How to make Clinical Trials Registry-India world class

Gayatri Saberwal

All clinical trial registries probably have some lacunae concerning the number and quality of their records. This article deals with the Clinical Trials Registry-India (CTRI). It describes 10 categories of problems with the data and some proposed solutions. In their registration practices, some trialists do not follow the rules or may break the law. The need to comply with WHO requirements is emphasized, as well as the need for CTRI to go beyond what is currently required by WHO. With improvement in the organizational functioning of CTRI, it could be amongst the best registries in the world.

Keywords: Clinical Trials.gov, CTRI, data integrity, ethics, registry, WHO.

BIOMEDICAL innovations need to be tested in human volunteers to determine their safety and efficacy. These tests are called clinical trials. A trialist needs to register a trial with a clinical trial registry. Such registries are repositories of information related to planned or ongoing trials. The records help patients to identify trials that they may wish to enroll in[1]. The registration also helps solve another problem. There is a known tendency to report trial results only if they are positive, leading to a bias in the literature. Registration ensures that all trials – regardless of the outcome – can be considered in any systematic review that could influence clinical practice. Further, the registry records provide a wealth of information to a range of stakeholders, such as medical doctors, scientists, funders and policymakers. The data in public trial registries have been useful in multiple ways, such as disseminating information, advancing science, improving healthcare, tracking biomedical innovation, tracking the globalization of trials, etc.[2]. As such, registries contribute to increased transparency and accountability of the clinical research ecosystem, to reducing research waste and to knowledge sharing[3].

The World Health Organization (WHO), Geneva, Switzerland, recognizes 18 public trial registries as data providers to its International Clinical Trials Registry Portal (ICTRP)[4]. Of these, 17, including the Clinical Trials Registry-India (CTRI), are also recognized as primary registries. ClinicalTrials.gov (CT.gov), the registry of the United States, is only recognized as a data provider. Such registries are prone to errors. The various categories of inaccuracies include missing data[5], nonsensical data[6], discrepancies in the records of the same trial in different registries[7], and so on. CTRI records have their share of such problems[8].

Gayatri Saberwal is in the Institute of Bioinformatics and Applied Biotechnology, Biotech Park, Electronics City Phase 1, Bengaluru 560 100, India.
e-mail: gayatri@ibab.ac.in

Here I summarize some of the challenges with the records or functioning of CTRI and suggest ways to handle them. The points detailed below fall into five categories: (a) problems with the data; (b) trials that are not in compliance with a rule, or a law, from the viewpoint of registration with CTRI; (c) the need for primary registries to comply with WHO requirements; (d) the desirability for CTRI to go beyond what is currently required by WHO, and (e) improving the organizational functioning of CTRI. Some issues have already been documented in the literature by my team or others. In other cases, which are based on our unpublished work, CTRI numbers are provided as illustrations. Yet others are in the form of suggestions.

Below, the fields in a trial record are italicized.

Problems with the data

Here, I enumerate various problems with the CTRI data, such as missing data, internal inconsistencies, confusion over definitions, incomplete or non-standard information, etc. These kinds of problems complicate the understanding of a given trial. Also, in case a data-analytics approach is used to analyse a large number of records, data need to be in a systematic form, without ambiguities, and therefore these problems are likely to compromise such analyses.

Missing data

Several fields have been found to have no entries. Missing data have been reported for the name of the principal investigator (PI), and *Countries of Recruitment*, *Name of Primary Sponsor*, etc.[8].

I want to discuss missing data for *Enrollment* specifically. In preliminary work a few years ago, we looked into whether participants from India have been over-represented in multinational interventional drug trials. We wanted to examine the planned versus actual recruitment from the

country. However, the CTRI records were not updated with the final enrollment for 281 of 606 (46%) trials. There have been fears that Indians could be exploited in multinational trials[9]. Given this, the planned and final enrollments must be available in the CTRI record on time.

A complex classification of 'Types of Study'

For interventional trials, an important fact that one would want to know is the nature of the intervention. CT.gov has 11 distinct categories, including biological, combination product, drug, genetic, etc. A given trial may list a particular category multiple times (such as when multiple drugs are being tested) or several of these categories. In CTRI, it appears that registrants enter information in a free text field. This has resulted in over 1000 categories, many of them atypical[8].

