web resource. BMJ [Internet].

2018 [cited 2021 Nov 26];362:3218. https://doi.org/10.1136/bmj.k3218.
22. WHO Global Observatory Clin trials. Number of clinical trials by year, country, region and income group [Internet]. 2022 [cited 2022 May 10]. https://www.who.int/observatories/global-observatory-on-health-research-and-development/monitoring/number-of-clinical-trials-by-year-country-who-region-and-income-group.
23. WHO News. Joint statement on public disclosure of results from clinical trials [Internet]. 2017 [cited 2022 May 27]. https://www.who.int/news/item/18-05-2017-joint-statement-on-registration.
24. Simes RJ. Publication bias: the case for an international registry of clinical trials. J Clin Oncol [Internet]. 1986 [cited 2016 Jul 13];4(10):1529–41. http://www.ncbi.nlm.nih.gov/pubmed/3760920.
25. Bass A. Side effects; A prosecutor, whistelblower, and a bestselling antidepressant on trial. 1st ed. Algonquin Books of Chapel Hill; 2008. 260 p.
26. Gibson L. GlaxoSmithKline to publish clinical trials after US lawsuit. BMJ. 2004;328(7455):1513.
27. DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the international committee of medical journal editors. JAMA [Internet]. 2004 [cited 2016 Jul 12];292(11):1363–4. http://www.ncbi.nlm.nih.gov/pubmed/15355936.
28. World Health Organisation. Ministerial Summit on Health Research Report by the Secretariat. 2005.
29. International Clinical Trials Registry Platform (ICRTP) [Internet]. [cited 2016 Jul 11]. https://www.who.int/clinical-trials-registry-platform.
30. Mahmić-Kaknjo M, Šimić J, Krleža-Jerić K. Setting the impact (Improve access to clinical trial data) observatory baseline. Biochem Med [Internet]. 2018;28(1):010201. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5806612/.
31. ISRCTN. ISRCTN Registry [Internet]. [cited 2022 Jun 2]. https://www.isrctn.com/.
32. NIH-NLM. ClinicalTrials.gov [Internet]. [cited 2022 May 27]. https://clinicaltrials.gov/.
33. Cho SM, Serghiou S, Ioannidis JPA, Klassen TP, Contopoulos-Ioannidis DG. Large pediatric randomized clinical trials in ClinicalTrials.gov download large pediatric randomized clinical trials in ClinicalTrials.gov. Pediatrics. 2021;148(3):e2020049771.
34. Chan AW, Tetzlaff JM, Altman DG, Dickersin K, Moher D. SPIRIT 2013: new guidance for content of clinical trial protocols. Lancet [Internet]. 2013 [cited 2022 Mar 26];381(9861):91–2. http://www.thelancet.com/article/S0140673612621606/fulltext.
35. Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J a, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ [Internet]. 2013;346:e7586. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3541470&tool=pmcentrez&rendertype=abstract.
36. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7.
37. Emam K el, Jonker E, Sampson M, Krleža-Jerić K, Neisa A. The use of electronic data capture tools in clinical trials: web-survey of 259 Canadian trials. J Med Internet Res [Internet]. 2009 [cited 2022 Mar 26];11(1). /pmc/articles/PMC2762772/.
38. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan AW, King MT. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols the SPIRIT-PRO extension. JAMA [Internet]. 2018;319(5):483–94. https://jamanetwork.com/.
39. Dai L, Cheng CW, Tian R, Zhong LL, Li YP, Lyu AP, et al. Standard Protocol Items for Clinical Trials with Traditional Chinese Medicine 2018: recommendations, explanation and elaboration (SPIRIT-TCM Extension 2018). Chin J Integr Med [Internet]. 2019 [cited 2022 Mar 27];25(1):71–9. https://pubmed.ncbi.nlm.nih.gov/30484022/.
40. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, Ashrafian H, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digital Health. 2020;2(10):e549–60.
41. Porcino AJ, Shamseer L, Chan AW, Kravitz RL, Orkin A, Punja S, et al. SPIRIT extension and elaboration for n-of-1 trials: SPENT 2019 checklist. BMJ [Internet]. 2020 [cited 2022 Mar 27]; https://doi.org/10.1136/bmj.m122. http://www.bmj.com/.
