BMC Medical Research Methodology

RESEARCH

Open Access

Identifying trials run in India that are registered in other clinical trial registries: a cross-sectional study

Rishima Borah[1], Anwesha Dhal Samanta[1], Jaishree Mendiratta[1], Manish Mishra[2] and Gayatri Saberwal[1*]

Abstract

Background Clinical trials play a crucial role in biomedical research, and it is important to register them in public registries to ensure transparency and prevent research waste. In this study, we wished to determine what steps need to be taken to identify every clinical trial run in India that has been registered in any of the (non-Indian) World Health Organization-recognised primary registries. Of the 16 registries, we studied all except that of the European Union, which will be studied separately.

Methods Two methodologies were employed for each registry, except for four that did not facilitate one or the other method. Methodology A involved downloading all the records in a registry and querying them. Methodology B involved conducting a search via the registry website.

Results Only four registries provided consistent results with both methodologies. Seven registries had different results from the two methodologies. Of these, in four cases, in Methodology A one field indicated that the study ran in India, while another indicated otherwise.

Conclusions The above-mentioned ambiguities should be addressed by the concerned registries. Overall, this study reinforces the need for improved data accuracy and transparency in clinical trial registries and emphasizes the importance of resolving complications faced by users while navigating the registries. Ensuring accurate and comprehensive registration of clinical trials is essential for meta-research and the use of such data by a variety of stakeholders.

Keywords Clinical trial, Lacunae in trial registries, India, Data integrity, Trial registry-metaresearch, Research waste

Background

Clinical trials are a crucial part of translational biomedical research and trial registries have become a crucial part of the global biomedical research infrastructure that enables various stakeholders to monitor registered studies on various counts. The World Health Organization (WHO) currently recognises 17 registries as primary registries that provide data to its International Clinical Trials Registry Platform, through which these registries can be searched [1].

If a trial is registered prospectively, that is, before the enrollment of the first participant, with complete and accurate information in the record, then (a) the study's existence cannot be hidden, and (b) the pre-specified outcomes should be accounted for when the results are announced. The date by which trial results ought to be declared will also be known. Further, pre-registration has been shown to protect against p-hacking (multiple, repeated analyses in search of a significant result) and selective reporting [2]. As such, the timely, correct, and complete recording of study details such as protocols,

*Correspondence:
Gayatri Saberwal
gayatri@ibab.ac.in
[1] Institute of Bioinformatics and Applied Biotechnology, Bengaluru, Karnataka, India
[2] Aganitha Cognitive Solutions, Hyderabad, Telangana, India

including data analysis plans and results, in a public registry contributes to the robustness of the registry and therefore to the robustness of research based on registry data.

Unfortunately, there are many lacunae in the holdings of trial registries. Known problems include studies that are not registered [3] or have false or incomplete data in a given record [4], trials for which results are not reported at all or not reported on time [5], discrepancies in the details of a given study that is registered in more than one registry [6], and hidden duplicates, which is the phenomenon of a given study that is registered in more than one registry without suitable identification as the same study [7, 8].

We have a specific interest in the records of trials run in India, i.e., where participants who were based in India were recruited to the study. In this work, we wished to determine what steps need to be taken to identify every study run in India that has been registered in any of the (non-Indian) WHO-recognised registries mentioned above. This work would enable various other studies, such as (a) whether any trial required to be registered in India had not been, although it was registered elsewhere; (b) whether Indian trials that were required to be prospectively registered, were registered elsewhere, but not in India, and were retrospectively registered; (c) conducting an audit of Indian trials registered in multiple registries, to check for consistency of information in the multiple records. To the best of our knowledge, the work reported here has never been conducted before, for any country's trials.

Since the EU Clinical Trials Register (EUCTR), the registry of the European Union, is a large registry with particular complexities, we have divided this work into two parts and will cover EUCTR separately. In this, the first part, we attempted to study the 15 registries aside from Clinical Trials Registry-India (CTRI) and EUCTR.