Further, in a study of 12,673 trials, the incidence of 'not available' in *Types of Study* was the second highest, after 'drug'[8] despite a significant fraction of trials being interventional. As such, even though this is an important field, filling it appears optional in CTRI records.

Internal inconsistencies

Trial records may also have internal inconsistencies[8]. An example is a record that lists BA/BE in *Type of Trial* but a *Phase* between 1 and 4.

Confusion over definitions

Although CTRI provides documentation to help trialists understand the various fields in a given trial record, it appears that registrants may still be confused. Illustratively, the entry in *Phase* should be between 1 and 4, but some trials list post-marketing surveys instead[8]. It has also been reported that interventional trials are listed in CTRI as observational trials due to a lack of understanding of the terms 'interventional' and 'observational'[10].

Incomplete or non-standard information

An example of non-standard information is found in the classification of cities. Illustratively, we have found that trials may have the following names for cities: West (as illustrated by CTRI/2018/02/012091), North West (CTRI/2017/11/010690), Central (CTRI/2020/09/027829), etc. These refer to parts of Delhi, with Delhi itself listed as a state.

Variations in the names of individuals or organizations

The name of a PI may be represented in various ways in different trial records. Box 1 lists trials that illustrate this problem. A person's name may also change with time, without indicating that the two names represent the same person. Further, multiple individuals may share a name. These issues complicate any attempt to identify all the trials that a given PI has run.

Organization names may also be represented in multiple ways, causing similar complications during analysis.

Variations in the classification of organizations

As done by other public trial registries, CTRI too provides a field to classify the primary sponsor. Several categories include contract research organisations, the pharmaceutical industry (Indian and global), research institutions, and others. However, there is lack of clarity on how to classify some organizations. For instance, is the Indian subsidiary of a multinational company a pharmaceutical industry – global or a pharmaceutical industry – Indian? How about an Indian pharmaceutical company that has become global? The All-India Institute of Medical Sciences has been classified as a government funding agency (CTRI/2020/06/026151), a government medical college (CTRI/2017/03/008137), a research institution (CTRI/2018/03/012510), as well as a research institution and hospital (CTRI/2017/12/010746). All these classifications are correct but unhelpful for any analysis of the number of sponsors of a given category.

CT.gov has a tighter set of four categories of sponsors: the US government's National Institutes of Health, other US federal agencies, industry, and all others (individuals, universities and organizations). The larger number of categories in CTRI is useful, but their utility needs to be increased by providing greater clarity on the definition of each category, by ensuring that a given organization is listed under only one category, and by minimising the use of the category 'Others'.

Incomplete or incorrect details of ethics committees

Each trial record in CTRI lists one or more sites and one or more ethics committees (ECs). However, a given site is not mapped to a given EC, and it can be challenging, if not impossible, to identify which EC cleared the trial at a given site. There are several other problems with data in the EC field. Illustratively (i) for a given trial, the number of ECs may exceed the number of sites, and (ii) an EC cannot be unambiguously identified because it is referred to in a casual manner (CTRI/2018/05/014249). We have documented about 30 categories of problems with the EC or EC-site data[11].

Messy data

Various issues may cause data to be unclear or confusing. Examples of messy data include the following: (a) The same acronym is used for different organizations. (b) An acronym is not spelled out anywhere in the record. (c) A

Box 1. Examples of the name of a principal investigator (PI), represented in various ways in different trial records.

CTRI number	PI
CTRI/2019/05/019303	Dr Amar Raikantiwar
CTRI/2019/05/019355	Dr Amar Raykantiwar
CTRI/2016/08/007192	Manisha Talekar
CTRI/2017/12/010837	Dr Talekar Manisha Tatobaji
CTRI/2017/01/007705	Dr Narendrakumar B. Mundhe
CTRI/2017/06/008790	Dr Narendrakumar Bhanudas Mundhe
CTRI/2020/09/027914	Dr Narendra B. Mundhe

trial in which the site was listed twice with the same PI. (d) A trial in which a site was duplicated in order to list two ECs, one of which belonged to the institution performing the study and the other to the Council to which the institution belongs[11]. Such issues could lead to errors, especially if a data analytics approach is used to handle the data.