42. EQUATOR. The EQUATOR Network | Enhancing the QUAlity and Transparency Of Health Research [Internet]. [cited 2022 May 26]. https://www.equator-network.org/.
43. https://www.icmje.org/about-icmje/faqs/clinical-trials-registration/
44. WHO trial registration data set version 1.3.1. [Internet]. [cited 2022 Mar 24]. https://www.who.int/clinical-trials-registry-platform/network/who-data-set.
45. Bhatia P, Mohammed S. Observational study design: extending the standard. J Anaesthesiol Clin Pharmacol [Internet]. 2021 [cited 2022 Jan 7];37(3):317–8. https://pubmed.ncbi.nlm.nih.gov/34759537/.
46. Bradley SH, Lloyd KE, Devito NJ. Automatic registration for UK trials. A welcome development, not a panacea. BMJ [Internet]. 2022 [cited 2022 Jan 29]; https://doi.org/10.1136/bmj.o41.
47. UMIN-ICDR Individual Case data repository [Internet]. [cited 2018 Aug 11]. http://www.umin.ac.jp/icdr/index.html.
48. Banzi R, Canham S, Kuchinke W, Krleza-Jeric K, Demotes-Mainard J, Ohmann C. Evaluation of repositories for sharing individual-participant data from clinical studies. Trials [Internet]. 2019 [cited 2022 Feb 12];20(1). https://pubmed.ncbi.nlm.nih.gov/30876434/.

49. Reveiz L, Chan AW, Krleza-Jerić K, Granados CE, Pinart M, Etxeandia I, et al. Reporting of methodologic information on trial registries for quality assessment: a study of trial records retrieved from the WHO search portal. PLoS One [Internet]. 2010;5(8):e12484. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0012484.
50. Reveiz L, Krleza-Jerić K, Chan AW, de Aguiar S. Do trialists endorse clinical trial registration? Survey of a PubMed sample. Trials [Internet]. 2007;8:30. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2147029&tool=pmcentrez&rendertype=abstract.
51. Harriman SL, Patel J. When are clinical trials registered? An analysis of prospective versus retrospective registration. Trials [Internet]. 2016;17:187. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4832501/.
52. Gray R, Brown E, Gray G. A review of prospective trial registration in the Journal of Advanced Nursing in 2018. J Adv Nurs [Internet]. 2019 Oct 1 [cited 2021 Nov 15];75(10):2051–3. https://onlinelibrary.wiley.com/doi/full/10.1111/jan.14090.
53. Venugopal N, Saberwal G. A comparative analysis of important public clinical trial registries, and a proposal for an interim ideal one. Plos One [Internet]. 2021. https://doi.org/10.1371/journal.pone.0251191.
54. Taylor NJ, Gorman DM. Registration and primary outcome reporting in behavioral health trials. BMC Med Res Methodol. 2022;22(1):41.
55. WMA Declaration of Helsinki—Ethical Principles for Medical Research Involving Human Subjects—WMA—The World Medical Association [Internet]. 2013 [cited 2018 Aug 19]. https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/.
56. Krleza-Jerić K, Lemmens T. 7th revision of the Declaration of Helsinki: good news for the transparency of clinical trials. Croat Med J [Internet]. 2009 [cited 2016 Jul 13];50(2):105–10. http://www.ncbi.nlm.nih.gov/pubmed/19399942.
57. Goodyear MDE, Krleza-Jeric K, Lemmens T. The Declaration of Helsinki. BMJ [Internet]. 2007 [cited 2016 Jul 13];335(7621):624–5. http://www.ncbi.nlm.nih.gov/pubmed/17901471.
58. Zarin DA, Tse T, Williams RJ, Rajakannan T. The status of trial registration eleven years after the ICMJE policy. N Engl J Med [Internet]. 2017;376(4):383–91. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5813248/.
59. Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. Ioannidis J, editor. PLoS Med [Internet]. 2008;5(11):e217. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2586350/.