Methods

The 15 WHO-recognized registries of this study are as follows: (i) ANZCTR: Australian New Zealand Clinical Trials Registry; (ii) ChiCTR: Chinese Clinical Trial Registry; (iii) CRIS: Clinical Research Information Service of Republic of Korea; (iv) DRKS: German Clinical Trials Register; (v) IRCT: Iranian Registry of Clinical Trials; (vi) ISRCTN: ISRCTN is not an acronym any more; (vii) TMCTR: International Traditional Medicine Clinical Trial Registry; (viii) JRCT: Japan Registry of Clinical Trials; (ix) LBCTR: Lebanese Clinical Trials Registry; (x) PACTR: Pan African Clinical Trials Registry; (xi) ReBEC: Brazilian Registry of Clinical Trials; (xii) REPEC: Peruvian Clinical Trials Registry; (xiii) RPCEC: Cuban Public Registry of Clinical Trials; (iv) SLCTR: Sri Lanka Clinical Trials Registry; and (xv) TCTR: Thai Clinical Trials Registry.

For each of these registries, we adopted two methodologies, where feasible. Each step of each methodology, for each registry, was performed by two authors independently. Methodology A consisted of downloading all the records in the registry. The records were web-scraped, and the information was structured in the form of tables. We queried these tables for 'India' and 'CTRI', since studies registered with the Indian registry have registration numbers starting with CTRI, and a foreign registry might cross-reference a trial registered in India using its CTRI number. We stored the outcomes of all of the successive steps in the processing of these records as tables in a local SQLite database. Methodology B involved conducting a search via the registry website.

In each of the cases where Methodologies A and B yielded different results, we wrote to the registry (the emails are available in Additional file 2), with a reminder two weeks later if necessary, bringing the details to their attention or seeking to understand the reason for the discrepancy in the result between the two methods. Some of the registries responded. Where relevant, we also wrote to the concerned registries, pointing out that conducting searches by Methodology A or B was not possible.

For each registry, we created a folder that contained the following: First, an SQLite database containing all the information about the downloaded records in the form of structured tables (an sqlite file). Second, a zipped folder containing (i) the scripts used to extract data (a txt file), (ii) the Methodology, Summary results, and Flowchart (a pdf file), wherein the Flowchart outlined the steps involved in identifying the trials that had run in India, and (iii) the Detailed results (an xls file). For each registry, all these files are available as supplementary material at the online data repository osf.io as Additional file 1 [9].

Methodological complications were encountered with certain registries, as follows. In each case, further details are available in the supplementary file that describes the methodology followed for that particular registry.

a) ChiCTR: A captcha code on the website prevented web crawlers from downloading all the records.
b) CRIS: There were three complications here. First, records could not be downloaded in the same way as for most of the other registries because the different links, which corresponded to various records, redirected us to the same page. Second, each record had multiple versions, which varied in their details, including the number of sites. Third, the country where each site was located was not listed. We had to manually examine the list of trial sites in order to assess which of them were located in India.

c) jRCT: This registry has a Japanese version and an official English version. However, one can easily translate the Japanese version into an unofficial English version. The official and unofficial English versions of a record may contain slightly different information. We used the official English version of the records.
d) PACTR: If we wished to download all the records, as we had done for most of the other registries, we would have had to extract all the records' URLs as listed on the website, and then we would have downloaded those records. However, in this registry, the successive webpages, that listed several studies each, had the same URL. Therefore, we had to find a different way of downloading all the trials, as we had done for CRIS.
e) ReBEC: The number of 'results found' did not correspond to the number of displayed records.
f) REPEC: While most records were accessible in English, some had not yet been translated from Spanish. Different protocols were necessary to access the records in English and those in Spanish.

Methodology B consisted of a search of the registry website using 'India' as a keyword. Preferably, this was an 'Advanced search', searching the '*Country of recruitment*' (or equivalent) field. But if that option was unavailable, then we performed a simple search. For jRCT and ReBEC, we could only use a simple search function. Not all registries enabled Methodology B, since some did not even have a simple search function. We were unable to use Methodology B for LBCTR, REPEC, and SLCTR since there were no suitable search options on the registry website. Also, we were unable to study TCTR by Methodology A because there was a common URL for all the result pages of the search for relevant TCTR records.

Finally, we wished to better understand why so many Indian studies were registered with other registries, and looked into those registered with ISRCTN.

Results

Table 1 summarizes the results obtained by Methodology A and Methodology B (where feasible).

