An examination of the use of clinical trials as a source of information in scientific research

Um olhar sobre o uso de estudos clínicos como fonte de informação nas pesquisas científicas

Bárbara Adriana Ferreira dos Santos TRIGUEIROS[1] 0000-0003-1950-3479

Andréa Rodrigues ÁVILA[1] 0000-0001-5221-261X

Fabricia Pires PIMENTA[1, 2] 0000-0002-2315-1974

Abstract

As part of the innovation process, clinical research generates valuable data to assess technological solutions in healthcare, linking scientific research and knowledge transfer to provide beneficial innovations for society. However, the generated clinical data do not appear to be adequately available to the scientific community and society. The present study seeks to analyze the effectiveness of databases of clinical records as a source of relevant information for scientific research. We conducted a comparative analysis of 27 correlations of clinical trials between three different technologies and their scientific articles by consulting two information sources. Making connections between the data from these sources proved to be challenging. In addition, a considerable time lag (40 months on average) was observed between the end of the study and the publication of the results. Among the completed studies, 56% had not published their results in any of the channels studied. In addition to the paucity of reported results, the poor publication record of clinical trials is further evident in the lack of information on these studies in scientific publications. Thus, databases of clinical records are a potential source of information and may come to represent a central tool in the search for new technological solutions in healthcare.

Keywords: Comparative Research. Technological Innovation. Metadata. Scientific Information.

Resumo

A pesquisa clínica, como parte de um processo de inovação, gera dados úteis para avaliação de diversas soluções tecnológicas em saúde, perfazendo um elo entre a pesquisa científica e a translação do conhecimento em inovações benéficas à sociedade. Entretanto, os dados clínicos gerados parecem não estar disponíveis adequadamente à comunidade científica e à sociedade. O presente estudo buscou analisar o potencial das bases de registros clínicos como fonte de informação relevante às pesquisas científicas e, para isso, foi realizada uma análise comparativa de 27 correlações de estudos clínicos de três diferentes tecnologias e seus artigos científicos em duas fontes de informação distintas. Nesse universo amostral, observou-se importante e limitante dificuldade em associar com exatidão os dados dessas fontes. Foi possível, ainda, observar um grande lapso temporal (40 meses em média) entre o encerramento de um estudo e a publicização dos resultados. Considerando os estudos finalizados, 56% não tiveram seus resultados publicados em nenhum dos canais estudados. A publicação dos estudos clínicos pode se mostrar ainda mais negligenciada quando se percebe que, além da ausência de resultados, não há qualquer informação desses estudos em publicações científicas. Assim, as bases de registros clínicos se apresentam como uma fonte de informação em potencial – e, por vezes, única – para suprir conhecimento na busca por novas soluções tecnológicas em saúde.

***Palavras-chave**: Pesquisa comparativa. Inovação tecnológica. Metadados. Informação científica.*

[1] Instituto Carlos Chagas, Fiocruz. Rua Prof. Algacyr Munhoz Mader, 3775, 81350-010, Curitiba, PR, Brasil. *Correspondência para*/Correspondence to: B.A.F.S. TRIGUEIROS. E-mail:< barbara.trigueiros@fiocruz.br>.

[2] Instituto Nacional de Ciência e Tecnologia de Inovação em Doenças de Populações Negligenciadas. Rio de Janeiro, RJ, Brasil.

Received in September 9, 2021 and approved on November 18, 2021.

Como citar este artigo/How to cite this article

Trigueiros, B. A. F. S.; Ávila, A. R.; Pimenta, F. P. An examination of the use of clinical trials as a source of information in scientific research. *Transinformação*, v. 34, e210065, 2022. https://doi.org/10.1590/2318-0889202234e210065

Introduction

The development of novel medicines, diagnostic methods, and treatments requires connecting various stages on the path from scientific research to commercialization (Ellwood; Williams; Egan, 2020). Specifically, connections must be made between basic research, technological development, Clinical Trials (CTs), and the commercialization of products.

Ideally, the large volume of scientific information generated at each stage should be communicated in scientific journals, patent documents, and CT registration platforms. However, as a social and political process, the transfer of information is not always homogeneous or equitable (Targino, 2001). Academic journals have been widely discussed as the vehicle of choice for communicating scientific discoveries. However, a variety of other channels are being used to transmit scientific information in spite of their lower profile within the scientific community (Pimenta, 2017).