60. CDISC [Internet]. [cited 2018 Aug 19]. https://www.cdisc.org/.
61. Morley RL, Edmondson MJ, Rowlands C, Blazeby JM, Hinchliffe RJ. Registration and publication of emergency and elective randomised controlled trials in surgery: a cohort study from trial registries. BMJ Open [Internet]. 2018 Jul 1 [cited 2022 Mar 29];8(7). https://pubmed.ncbi.nlm.nih.gov/29982216/.
62. Greiner B, Corcoran A, Wheeler D. Clinical trial registry searches are under-utilized in systematic reviews from critical care journals: a bibliometric analysis. J Crit Care. 2021;63:175–8.
63. Riedl D, Rothmund M, Darlington AS, Sodergren S, Crazzolara R, de Rojas T. Rare use of patient-reported outcomes—in childhood cancer clinical trials—a systematic review of clinical trial registries. Eur J Cancer [Internet]. 2021 [cited 2021 Nov 14];152:90–9. http://creativecommons.org/licenses/by-nc-nd/4.0/.
64. Yerokhin VV, Carr BK, Sneed G, Vassar M. Clinical trials registries are underused in the pregnancy and childbirth literature: a systematic review of the top 20 journals. BMC Res Notes. 2016;9(1):475.
65. Dear RF, Barratt AL, McGeechan K, Askie L, Simes J, Tattersall MHN. Landscape of cancer clinical trials in Australia: using trial registries to guide future research. Med J Aust [Internet]. 2011 [cited 2022 Mar 29];194(8):387–91. https://pubmed.ncbi.nlm.nih.gov/21495937/.
66. Ecker BL, Brajcich BC, Ellis RJ, Ko CY, D'Angelica MI. Registry-based randomized clinical trials in surgery: working with ACS-NSQIP and the AHPBA to conduct pragmatic trials. J Surg Oncol [Internet]. 2022 [cited 2022 Jan 4];125(1):89–92. https://onlinelibrary.wiley.com/doi/10.1002/jso.26742.
67. Page MJ, Mckenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ [Internet]. 2021. https://doi.org/10.1136/bmj.n71.
68. Reddy AK, Anderson JM, Gray HM, Fishbeck K, Vassar M. Clinical trial registry use in orthopaedic surgery systematic reviews. J Bone Joint Surg Am [Internet]. 2021 [cited 2022 Mar 28];103(10):e41. https://pubmed.ncbi.nlm.nih.gov/33983151/.
69. Food and Drug Administration Amendments Act (FDAAA) of 2007 [Internet]. Office of the Commissioner; [cited 2018 Aug 21]. https://www.fda.gov/regulatoryinformation/lawsenforcedbyfda/significantamendmentstothefdcact/foodanddrugadministrationamendmentsactof2007/default.htm.
70. Prospero-International prospective register of systematic reviews [Internet]. [cited 2018 Aug 20]. https://www.crd.york.ac.uk/prospero/.
71. Krleža-Jeric K, Lemmens T, Reveiz L, Cuervo LG, Bero LA. Prospective registration and results disclosure of clinical trials in the Americas: a roadmap toward transparency. Rev Panam Salud Publica [Internet]. 2011 Jul [cited 2016 Jun 15];30(1):87–96. http://www.ncbi.nlm.nih.gov/pubmed/22159656.
72. Australia Clinical Trials Toolkit | Australian Clinical Trials [Internet]. [cited 2018 Aug 21]. https://www.australianclinicaltrials.gov.au/clinical-trials-toolkit#overlay-context=home.

73. Pienaar E. Cochrane SA webinar: clinical trial registration and the Pan African Clinical Trials Registry—YouTube [Internet]. 2018 [cited 2022 May 7]. https://www.youtube.com/watch?v=BIMHoXXqyTA.
74. Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans [Internet]. TCPS 2. 2014. www.nserc-crsng.gc.ca.
75. Australia New Zealand Clinical Trials Registry. Data item definition/explanation [Internet]. [cited 2018 Aug 13]. http://www.anzctr.org.au/docs/ANZCTRDatafieldexplanationV5.pdf.