To be noted, using Methodology A, although usually it was the *Countries of recruitment* field that indicated that

Table 1 For each of the 15 registries that were part of this study, the number of trials that had run in India, as identified by Methodology A or B

	Registry name or acronym	Methodology A (web scraping)			Methodology B (registry search function)
		Methodology A, overall	From the Countries of recruitment field	From other fields	
I. Registries for which both Methodology A and Methodology B could be used					
(a) Those for which Methodology A and Methodology B yielded the same result					
1.	DRKS	30	30	0	30
2.	ISRCTN	358	358	0	358
3.	ITMCTR	0	0	0	0
4.	RPCEC	0	0	0	0
(b) Those for which Methodology A and Methodology B did not yield the same result					
5.	ANZCTR	123	121	2	121
6.	ChiCTR	17	16	1	16
7.	CRIS	2	2 (from *Study site*)	0	0
8.	IRCT	1	0	1	2
9.	jRCT	111	111	0	78
10.	PACTR	7	4	3	4
11.	ReBEC	13	13	0	1
II. Registries for which only Methodology A could be used					
(a) Those for which the field *Countries of recruitment* yielded all the trials that had run in India					
12.	LBCTR	23	23	0	Not feasible
(b) Those for which the field *Countries of recruitment* did not yield all the trials that had run in India					
13.	REPEC	346	345	1	Not feasible
14.	SLCTR	29	27	2	Not feasible
III Registry for which only Methodology B could be used					
15.	TCTR	Not feasible	-	-	6

a trial had run in India, sometimes this information was only available from other fields.

The registries that gave us perfect results with the same results by Methodologies A and B were DRKS, ISRCTN, ITMCTR, and RPCEC. LBCTR did not enable Methodology B. However, in this registry, all the trials identified through Methodology A were from the field *Country of recruitment*, and therefore this was a good set of records as well.

Figure 1 summarizes the results for registries where both Methodology A and Methodology B were feasible.

We identified the studies (listed in Additional file 3) where some of the trials that had run in India were identified by fields other than *Countries of recruitment*. These trials were in the registries ANZCTR, ChiCTR, IRCT, and PACTR from Table 1-Ib and REPEC and SLCTR from Table 1-IIb. The various issues (detailed in Additional file 3) were as follows: The *Country of recruitment* did not list India, did not list any country, or did not exist at all. Nevertheless, the *Brief summary* (2 trials) the *Public title* (1 trial), *Official scientific title* (1 trial) or the *Inclusion Criteria* (1 trial) indicated that the trial had run in India; both the Applicant's and the Ethics committee's addresses were in India (1 trial); or a CTRI ID was listed as a Secondary ID (3 trials) indicating that the trial had most likely run in India.

Aside from these discrepancies, we also found certain idiosyncrasies with particular registries that we list in the 'Methodology, Summary results, and Flowchart' file for each registry.

The maximum number of records (358) were hosted by ISRCTN. In order to better understand why so many Indian studies were registered with other registries, as exemplified by ISRCTN, we examined these records. The results are available in Additional file 4. In brief, of the 358 trials, 169 ran in other countries as well. Of the 189 that had run only in India, 43 were registered before it became compulsory to register with CTRI, and 70 of the rest had foreign funders or sponsors. Of the remaining 76 cases, 4 were registered prospectively, and 72 retrospectively.

Finally, we note that some of the registries host far more Indian trials than others. ISRCTN hosts almost three times the number hosted by ANZCTR, and only ISRCTN, ANZCTR and jRCT host trials in the triple digits. DRKS, ChiCTR and ReBEC records are in the double digits, PACTR, IRCT and CRIS in the single digits and ITMCTR and RPCEC have zero holdings.

Discussion

As mentioned above, we had planned to study 15 of the registries that WHO recognizes. From earlier work [10], we knew that a given registry may not have an advanced search option, or even a suitable basic one. We therefore decided to analyse the records after downloading all the records in each registry. However, on cross-checking our