Misleading information

Below, I list three categories of misleading information pertaining to: (i) the prospective or retrospective registration status of a trial; (ii) the issue of secondary IDs of a trial not being listed, resulting in 'hidden duplicates', and (iii) the results from a faulty search function.

Prospective or retrospective registration: Since 1 April 2018, it has been mandatory for trials registering with CTRI to register prospectively, that is, before the first participant is enrolled in the trial[12]. Each record states whether the trial was prospectively or retrospectively registered, but the information is sometimes misleading[13].

Secondary IDs: Another issue relates to 'hidden duplicates'. If a trial run in India is registered with a foreign registry, the ID issued by the latter registry should be entered in the CTRI record of the trial as a secondary ID, and vice versa. In a systematic review, a study registered in more than one registry but not cross-referenced in this way would be counted multiple times, leading to a bias in the literature. As such, although the absence of the secondary ID is a category of 'missing data', it deserves separate emphasis since it contributes to the larger problem of 'hidden duplicates'.

We identified 2908 trials in CT.gov that recruited from India but did not list a CTRI number[14]. The converse situation, where CTRI records do not list the CT.gov ID, is also likely to exist.

Faulty search function: I would like mention the faulty search function of CTRI specifically. Illustratively, a search for the number of trials a particular hospital ran yielded incorrect results[8]. A faulty search function is worse than an inoperable one since the result is misleading.

Above, I have listed 10 categories of errors. As mentioned, these could confuse in understanding an individual record or for a data analytics-based understanding of many records. Solutions must be put forth for the many challenges with data in these records, and I list some of them below.

Some proposed solutions

Earlier, we suggested several solutions to some of the problems listed above, such as making more fields compulsory, decreasing the number of free text fields, increasing the number of drop-down menus, and increasing the use of logic rules[8].

CTRI would benefit from the guidance of a highly skilled and experienced IT professional, who is likely to make additional suggestions. Also, if registrants were required to undergo mandatory training before they were permitted to register for their trials, this should help them generate more accurate records.

Some trials do not follow the rules, or may break the law

As mentioned above, several steps could be taken to improve the quality of data in CTRI. I next discuss two situations where CTRI records have not always complied with the relevant rules.

Registering regulatory trials with CTRI: Since 15 June 2009, it has been required that regulatory trials running in India register with CTRI[10]. However, it is possible that such trials have been registered with another registry, such as CT.gov, but not with CTRI. In a modelling study, we estimated that of 581 relevant trials registered with CT.gov, between 50 and 300 trials were not registered with CTRI, although the law requires it[14]. Although CTRI cannot ensure that such trials are registered, the Indian regulator, the Central Drugs Standard Control Organization (CDSCO), could do so.

Retrospectively registered trials: Since 1 April 2018, CTRI has required that trials register prospectively. We have found that all trials have not done so. Such trials have broken this rule[13].

Complying with the WHO requirements of primary registries

I now come to a third category of issues with trial registries, which is complying with the WHO requirements of primary registries. We have performed a comparative analysis of the registries that are data providers to ICTRP[4]. This analysis – based on various criteria that reflect the WHO recommendations through its International Standards of Clinical Trial Registries – ranked CTRI 11th out of the 18 registries. This reflects the significant improvement that CTRI could make with regard to: (i) information provided in each record; (ii) data download modalities, and (iii) details of the audit trail, etc. to become world class. Although on some issues, CTRI is technically compliant with the WHO requirements, other registries may have implemented a better system, and CTRI could aim to do likewise. In particular, I would like to highlight the importance of the *Results* field. CTRI does not have this field at present, but this would be a vital part of an excellent registry. Likewise, it does not have a *Data sharing plan* field. The WHO guidelines specify the need to have both of these fields[15].

Going beyond what is currently required by the WHO

In the early days, when trial registries were being conceptualized, there were discussions about what a record in a registry should look like[3]. Over the years, norms have been established about what data a trial registry should hold[15]. What additional steps must be taken, so that India has one of the best registries in the world? From the day it was established, CTRI has required trialists to provide details of the EC(s) when registering a study[16]. This was even before WHO recommended that registries make such information available. CTRI can build on this tradition and take certain steps – sometimes in collaboration with CDSCO – that will enable it to be a front-ranking registry. There are three issues that I discuss here.