76. Zarin DA, Tse T, Ide NC. Trial registration at ClinicalTrials.gov between May and October 2005. N Engl J Med [Internet]. 2005;353(26):2779–87. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1568386/.
77. Askie LM, Hunter KE, Berber S, Langford A, Tan-Koay AG, Vu T, Sausa R, Seidler AL, Ko H SRJ. The Clinical Trials Landscape in Australia 2006–2015 [Internet]. Sydney: Australian New Zealand Clinical Trials Registry; 2017 [cited 2018 Aug 18]. 83 p. http://www.anzctr.org.au/docs/ClinicalTrialsInAustralia2006-2015.pdf#page=1&zoom=auto,557,766.
78. Zarin DA, Califf RM. Trial reporting and the clinical trials enterprise. JAMA Intern Med [Internet]. 2021 [cited 2022 Mar 14];181(8):1131–2. https://pubmed.ncbi.nlm.nih.gov/34028521/.
79. Speich B, Gloy VL, Klatte K, Gryaznov D, Taji Heravi A, Ghosh N, et al. Reliability of trial information across registries for trials with multiple registrations: a systematic review + supplemental content. JAMA Network Open [Internet]. 2021 [cited 2022 Jan 29];4(11):2128898. https://jamanetwork.com/.
80. 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 [Internet]. 2022 [cited 2022 Mar 19]. https://pubmed.ncbi.nlm.nih.gov/35144496/.
81. Paludan-Müller AS, Maclean-Nyegaard IR, Munkholm K. Substantial delays in clinical data published by the European Medicines Agency—a cross sectional study. J Clin Epidemiol [Internet]. 2022;146:68–76. https://doi.org/10.1016/j.jclinepi.2022.02.004.
82. Ghersi D, Clarke M, Berlin J, Gülmezoglu AM, Kush R, Lumbiganon P, et al. Reporting the findings of clinical trials: a discussion paper. Bull World Health Organ [Internet]. 2008;86(6). http://www.icmje.
83. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database—update and key issues. N Engl J Med [Internet]. 2011;364(9):852–860. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3066456/.
84. Speich B, Gryaznov D, Busse JW, Gloy VL, Lohner S, Klatte K, et al. Nonregistration, discontinuation, and nonpublication of randomized trials: A repeated metaresearch analysis. Kesselheim AS, editor. PLoS Med [Internet]. 2022 [cited 2022 May 20];19(4):e1003980. https://dx.plos.org/10.1371/journal.pmed.1003980.
85. Riedel N, Wieschowski S, Bruckner T, Holst MR, Kahrass H, Nury E, et al. Results dissemination from completed clinical trials conducted at German university medical centers remained delayed and incomplete. The 2014–2017 cohort. J Clin Epidemiol [Internet]. 2022 [cited 2022 May 5];144:1–7. https://pubmed.ncbi.nlm.nih.gov/34906673/.
86. Moorthy VS, Karam G, Vannice KS, Kieny MP. Rationale for WHO's new position calling for prompt reporting and public disclosure of interventional clinical trial results. PLoS Med. 2015;12(4):e1001819.
87. Krleza-Jeric K. International dialogue on the public reporting of clinical trial outcome and results—PROCTOR meeting. Croat Med J. 2008;49:267–8.
88. Krleža-Jerić K, Gabelica M, Banzi R, Krnić-Martinić M, Pulido B, Mahmić-Kaknjo M, et al. IMPACT Observatory: tracking the evolution of clinical trial data sharing and research integrity. Biochem Med (Zagreb) [Internet]. 2016 [cited 2022 Apr 3];26(3):308–17. https://pubmed.ncbi.nlm.nih.gov/27812300/.
89. Mahmić-Kaknjo M, Šimić J, Krleža-Jerić K. Setting the IMPACT (Improve access to clinical trial data) observatory baseline. Biochem Med [Internet]. 2018 [cited 2022 Jun 4];28(1):010201. https://www.biochemia-medica.com/en/journal/28/10.11613/BM.2018.010201.
90. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. Lancet [Internet]. 2014 [cited 2016 Jul 14];383(9913):257–66. http://www.ncbi.nlm.nih.gov/pubmed/24411650.