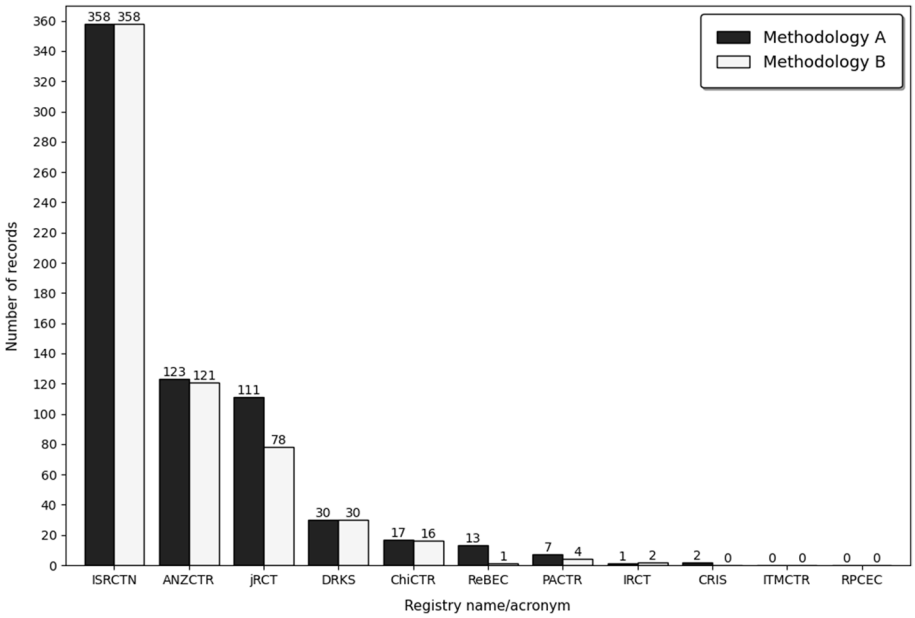

Fig. 1 The results obtained by Methodology A (web scraping) and Methodology B (registry search function) from registries where both were feasible

results with searches of the registry website, we noticed discrepancies in some registries. Therefore, we systematically used both methodologies (where possible) for each registry, and quantified and detailed the discrepancies in the results.

In some of the trials, where *Country of recruitment* did not list India, but other fields indicated that the study had run in India, it is possible that India was a planned location, but was subsequently dropped. However, unless the data in these other fields is reconciled after any changes, or there is a clarificatory note such as in *Brief Summary*, it is unclear whether or not the study ran in India. Although ChiCTR does have a *Recruitment country* field, this was blank in the record of interest, although other fields indicated that the trial had run in India. This speaks to the issue of 'incomplete records' that have been noted in registries earlier [4, 11]. IRCT had no *Recruitment country* field, and therefore the keyword 'India' was not present in a record although this record was picked up by Methodology B. Presumably 'India' was present in the metadata. But the results ought to be the same irrespective of the methodology followed. Since most registries do have *Country of recruitment*, it would be best if IRCT, too, has this field.

In three cases, it was only the CTRI numbers that identified them as registered with the Indian registry. This was confirmed on the CTRI website. However, it should not be necessary to check the CTRI record to try and understand a record in any other registry.

Coming to other records, as mentioned in the 'Methodology, Summary results, and Flowchart' file pertaining to LBCTR, this registry had the unusual situation of one record being deleted over a few months. The WHO recommends that once a record is entered in the database, it should not be deleted except in exceptional circumstances [12]. As such, this record should probably not have been deleted. In another LBCTR record, all the countries of recruitment were missing from a record that had listed them earlier. WHO recommends a publicly available 'audit trail' for each record so that researchers can understand the history of a given record, but the *Changes History* field of this trial did not list this change.

Methodologically speaking, ChiCTR, CRIS, jRCT, PACTR, ReBEC, and REPEC had twists, as mentioned above. It would probably be best if their anomalies were smoothed out, since it makes navigating their records more complicated. As noted, a problem with the ReBEC searches was that the number of 'results' corresponded to the number of occurrences of 'India'. Therefore, on a simple search for 'India', the website stated that there were four results but displayed only three records because one record contained'India' twice. Subsequent to our reaching out to ReBEC, this problem has been rectified.

In most, if not all, cases, all the data in the records must be whatever the sponsor or responsible party submitted. For instance, the managers of SLCTR mentioned that "SLCTR publishes the Trial Registration Data Set submitted by the Principal Investigator after review. However, each data field is not independently verified prior to registration." Although there may be ways to further improve the way registry staff cross-check records before accepting them, the onus of having correct records primarily lies with the registrant. Johns Hopkins and other universities have created dedicated teams to help registrants. The team at Johns Hopkins University reports that they brought down their potentially problematic trial records from 44 to 2% over a 5-year period [13].