Making other information available: In the interest of full transparency, the 'OpenTrials' initiative has sought to bring together all documentation on every trial on one platform[17]. Ideally, this should happen on an official platform such as a registry. As such, the protocol, consent form template, details of insurance for the participants, provisions for post-trial access, etc. could all be included in this 'futuristic' version of CTRI.

Working with CDSCO to enable public access to regulatory documents: Registry data are not peer-reviewed. Also, it is usually not possible to check them for authenticity against the raw data. It is known that there may be discrepancies in the data of a trial that is registered in more than one registry[7]. Discrepancies have also been noted in comparisons of trial data between (i) a registry and the subsequent publication(s)[18,19]; (ii) a registry and documents of the US Food and Drug Administration (FDA)[20,21], and (iii) a registry, the FDA documents and publications[22]. Further, it is known that there have been significant discrepancies in efficacy end-points that harm outcomes in the protocol which the EC has approved versus those reported in the publication[23]. Ideally, it should be possible to audit the registry data pertaining to a given trial through comparison with data on the same trial from other sources.

This should be partially possible if the data are in a publicly available regulatory document. Notably, in the US, documents pertaining to approved drugs – the Medical Review, Multidiscipline Review or Clinical Review – are often publicly available. These documents provide details of the clinical trials that were the basis for the drug's approval. However, CDSCO does not make such documents public. To enable such audits, India should make similar regulatory documents of all approved drugs and devices readily available to the public.

Working with CDSCO to ensure good CTRI records of regulatory trials: In response to our publication about some deficiencies in the CTRI records[8], CTRI staff argued that they are somewhat helpless in sub-standard records if the trialists do not cooperate[24]. CDSCO could refrain from clearing new applications until the CTRI records are correct, at least for regulatory trials. This approach could be used even to correct the older records whose sponsors are seeking fresh approvals. CTRI could seek to collaborate with CDSCO to achieve this goal.

Improving organizational functioning of CTRI

Finally, CTRI is a non-permanent activity of the Indian Council of Medical Research, and its staff have been 'temporary' for up to 15 years[25]. This led to the staff going to court in 2022. For CTRI to be a world-class registry, it needs to be consolidated as a permanent activity with re-energized staff who have, say, five-year contracts.

Conclusion

I have summarized some of the challenges with CTRI and how they could be handled. Given the many uses of trial registry data, it behooves us to ensure that registries capture high-quality data related to every trial that is mandated to be registered. If these suggestions are acted upon, CTRI would fulfil the goals of transparency and accountability and be amongst the best registries in the world. Further, the issues discussed above are unlikely to affect only the

records of CTRI. All 18 registries recognized by WHO would benefit from the proposed corrective or new measures in case they do not already have them in place.

Competing interests: G.S. has acted as a consultant to Dinesh Thakur on the various ways in which the workings of CTRI could be improved. This article is partially an outcome of that work.