As part of the innovation process, clinical research generates valuable data for assessing new treatments, diagnoses, drug effects, vaccines, diets, medical devices, and the detection of side effects. This research therefore bridges gaps between basic research and knowledge transfer, resulting in beneficial innovations for society (Ellwood; Williams; Egan, 2020).

According to the resolution RDC N° 39/2008 Anvisa-MS from the Drug Surveillance Agency of the Brazilian Ministry of Health, a clinical trial is:

> any research on humans involving therapeutic and diagnostic intervention that uses registered or registrable products to discover or verify their pharmacodynamic, pharmacokinetic, pharmacological, clinical and/or adverse side effects, in addition to determining their safety and/or efficacy (Brasil, 2008, online).

The main goals of clinical research include verifying the safety and efficacy of products and technologies (phase I), assessing their effectiveness and investigating their side effects (phase II), confirming their efficacy, and monitoring adverse reactions (phase III). The final phase (phase IV) involves commercialization and pharmacovigilance of the investigated product (Brasil, 2008). Hence, the benefits of clinical research go beyond innovations that are available to the public, since such research is the result of a complex and multidisciplinary process involving several agents like researchers, physicians, regulatory bodies, ethics committees, volunteers, and sponsors (Quental; Salles Filho, 2006).

In Brazil, all CTs must be evaluated by an institutional review board called the Research Ethics Committee (*Comitê de Ética em Pesquisa*; CEP) and, if necessary, by a second board, the National Commission for Research Ethics (*Comissão Nacional de Ética em Pesquisa*; CONEP) (Gouy; Porto; Penido, 2018). Clinical studies must also be registered in a database to avoid the duplication of trials and selective publishing of results, in addition to providing patients and the public with transparency and access to information (Tse; Fain; Zarin, 2018). Whether in phase I, II, III, or IV, CTs must be registered in the Brazilian Registry of Clinical Trials (*Registro Brasileiro de Ensaios Clínicos*; ReBEC) database in accordance with ANVISA resolution RDC 36 of June 27, 2012 (Brasil, 2012, online).

The registration of CTs begins with the submission of the experimental design, including the intervention plan in general terms. At the end of the study, the publication of results is considered extremely important for the scientific community and society. Traditionally, scientific journals have been the primary vehicles for communicating results, with a parallel role fulfilled by clinical trial databases. However, many CTs are never published (Shrivastava; Shrivastava; Ramasamy, 2018).

In this context, we seek to demonstrate the potential benefits of databases for CTs as a source of relevant information for developing technological solutions in healthcare.

Methodological Procedures

Selection and identification of technologies

Three technologies were selected for their relevance as potential solutions for public health issues and institutional patent landscape analyses (not yet published). Our study is based on published metadata on these three technologies. We used a scientometrics approach to understand the process of organizing scientific communication (Van Raan, 1997; Vanti, 2002).

The first technology is Eryaspase, a donor-derived red blood cell (RBC) encapsulation technology for the enzyme L-asparaginase. This technology aims to reduce the toxicity observed in treating Acute Lymphoblastic Leukemia (ALL), as well as increase the ability of other oncological treatments to reduce toxicity. Eryaspase is in phase III of CTs conducted by Erytech Pharma SA.

The second technology describes the active compound GSK3186899, similar to current anti-Leishmania drugs. This compound is in phase I of CTs, which were initiated in April 2019, conducted by GlaxoSmithKline plc.

The third technology is the compound DNDI 0690, which is being studied through a partnership between the TB Alliance and Drugs for Neglected Diseases (DNDi). This compound showed excellent in vitro activity against visceral and cutaneous leishmaniasis, proceeding to phase I CTs in March 2020.

Search strategy in clinical studies databases

Metadata from CTs were identified from the International Clinical Trials Registry Platform (ICTRP), an initiative of the World Health Organization (World Health Organization, 2018). The ICTRP platform was created to establish "a network of international clinical trials registries," ensuring a single point of access and unambiguous identification of CTs (https://www.who.int/clinical-trials-registry-platform/about).

Searches were carried out in May 2021 without time limitation, using the simple interface and entering keywords identified as relevant for each of the three technologies in the 'Search' field (Table 1).