We now come to the analysis of the trials registered with ISRCTN. We need not be concerned about studies that were registered with ISRCTN before it became compulsory to register with CTRI. For trials that were run in other countries or where the sponsors or funders were foreign organizations, there may have been other compulsions to register with a foreign registry. However, there are two registration dates that we must keep in mind: (i) Since 1 April 2018, trials must register with CTRI prospectively. It is understandable that if a study could no longer be registered with CTRI because it was not registered prospectively by this date, it would register with another primary registry [14]. This seems to have been the case for most of the trials for which there was no other obvious explanation. (ii) Since March 2019, it has been a legal requirement that regulatory trials run in India be registered with CTRI [15]. Therefore, in future, all the studies registered with ISRCTN or any other registry should be examined to see whether they have broken Indian law. In earlier work in which we assessed the records of the United States' registry, ClinicalTrials.gov, we showed that there do seem to be such cases [8].

Finally, we note that some of the registries host records for far more trials that have run in India than others. This information should be of interest to researchers or other users of registries.

In summary, several registries had idiosyncrasies that needed to be worked through before a comprehensive assessment of their Indian trials could be undertaken. For the four registries where only one methodology could be performed, it is not known how robust the findings are. In general, because a large fraction of studies are not published [16]], meta-researchers are being increasingly urged to search trial registries in addition to publications for relevant data. However, registry managers [17] and other researchers [18] have found that such searches need to be customised for particular registries. In order to facilitate users, registries should take steps to resolve such complications if possible.

Only four registries provided the same result using both methods, and six registries had discrepancies within a given record, wherein one field indicated that the trial had run in India and another field indicated that it had not. These registries should find ways to enable unambiguous records. Either changes to the records need to be reflected throughout the record, or clarification provided. Records also need to be up to date and complete. All registries should aim to have fields that are prescribed by WHO [12] and that are common to most registries. Registries should also follow WHO guidelines on not deleting information from records (unless there is an audit trail) or the record itself. Finally, it is important that institutions provide assistance to their registrants to file accurate records. Although the number of discrepant records was small, it is important to be aware of the phenomenon, and to take it into account for future trial-related meta-research.

Limitations

In this study, we aimed to study 15 registries using two methods each. We were unable to study one registry using Methodology A and three using Methodology B. As such, we may have missed some of the complexities in the data in these registries. Further, we only studied one issue, that of whether or not the data indicated that the trial had run in India. We cannot extrapolate these insights to other issues.

Conclusions

We wished to determine what steps need to be taken to identify every clinical study run in India that has been registered in 15 (non-Indian) registries recognised as primary registries by WHO. Only in four registries did we obtain the same result when we used two methods to examine the data per registry. In the other registries, either one method could not be used, or there were discrepant results using the two methods. In the interest of accuracy and transparency, each registry must develop methods to prevent such ambiguity. These improvements will be of great help to researchers conducting meta-research or using these registries in other ways.

Abbreviations

ANZCTR	Australian New Zealand Clinical Trials Registry
ChiCTR	Chinese Clinical Trial Registry
CRIS	Clinical Research Information Service of Republic of Korea
CTRI	Clinical Trials Registry – India
DRKS	German Clinical Trials Register
EUCTR	European Union Clinical Trials Register
IRCT	Iranian Registry of Clinical Trials
ITMCTR	International Traditional Medicine Clinical Trial Registry
JRCT	Japan Registry of Clinical Trials
LBCTR	Lebanese Clinical Trials Registry
PACTR	Pan African Clinical Trials Registry
ReBEC	Brazilian Registry of Clinical Trials
REPEC	Peruvian Clinical Trials Registry
RPCEC	Cuban Public Registry of Clinical Trials
SLCTR	Sri Lanka Clinical Trials Registry
TCTR	Thai Clinical Trials Registry
WHO	World Health Organization

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12874-024-02336-w.

Additional file 1. For each registry, where feasible, the SQLite database, scripts, methodology, summary results, flowchart, and detailed results.

Additional file 2. Correspondence with nine registries.

Additional file 3. The trials that had run in India that were identified by fields other than *Countries of recruitment*.

Additional file 4. Analyzing the records of Indian trials registered with ISRCTN.

Acknowledgements

We thank Ms. Claire Veryard, of ISRCTN, for discussion.

Authors' contributions

GS conceptualized the study, acquired funding, helped develop the methodology, conducted the investigation and analysis, validated the results, coordinated and supervized the project, and wrote the original draft of the manuscript. RB and ADS helped develop the methodology, conducted the investigation and analysis, and validated the results. JM and MM helped develop the methodology, and conducted the investigation and analysis. All authors contributed to the final manuscript and approved it.