1. Dickersin, K. and Rennie, D., Registering clinical trials. *JAMA*, 2003, **290**, 516–523.
2. Saberwal, G., The many uses of data in public clinical trial registries. *Curr. Sci*., 2021, **120**, 1686–1691.
3. Krleža-Jerić, K., Clinical trial registration: the differing views of industry, the WHO, and the Ottawa Group. *PLoS Med*., 2005, **2**, e378.
4. Venugopal, N. and Saberwal, G., A comparative analysis of important public clinical trial registries, and a proposal for an interim ideal one. *PLoS ONE*, 2021, **16**, e0251191.
5. Zarin, D. A., Tse, T., Williams, R. J. and Rajakannan, T., Update on trial registration 11 years after the ICMJE policy was established. *N. Engl. J. Med*., 2017, **376**, 383–391.
6. Zarin, D. A., Tse, T., Williams, R. J., Califf, R. M. and Ide, N. C., The ClinicalTrials.gov results database – update and key issues. *N. Engl. J. Med*., 2011, **364**, 852–860.
7. Fleminger, J. and Goldacre, B., Prevalence of clinical trial status discrepancies: a cross-sectional study of 10,492 trials registered on both ClinicalTrials.gov and the European Union Clinical Trials Register. *PLoS ONE*, 2018, **13**, e0193088.
8. Pillamarapu, M., Mohan, A. and Saberwal, G., An analysis of deficiencies in the data of interventional drug trials registered with Clinical Trials Registry-India. *Trials*, 2019, **20**, 535.
9. Nundy, S. and Gulhati, C. M., A new colonialism? – Conducting clinical trials in India. *N. Engl. J. Med.*, 2005, **352**, 1633–1636.
10. Pandey, A., Aggarwal, A., Maulik, M., Gupta, J. and Juneja, A., Challenges in administering a clinical trials registry: lessons from the Clinical Trials Registry-India. *Pharmaceut. Med.*, 2013, **27**, 83–93.
11. Chakraborty, I., Shreya, A., Mendiratta, J., Bhan, A. and Saberwal, G., An analysis of deficiencies in the ethics committee data of certain interventional trials registered with the Clinical Trials Registry-India. *PLoS Global Public Health*, 2022, **2**, e0000617.
12. http://ctri.nic.in/Clinicaltrials/alert.php (accessed on 2 March 2023).
13. Chakraborty, I. and Saberwal, G., CTRI has required prospective registration since 1 April 2018 – this has not always happened. *Indian J. Med. Ethics.*, 2022, **7**, 312–314.
14. Kumari, S., Mohan, A. and Saberwal, G., Hidden duplicates: 10s or 100s of Indian trials, registered with ClinicalTrials.gov, have not been registered in India, as required by law. *PLoS ONE*, 2020, **15**, e0234925.
15. WHO, International standards for clinical trial registries: the registration of all interventional trials is a scientific, ethical and moral responsibility. World Health Organization, Geneva, Switzerland, 2018.
16. Tharyan, P., Prospective registration of clinical trials in India: strategies, achievements and challenges. *J. Evid. Based Med.*, 2009, **2**, 19–28.
17. Goldacre, B. and Gray, J., OpenTrials: towards a collaborative open database of all available information on all clinical trials. *Trials*, 2016, **17**, 164.
18. Lv, J.-W. *et al*., Do published data in trials assessing cancer drugs reflect the real picture of efficacy and safety? *J. Natl. Compr. Cancer Netw*., 2017, **15**, 1363–1371.
19. Talebi, R., Redberg, R. F. and Ross, J. S., Consistency of trial reporting between ClinicalTrials.gov and corresponding publications: one decade after FDAAA. *Trials*, 2020, **21**, 675.
20. Zarin, D. A., Fain, K. M., Dobbins, H. D., Tse, T. and Williams, R. J., 10-Year update on study results submitted to ClinicalTrials.gov. *N. Engl. J. Med.*, 2019, **381**, 1966–1974.
21. Choudhury, M. C., Chakraborty, I. and Saberwal, G., Discrepancies between FDA documents and ClinicalTrials.gov for orphan drug-related clinical trial data. *PLoS Global Public Health*, 2022, **2**, e0000261.
22. Pradhan, R. and Singh, S., Comparison of data on serious adverse events and mortality in ClinicalTrials.gov, corresponding journal articles, and FDA medical reviews: cross-sectional analysis. *Drug Saf.*, 2018, **41**, 849–857.
23. Chan, A.-W., Hróbjartsson, A., Haahr, M. T., Gøtzsche, P. C. and Altman, D. G., Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*, 2004, **291**, 2457–2465.
24. Maulik, M. *et al*., Letter on: An analysis of deficiencies in the data of interventional drug trials registered with Clinical Trials Registry-India. *Trials*, 2020, **21**, 38.
25. Mudur, G. S., Indian Council of Medical Research public database not 'permanent'. *The Telegraph*, online, 19 June 2022; https://www.telegraphindia.com/india/indian-council-of-medical-research-public-database-not-permanent/cid/1870722 (accessed on 2 March 2023).

ACKNOWLEDGEMENTS. G.S. thanks Mounika Pillamarapu, Jaishree Mendiratta and Anwesha Dhal (all while working at Institute of Bioinformatics and Applied Biotechnology) for technical assistance. This work was also partially enabled by support from the Department of Electronics, Information Technology, Biotechnology and Science and Technology, Government of Karnataka.

Received 8 December 2022; revised accepted 10 February 2023

doi: 10.18520/cs/v124/i7/785-789