Table 1 - Search strategies used to identify clinical trials and the respective case studies in the ICTRP database.

Number	Search strategy	Technology
I	asparaginase red blood cell	Eryaspase
II	asparaginase and erytec	Eryaspase
III	eryaspase	Eryaspase
Ia	(leishmaniasis or leishmania infection) and GSK3186899	GSK3186899
IIa	GSK3186899	GSK3186900
Ib	(leishmaniasis or leishmania infection) and DNDI 0690	DNDI-0690
IIb	DNDI 0690	DNDI-0690

Source: Prepared by the authors (2021).

After performing the searches, the metadata were collected from the fields "Main ID," "Secondary ID," "Public Title," "Scientific Title," "Sponsor," "Phase," "Contacts," "Affiliation," "Intervention," and "Results," which was then organized in an Excel table with duplicates manually removed. Additionally, when absent, some metadata content was supplemented with the results of searches in the primary databases of CT records from links of the ICTRP platform.

Identification of scientific articles related to clinical trials

As described in section 2, the data collected for each CT were used to search for related scientific articles (SAs) in Medline databases (via Pubmed), Capes Journal Portal, and Google Scholar. The SA metadata were organized into the categories "Title," "Authors," "Affiliation," "Disclosures," "Conflict of Interest," "Abstract," "Introduction," "Material and Methods," "Results," "Discussion," and "Conclusion" and grouped according to the respective CT. Numerical codes were used for CTs, while SAs were assigned an alphabetic code.

Comparative analysis between metadata from CTs and scientific articles

A comparative analysis was carried out of the CT and SA metadata from the three technological solutions using the criteria listed below. The CT metadata in the categories of "Main ID," "Secondary ID," "Phase," "Sponsor," "Intervention," "Contacts," and "Result" was considered to be present in the SA if they were contained in any part of the related article. The "Public Title" and "Scientific Title" from the CT metadata were considered to be present in the SA when two or more identical or technically similar keywords were present in the article's title. The "Affiliation" field in the CT metadata were considered to be present in the SA if the article's "Affiliation," "Disclosures," or "Conflict of interest" sections contained matching information.

Analysis of the state of CTs and publication of results

Initially, the metadata on the completion date of the CTs in the primary databases was used to determine studies that had been completed and those that were still in progress.

We used the following fields from CT databases to identify the publication of results: "Results," "Posted Results," "Trial Results," and "Study Information." In the SA databases, we searched through the following fields to assess whether results had been published: "Publication Date," "Publication Type," "Abstract," "Materials and Methods," and "Results."

Finally, these data were used to verify the time elapsed between CT completion and the publication of results. A simple average was used when CT results were published more than once.

Results

Identification of CTs and SAs

After removing duplicates, the searches resulted in 17 records of Eryaspase CTs in the ICTRP database (Table 1). After the first search (I), the term "erytech" was observed in the "Primary Sponsor" field. Based on this information, a new search (II) was designed to increase search recall (I). The new search resulted in 12 CT records, nine of which were already present in the previous search.

An examination of the "Intervention" field during our analysis of search (II) uncovered three different names used for Eryaspase: "l-asparaginase encapsulated in RBC," "graspa," and "eryaspase." Thus, search (III) was designed, from which five more studies were added to the sample, for a total of 17 CT records for Eryaspase.

The searches for CTs for the compounds GSK3186899 and DNDI 0690 used a strategy based on combinations of keywords with concepts related to the disease and the compounds (searches Ia and Ib, respectively) or only the compounds (searches IIa and IIb) (Table 1). Both searches produced the same outcome. A single CT was identified for GSK3186899, while two were found for DNDI-0690.

Table 2 shows the results of the searches for SAs related to the identified CTs. CTs are represented by sequential numbers and SAs by sequential letters, with a corresponding alphanumeric code denoting each SA-CT combination.

A total of 27 SAs were found for Eryaspase, while no SAs were found describing the results of CTs on GSK3186899 and DNDI-0690.

Table 2 - Correlation between clinical trials and the corresponding SAs for Eryaspase, GSK3186899, and DNDI 0690 technologies.