Funding

This work was largely supported by the Department of Electronics, Information Technology, Biotechnology, and Science & Technology of the Government of Karnataka, India. The funder had no specific role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

The dataset(s) supporting the conclusions of this article are available in the OSF repository repository, [DOI: https://doi.org/10.17605/OSF.IO/4VDP3, https://osf.io/4vdp3] or are included within the article and its supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Manish Mishra works for Aganitha (https://www.aganitha.ai/solutions/knowledge-hubs/), which engineers and markets Knowledge Hubs for research and clinical data. Jaishree Mendiratta currently works for Novotech (https://novotech-cro.com/) that provides clinical development services. However, neither company has a commercial interest in the outcomes of this study. The other authors declare that they have no competing interests.

Received: 31 August 2023 Accepted: 6 September 2024
Published online: 12 September 2024

References

1. World Health Organization. International Clinical Trials Registry Platform. ICRTP Registry Network. Primary registries. Available from: https://www.who.int/clinical-trials-registry-platform/network/primary-registries. Cited 3 Jul 2023.
2. Decker C, Ottaviani M. Preregistration and credibility of clinical trials. medRxiv. 2023;2023.05.22.23290326. https://doi.org/10.1101/2023.05.22.23290326. Cited 30 Aug 2023.
3. Denneny C, Bourne S, Kolstoe SE. Registration audit of clinical trials given a favourable opinion by UK research ethics committees. BMJ Open. 2019;9(2):e026840.
4. Pillamarapu M, Mohan A, Saberwal G. An analysis of deficiencies in the data of interventional drug trials registered with clinical trials Registry - India. Trials. 2019;20(1):409.
5. Anonymous. "No more excuses" as major global research funders take strong lead on clinical trial transparency. In: All trials. 2017. Available from: https://www.alltrials.net/news/funders-agree-to-who-standards/. Cited 9 Aug 2023.
6. 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. JAMA Netw Open. 2021;4(6):e2128898.
7. van Valkenhoef G, Loane RF, Zarin DA. Previously unidentified duplicate registrations of clinical trials: an exploratory analysis of registry data worldwide. Syst Rev. 2016;5(1):119.
8. Kumari S, Mohan A, 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(6):e0234925.
9. Borah R, Samanta AD, Mendiratta J, Mishra M, Saberwal G. Trials in India registered in other registries. OSF Database. 2023. https://doi.org/10.17605/OSF.IO/4VDP3.
10. Venugopal N, Saberwal G. A comparative analysis of important public clinical trial registries, and a proposal for an interim ideal one. PLoS One. 2021;16(6):e0251191.
11. DeVito NJ, Morley J, Smith JA, Drysdale H, Goldacre B, Heneghan C. Availability of results of clinical trials registered on EU Clinical Trials Register: cross sectional audit study. BMJ Med. 2024;3(1):e000738. Available from: https://bmjmedicine.bmj.com/content/3/1/e000738.
12. International standards for clinical trial registries v 3.0. 2021. Available from: https://apps.who.int/iris/bitstream/handle/10665/76705/9789241504294_eng.pdf?sequence=1&isAllowed=y. Cited 26 Jan 2021.
13. Keyes A, Mayo-Wilson E, Nuamah P, Lalji A, Tetteh O, Ford DE. Creating a program to support registering and reporting clinical trials at Johns Hopkins University. Acad Med. 2021;96(8):529–33.
14. Chakraborty I, Saberwal G. CTRI requirement of prospective trial registration: not always consistent. Indian J Med Ethics. 2022;VII(4):312–4.
15. Ministry of Health and Family Welfare. Govt of India. Notification. The Gazette of India: Extraordinary, Part II, Section 3, Subsection (i), New Delhi. 2019. p. 1-264. Ministry of Health GSR Notification #227; 2019. Available from: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/NewDrugs_CTRules_2019.pdf.
16. Schmucker C, Schell LK, Portalupi S, Oeller P, Cabrera L, Bassler D, et al. Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. PLoS One. 2014;9(12):e114023.
17. Hunter KE, Webster AC, Page MJ, Willson M, McDonald S, Berber S, et al. Searching clinical trials registers: guide for systematic reviewers. BMJ. 2022;377:e068791.
18. Alqaidoom Z, Nguyen P-Y, Awadh M, Page M. Impact of searching clinical trials registers in systematic reviews of pharmaceutical and non-pharmaceutical interventions: reanalysis of meta-analyses. Res Synth Methods. 2023;14(1):52–67.

Publisher's Note