91. Ohmann C, Banzi R, Canham S, Battaglia S, Matei M, Ariyo C, et al. Sharing and reuse of individual participant data from clinical trials: principles and recommendations. BMJ Open [Internet]. 2017;7(12):e018647. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5736032/.
92. World Health Organization. WHO statement on public disclosure of clinical trial results [Internet]. 2015 [cited 2016 Jul 11]. http://www.who.int/ictrp/results/reporting/en/.
93. Woodcock J. FDA Takes action for failure to submit required clinical trial results information to ClinicalTrials.gov | FDA [Internet]. 2021 [cited 2022 Mar 30]. https://www.fda.gov/news-events/press-announcements/fda-takes-action-failure-submit-required-clinical-trial-results-information-clinicaltrialsgov.
94. European Medicines Agency. European Medicines Agency policy on publication of clinical data for medicinal products for human use [Internet]. 2019 [cited 2022 May 12]. https://www.ema.europa.eu/en/documents/other/european-medicines-agency-policy-publication-clinical-data-medicinal-products-human-use_en.pdf.

95. DeVito NJ, Goldacre B. New EU trial reporting regulations must be enforced. BMJ. 2022;376:o410.
96. Strzebonska K, Wasylewski MT, Zaborowska L, Riedel N, Wieschowski S, Strech D, et al. Results dissemination of registered clinical trials across Polish academic institutions: a cross-sectional analysis. BMJ Open. 2020;10:34666.
97. Mayo-Wilson E, Heyward J, Keyes A, Reynolds J, White S, Atri N, et al. Clinical trial registration and reporting: a survey of academic organizations in the United States. BMC Med. 2018;16(1):60.
98. European Medicine Agency. Clinical Trials Information System | European Medicines Agency [Internet]. 2022 [cited 2022 May 4]. https://www.ema.europa.eu/en/human-regulatory/research-development/clinical-trials/clinical-trials-information-system.
99. EU Trials Tracker—Who's not sharing clinical trial results? [Internet]. [cited 2022 Mar 30]. https://eu.trialstracker.net/.
100. Dal-Ré R, Goldacre B, Mahillo-Fernández I, DeVito NJ. European non-commercial sponsors showed substantial variation in results reporting to the EU trial registry. J Clin Epidemiol. 2022;(142):161–70.
101. Gao L, Guo R, Han Z, Liu J, Chen X. Clinical trial reporting. Lancet [Internet]. 2020 [cited 2022 May 7];396(10261):1488–9. http://www.thelancet.com/article/S0140673620322650/fulltext.
102. Chan AW, Hróbjartsson A. Promoting public access to clinical trial protocols: challenges and recommendations. Trials [Internet]. 2018. https://doi.org/10.1186/s13063-018-2510-1.
103. Bian ZX, Wu TX. Legislation for trial registration and data transparency. Trials [Internet]. 2010;11(1):64. https://doi.org/10.1186/1745-6215-11-64.
104. Collins F. Has the revolution arrived? Nature [Internet]. 2010;464(7289):674–5. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5101928/.
105. Collins FS, Green ED, Guttmacher AE, Guyer MS. A vision for the future of genomics research. Nature [Internet]. 2003;422:835. https://doi.org/10.1038/nature01626
106. Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. Ethiop J Health Sci. 2017;27(4):315–8.
107. Gøtzsche PC. Why we need easy access to all data from all clinical trials and how to accomplish it. Trials [Internet]. 2011;12:249. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3264537/.
108. Rockhold F, Bromley C, Wagner EK, Buyse M. Open science: the open clinical trials data journey. Clin Trials [Internet]. 2019 [cited 2022 Mar 19];16(5):539–46. https://journals.sagepub.com/doi/10.1177/1740774519865512.