1 de 5

	N° CT	(Main ID / Secondary ID	Primary registration base	Public title	Search	Title(s) of the identified scientific article(s)	Letter AC	N° of CT and letter of SA
Eryaspase	**1**	NCT00723346/ GRASPALL 20S05-01	ClinicalTrials.gov	Administration of Allogenic Red Blood Cells Loaded L-asparaginase in Cases of Relapse of Acute Lymphoblastic Leukaemia GRASPALL.	I e II	Erythrocyte encapsulated l-asparaginase (GRASPA) in acute leukemia (doi.org/10.2217/ijh-2016-0002)	a	(1a)
						L-asparaginase loaded red blood cells in…GRASPALL 2005-01 randomized trial (doi.org/10.1111/j.1365-2141.2011.08588.x).	b	(1b)
						GRASPALL 2005.01 Clinical Study: L-Asparaginase Loaded into Red Blood Cells…Relapsed Acute Lymphoblastic Leukaemia (ALL) (doi.org/10.1182/blood.V112.11.306.306)	c	(1c)
						L-Asparaginase Loaded Inside Red Cells Has An Acceptable Tolerability Profile On Bilirubin Value (doi.org/10.1182/blood.V122.21.2642.2642).	d	(1d)
	2	EUCTR2009-012584-34-BE/ GRASPALL 2009-06	EUCTR	Clinical trial with GRASPA, Red Blood cells encapsulating L-Asparaginase, in patients affected by Acute Lymphoblastic leukemia at relapse.	I e II	Erythrocyte encapsulated l-asparaginase (GRASPA) in acute leukemia (doi.org/10.2217/ijh-2016-0002).	a	(2a)
	3	NCT01523808/ GRASPANC 2008-02	ClinicalTrials.gov	Administration of GRASPA (Suspension of Erythrocytes Encapsulating L-asparaginase) in Patients With Pancreatic Cancer.	I e II	Asparagine Synthetase Expression and Phase I Study With L-Asparaginase Encapsulated in Red Blood Cells in Patients With Pancreatic Adenocarcinoma (doi.org/10.1097/MPA.0000000000000394).	e	(3e)
	4	NCT01810705/ GRASPA-AML-2012-01	ClinicalTrials.gov	GRASPA Treatment for Patients With Acute Myeloblastic Leukemia ENFORCE.	I e II	Erythrocyte encapsulated l-asparaginase (GRASPA) in acute leukemia (doi.org/10.2217/ijh-2016-0002).	a	(4a)
						GRASPA-AML 2012-01 study (NCT01810705): A multicenter,open, randomized phase 2b trialevaluating ERY001… for intensive chemotherapy (doi.org/10.1200/jco.2015.33.15_suppl.tps709)	f	(4f)

Table 2 - Correlation between clinical trials and the corresponding SAs for Eryaspase, GSK3186899, and DNDI 0690 technologies.

	N° CT	(Main ID / Secondary ID	Primary registration base	Public title	Search	Title(s) of the identified scientific article(s)	Letter AC	N° of CT and letter of SA
Eryaspase	**5**	NCT01910428/ GRASPALL 2012-09	ClinicalTrials.gov	L-asparaginase Encapsulated in Red Blood Cells (Eryaspase) for Treatment of Adult Patients With ALL or LBL	I e II	Not found		
	6	EUCTR2012-002026-78-FI; GRASPA-AML-2012-01	EUCTR	Clinical trial with GRASPA, Red Blood cells encapsulating L-Asparaginase, in patients affected by Acute Myeloid leukemia	I e II	Erythrocyte encapsulated l-asparaginase (GRASPA) in acute leukemia (doi.org/10.2217/ijh-2016-0002)	a	(6a)
	7	EUCTR2013-004262-34-FR/ GRASPANC2013-03	EUCTR	Clinical trial with L-asparaginase encapsulated in erythrcoytes in patients affected by metastatic pancreatic cancer after first line treatment	I e II	Not found		
	8	NCT02195180/ GRASPANC 2013-03	ClinicalTrials.gov	Efficacy and Safety of L-asparaginase Encapsulated in RBC Combined With Gemcitabine or FOLFOX in 2nd Line for Progressive Metastatic Pancreatic Carcinoma	I e II	Circulating Tumor DNA is Prognostic and Potentially Predictive of Eryaspase Efficacy in Second-line in Patients with Advanced Pancreatic Adenocarcinoma (doi.org/10.1158/1078-0432.CCR-20-0950)	g	(8g)
						Erythrocyte-encapsulated asparaginase (eryaspase) combined with chemotherapy in second-line treatment of advanced pancreatic cancer: An open-label, randomized Phase IIb trial (doi.org/10.1016/j.ejca.2019.10.020)	h	(8h)
						A Phase 2b of eryaspase in combination with gemcitabine or FOLFOX as second-line therapy in patients with metastatic pancreatic adenocarcinoma (NCT02195180) (doi.org/10.1093/annonc/mdx369.005)	i	(8i)
	9	EUCTR2018-002211-10-ES/ GRASPA-TNBC-2018-02; NCT03674242	EUCTR	Study to determine whether the addition of eryaspase to gemcitabine and carboplatin will reduce the tumor burden and stabilizate the tumor progression	I e II	TRYbeCA-2: A randomized phase II/III study of eryaspase in combination with gemcitabine and carboplatin chemotherapy versus chemotherapy alone as first-line treatment in patients with metastatic or locally recurrent triple-negative breast cancer (doi.org/10.1093/annonc/mdz242.076)	j	(9j)

Table 2 - Correlation between clinical trials and the corresponding SAs for Eryaspase, GSK3186899, and DNDI 0690 technologies.

	Nº CT	(Main ID / Secondary ID	Primary registration base	Public title	Search	Título(s) of the identified scientific article(s)	Letter AC	Nº of CT and letter of SA
	10	NCT01518517/ GRASPALL 2009-06	ClinicalTrials.gov	GRASPA (Erythrocytes Encapsulating L-asparaginase) in Patients With Relapse of Acute Lymphoblastic Leukemia GRASPIVOTALL	II	Erythrocyte encapsulated l-asparaginase (GRASPA) in acute leukemia (doi.org/10.2217/ijh-2016-0002)	a	(10a)
						Pharmacokinetic and Pharmacodynamic Characterization of Graspa Versus Native L-Asparaginase in Combination with Cooprall Chemotherapy in a Phase 3 Randomized Trial for theTreatment of Patients with Relapsed Acute Lymphoblastic Leukemia (NCT01518517) (doi.org/10.1182/blood.V126.23.2492.2492)	l	(10l)
						Pharmacodynamic characterization of eryaspase (L-asparaginase encapsulated in red blood cells) in combination with chemotherapy in a phase 2/3 trial in patients with relapsed acute lymphoblastic leukemia (NCT01518517) (doi.org/10.1200/JCO.2018.36.15_suppl.7049)	m	(10m)
						Updated Clinical Activity of Graspa Versus Native l-Asparaginase in Combination with Cooprall Regimen in Phase 3 Randomized Trial in Patients with Relapsed Acute Lymphoblastic Leukemia (NCT01518517) (doi.org/10.1182/blood.V126.23.3723.3723)	n	(10n)
						Evaluation of the Impact of the Presence of Neutralizing L-Asparaginase Antibodies on the Efficacy and Safety of Graspa in Phase 3 Randomized Trial Versus Native L-Asparaginase in Patients with Relapsed Acute Lymphoblastic Leukemia (NCT01518517) (doi.org/10.1182/blood.V126.23.3734.3734)	o	(10o)
						Drug monitoring of ERY001 (erythrocyte encapsulated L-asparaginase) and native L-asparaginase (L-ASP) in combination with COOPRALL regimen in Phase 3 randomized trial in patients with relapsed acute lymphoblastic leukemia (doi.org/10.1200/jco.2015.33.15_suppl.e18036	p	(10p)

Table 2 - Correlation between clinical trials and the corresponding SAs for Eryaspase, GSK3186899, and DNDI 0690 technologies.