109. Rockhold F, Nisen P, Freeman A. Data sharing at a crossroads. NEJM. 2016;375:1115–7.
110. Re3Data; Registry of Research Data Repositories [Internet]. [cited 2018 Aug 11]. www.Re3data.org.
111. Edinburgh DataShare [Internet]. [cited 2018 Aug 18]. https://datashare.is.ed.ac.uk/.
112. DRUM. Data Repository for U of M (DRUM) [Internet]. [cited 2022 Jun 12]. https://conservancy.umn.edu/drum.
113. Krleza-Jeric K, Hrynaszkiewicz I. Environmental scan of repositories of clinical research data: how far have we got with public disclosure of trial data? [Internet]. figshare; 2018. https://figshare.com/articles/Environmental_Scan_of_Repositories_of_Clinical_Research_Data_How_Far_Have_We_Got_With_Public_Disclosure_of_Trial_Data_/5755386.
114. Gabelica M, Martinic MK, Luksic D, Krleza-Jeric K. Clinical trial transparency and data repositories; an environmental scan of the IMPACT (Improving access to clinical trial data) Observatory [Internet]. figshare; 2018. https://figshare.com/articles/Clinical_trial_transparency_and_data_repositories_an_environmental_scan_of_the_IMPACT_Improving_Access_to_Clinical_Trial_Data_Observatory/7390559.
115. Krleža-Jerić K, Gabelica M, Banzi R, Martinić MK, Pulido B, Mahmić-Kaknjo M, et al. IMPACT observatory: tracking the evolution of clinical trial data sharing and research integrity. Biochem Med (Zagreb) [Internet]. 2016;26(3):308–17. http://www.ncbi.nlm.nih.gov/pubmed/27812300%0A, http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5082220.
116. NIH. NIH Office of Data Science Strategy Announces New Initiative to Improve Access to NIH-funded Data | Data Science at NIH [Internet]. 2022 [cited 2022 May 28]. https://datascience.nih.gov/news/nih-office-of-data-science-strategy-announces-new-initiative-to-improve-data-access.
117. The Dataverse Project [Internet]. [cited 2018 Aug 18]. https://dataverse.org/.
118. Harvard Dataverse [Internet]. [cited 2018 Aug 19]. https://dataverse.harvard.edu/.
119. Research Data Alliance RDA [Internet]. [cited 2018 Aug 16]. https://www.rd-alliance.org/about-rda/who-rda.html.
120. CoreTrustSeal [Internet]. [cited 2018 Jun 27]. https://www.coretrustseal.org/about/.
121. Pampel H, Vierkant P, Scholze F, Bertelmann R, Kindling M, Klump J, et al. Making research data repositories visible: the re3data.org registry. PLoS One. 2013;8(11):e78080.
122. Persistent Identifier [Internet]. [cited 2018 Aug 18]. https://en.wikipedia.org/wiki/Persistent_identifier.
123. DataCite [Internet]. [cited 2018 Aug 11]. https://www.datacite.org/index.html.
124. Scaiano M, Middleton G, Arbuckle L, Kolhatkar V, Peyton L, Dowling M, et al. A unified framework for evaluating the risk of re-identification of text de-identification tools. J Biomed Inform [Internet]. 2016 [cited 2022 May 10];63:174–83. https://doi.org/10.1016/j.jbi.2016.07.015.

125. Willenborg L, de Waal T. Elements of statistical disclosure control. New York: Springer New York; 2001. (Lecture Notes in Statistics; vol. 155).
126. European Medicines Agency. European Medicines Agency policy on publication of clinical data for medicinal products for human use: policy 0070 [Internet]. 2014 [cited 2022 May 10]. www.ema.europa.eu/contact.
127. Health Canada. Guidance document on public release of clinical information: profile page—Canada.ca [Internet]. 2019 [cited 2022 May 10]. https://www.canada.ca/en/health-canada/services/drug-health-product-review-approval/profile-public-release-clinical-information-guidance.html.
128. El Emam K, Lucy Mosquera RH. Practical synthetic data generation balancing privacy and the broad availability of data. Sebastopol: O'Reilly Media, 2020 [Internet]. 2020 [cited 2022 May 10];166. https://www.oreilly.com/library/view/practical-synthetic-data/9781492072737/.
129. El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. PLoS One [Internet]. 2011. www.plosone.org.
130. MDR. ECRIN-MDR Wiki [Internet]. 2020 [cited 2022 Jun 14]. http://ecrin-mdr.online/index.php/Project_Overview.