	N° CT	(Main ID / Secondary ID	Primary registration base	Public title	Search	Title(s) of the identified scientific article(s)	Letter AC	N° of CT and letter of SA
Eryaspase	**10**	NCT01518517/ GRASPALL 2009-06	ClinicalTrials.gov	GRASPA (Erythrocytes Encapsulating L-asparaginase) in Patients With Relapse of Acute Lymphoblastic Leukemia GRASPIVOTALL	II	Clinical activity of ERY001 (erythrocyte encapsulated l-asparaginase) and native l-asparaginase (L-ASP) in combination with COOPRALL regimen in phase III randomized trial in patients with relapsed acute lymphoblastic leukemia (ALL) (doi.org/10.1200/jco.2015.33.15_suppl.7004)	q	(10q)
	11	NCT01523782/ GRASPALL/ GRAALLSA2-2008	ClinicalTrials.gov	Administration of GRASPA (Suspension of Erythrocytes Encapsulating L-asparaginase) in Elderly Patients With First Line Acute Lymphoblastic Leukemia	II	Erythrocyte encapsulated l-asparaginase (GRASPA) in acute leukemia (doi.org/10.2217/ijh-2016-0002)	a	(11a)
						A Phase 2 study of L-asparaginase encapsulated in erythrocytes in elderly patients with Philadelphia chromosome negative acute lymphoblastic leukemia: The GRASPALL/GRAALL-SA2-2008 study (doi.org/10.1002/ajh.24093)	r	(11r)
						L-Asparaginase Loaded Inside Red Cells Has An Acceptable Tolerability Profile On Bilirubin Value (doi.org/10.1182/blood.V122.21.2642.2642)	d	(11d)
	12	NCT02197650/ GRASPALL 2012-10-EAP	ClinicalTrials.gov	Expanded Access Program: Safety of Erythrocytes Encapsulating L-asparaginase (GRASPA®) in Combination With Polychemotherapy in Patients Under 55 Years Old With Acute Lymphoblastic Leukemia (ALL) at Risk to Receive Other Formulation of Asparaginase EAP	II	Expanded Access Program of Graspa for Treatment of Patients with Acute Lymphoblastic Leukemia Unable to Receive Other Form of L-Asparaginase - a Status Update (NCT02197650) (doi.org/10.1182/blood.V126.23.4877.4877)	s	(12s)
						L-Asparaginase Allergic Patients Treated with L-Asparaginase Loaded into Red Blood Cells in an Expanded Access Program. Report of Four Cases (doi.org/10.1182/blood.V124.21.937.937)	t	(12t)
	13	EUCTR2016-004451-70-DK; NOR-GRASPALL-2016	EUCTR	Eryaspase treatment for children and young adults with leukemia and hypersensitivity to PEG-asparaginase.	III	Not found		
	14	EUCTR2016-004451-70-NO; 2016-004451-70-DK NOR-GRASPALL-2016	EUCTR	Eryaspase treatment for children and young adults with leukemia and hypersensitivity to PEG-asparaginase.	III	Not found		

Table 2 - Correlation between clinical trials and the corresponding SAs for Eryaspase, GSK3186899, and DNDI 0690 technologies.

	N° CT	(Main ID / Secondary ID	Primary registration base	Public title	Search	Title(s) of the identified scientific article(s)	Letter AC	N° of CT and letter of SA
	15	NCT03267030/ NOR-GRASPALL-2016	ClinicalTrials.gov	Asparaginase Encapsulated in Erythrocytes for Patients With ALL and Hypersensitivity to PEG-asparaginase	III	Not found		
Eryaspase	**16**	NCT03665441; GRASPANC 2018-01	ClinicalTrials.gov	Study of Eryaspase in Combination With Chemotherapy Versus Chemotherapy Alone as 2nd-Line Treatment in PAC Trybeca-1	III	TRYbeCA-1: A randomized, phase 3 study of eryaspase in combination with chemotherapy versus chemotherapy alone as second-line treatment in patients with pancreatic adenocarcinoma (NCT03665441) (doi.org/10.1093/annonc/mdz155.097)	u	(16u)
	17	NCT04292743/ STUDY00002008	ClinicalTrials.gov	Eryaspase With Modified FOLFIRINOX in Advanced Pancreatic Ductal Adenocarcinoma	III	A phase I dose escalation study of eryaspase in combination with modified FOLFIRINOX in locally advanced and metastatic pancreatic ductal adenocarcinoma (doi.org/10.1200/JCO.2021.39.3_suppl. TPS453	v	(17v)
GSK318689	**18**	NCT03874234; 208436	ClinicalTrials.gov	Safety, Tolerability and Pharmacokinetics (PKs) Investigation of GSK3186899 in Healthy Subjects	Ia e IIa	Not found		
DNDI 0690	**19**	NCT03929016/ 2018-002021-35; DNDi-0690-01; QSC20093229016	ClinicalTrials.gov	Single Oral Dose Escalation Study of DNDI-0690 in Healthy Subjects	Ib e IIb	Not found		
	20	ISRCTN30122193; 2020-003963-24; DNDi-0690-02 / RD 777/34920 / IRAS 288914	ISRCTN	A study to investigate the safety, tolerability and activity of multiple ascending doses of DNDI-0690 in healthy volunteers including assessment of heart and kidney function	Ib e IIb	Not found		

Total clinical studies = 20
Total scientific articles=22
Full list of clinical trials and their scientific articles = 27 (scientific article 'a' is a review and contains data from more than one clinical trial, and article 'd' includes data from two clinical trials).

Source: Prepared by the authors (2021).

Analysis of relationships between the metadata of CTs and SAs

After the collection, organization, and analysis of the metadata, we observed that all CTs had data for the "Main ID," "Secondary ID," "Public Title," "Scientific Title," "Sponsor," "Phase," and "Intervention" fields. However, CTs 1, 3, and 15 were missing information in the "Contacts," "Affiliation," and "Results" fields, respectively, even after complementary verification in the primary databases.

Table 3 shows an overview of data presence or absence for each metadata field in the records of CTs and their respective SAs. For the most part, data are present for all metadata fields in the clinical records and the corresponding SAs. Specifically, CT data in the "Main ID" and "Secondary ID" fields occur in the SA fields "Title," "Affiliation," "Abstract," "Materials and Methods," and "Results." Likewise, CT metadata from the "Affiliation" and "Sponsor" fields was found in the "Affiliation," "Conflict of Interest," "Abstract," "Introduction," "Materials and Methods," "Results," and "Discussion" fields of SAs. However, mismatches in some metadata content in the SA and CT records make the searching process more laborious.

Despite being the primary identifier of CT records, the "Main ID" data is not present in all scientific articles. When present, it is primarily located in the "Title" and "Abstract" sections. The lack of this data in SAs is a significant obstacle, making it difficult to trace data from CTs.

The "Secondary ID" proved to be an important data point for identifying CTs, as it was cited in eight SAs that did not cite the "Main ID." Thus, efforts to trace CTs in the scientific literature should consider both the "Main ID" and the "Secondary ID," since the latter is frequently cited in the "Title," "Abstract," and, particularly, in the "Results" section of SAs.

The public and scientific titles of CTs tended to be close to the titles of the corresponding SAs. This indicates that these metadata can be more effective at obtaining high recall rates despite their low precision.

Finally, although CT results were only rarely available from the ICTRP and primary databases, they could sometimes be found in multiple SA metadata fields, especially "Abstract," "Results," "Discussion," and "Conclusion."

Another critical aspect to highlight is the absence of content for the "Results" metadata in 15 CT records, indicating a significant gap in scientific knowledge. Hence, only five records among the 20 CT analyzed presented data on results, and only four were reported in eight SAs. Thus, the 166.6% correlation for data on results in Table 3 was mainly due to the data repetition in the metadata of the same scientific article.

Information was analyzed in the primary databases of CTs in order to complement the search for the "Results" metadata; however, no new information was available.

Table 3 - Correlation between data from the clinical trials metadata in the respective scientific articles.

1 de 3

		Scientific Articles Metadata										n; %
		Title	Author name	Affiliation/ communication	Conflict of interest	Abstract	Introduction	Material and Methods	Results	Discussion	Conclusion	
Clinical trial metadata	Public/scientific title	2a/ 1c/ 4a/ 6a/ 10a/ 11a/ 1b/ 3e/ 4f/ 8h/ 8i/ 8g/ 9j/ 10l/ 10m/ 10n/ 10o/ 10p/ 10q/ 11d/ 11r/ 12s/ 12t/ 16u/ 17v										27; 92,6

Table 3 - Correlation between data from the clinical trials metadata in the respective scientific articles.

		Scientific Articles Metadata									n; %
		Title	Author name	Affiliation/ communication	Conflict of interest	Abstract	Introduction	Material and Methods	Results	Discussion	Conclusion	
Clinical trial metadata	Main ID	4f/ 12s/ 8i/ 10l/ 10m/ 10n/ 10o/ 16u/ 17v/				1c/ 8h/ 8g/ 9j/ 10m/ 10p/ 10q/ 11r/ 12t		3e	4f			27; 74
	Secondary ID	1b/ 4f/ 11r		4f		4f/ 11d/ 11r		3e/ 11r	1a/ 2a/ 4a/ 6a/ 10a/ 11a			27; 55,5
	Contacts		1c/ 4a/ 1b/ 8h/ 8i/ 8g/ 9j/ 10l/ 10n/ 12s/ 10o/ 10p/ 10q/ 11d/ 11r/ 12t/ 16u/ 17v									26; 69,2
	Affiliation			1a/ 1b/ 8h/ 9j	2a/ 6a	6a/ 9j	6a		6a			17; 58,8
	Sponsor			8g/ 1c/ 8i/ 9j/ 10l/10m/ 10n/ 10o/ 10p/ 10q/ 11d/ 11r/ 12t/ 16u/ 17v	1b/ 1a/ 2a/ 4a/ 6a/ 10a/ 8h/ 11a/ 11r/ 12s	1a/ 2a/ 4a/ 6a/ 9j/ 10a/ 11a/ 12s/ 16u	6a/ 1a/ 4a/ 10a/ 11a/ 11r	8h	1a/ 2a, / 4a/ 6a/ 10a/ 11a	8h		27; 177,8
	Phase	3e/ 4f/ 8h/ 9j/ 10l/ 10m/ 10n/ 10o/ 10p/ 16u/ 10q/ 11r/ 17v				1b/ 3e/ 4f/ 8h/ 8g/ / 9j/ / 10l/ 10m/ 10n/ 10o/ 16u/ 10p/ 10q/ 11r	1b/ 3e/ 8h	1b/ 3e/ 8h/ 11r	1a/ 2a/ 4a/ 10a/ 11a/ 11r/ 3e	8h/ 11r		27; 159,2

Table 3 - Correlation between data from the clinical trials metadata in the respective scientific articles.

		Scientific Articles Metadata										n; %
		Title	Author name	Affiliation/ communication	Conflict of interest	Abstract	Introduction	Material and Methods	Results	Discussion	Conclusion	
Clinical trial metadata	Intervention					1b/ 1c/ 1d/ 3e/ 4f/ 8g/ 8h/ 9i/ 10l/ 10m/ 10n/ 10o/ 10p/ 10q/ 11d/ 11r/ 12s/ 12t/ 16u/ 17v	1b/ 11r	1a/ 1b/ 1c/ 1d/ 2a/ 3e/ 4a/ 6a/ 8h/ 10a/ 11r	1a/ 1b/ 2a/ 3e/ 8h/ 11r	3e/ 8h/ 11r	3e	27; 159,2
	Results					3e/ 8g/ 8h/ 11d/ 11r			2a/ 3e/ 8h/ 11r	3e/ 8h/ 11r	3e	9; 166,6

Note: The indices were calculated considering the quantity of a given data from metadata in scientific articles by the quantity of the same data (identical or technically similar) from metadata in clinical trials; the 'n' was calculated by the presence of data from metadata in clinical trials, corrected by the number of publications generated from these clinical trials. The rates that surpassed 100% indicate that the information from the clinical trial was present in more than one metadata of the respective scientific publication.

Source: Prepared by the authors (2021).

Assessment publication of results from CTs

After observing the low rate of publication of results in the CT databases (n=5), we sought to identify whether the trials were already completed or in progress. We therefore searched for metadata on the completion date of CTs in our sample to identify ones that had already been completed. Of the 20 CTs recorded, 16 were listed as completed (Figure 1A). For the 16 completed CTs, we observed that nine (56%; trials 4, 5, 7, 13, 14, 15, 16, 17, and 19) had no published results of any kind. Hence, only seven studies (44%) were published, with only four (25%; trials 2, 3, 8, and 11) in both SA and CT databases. Meanwhile, results for two CTs (13%; trials 1 and 10) could only be found in the SA databases, while one (6%; study 6) could only be found in the CT databases (Figure 1B). Of the CTs with no published results (56%) in any database (Figure 1B), three CTs (trials 4, 16, and 17) included information on the experimental design in summarized publications. For the other six records, data were only available in CT databases (Figure 1C). These data indicate that the non-publication of CT results in the scientific literature or clinical databases makes such results unavailable to society, drastically restricting access to information.

https://doi.org/10.1590/2318-0889202234e210065

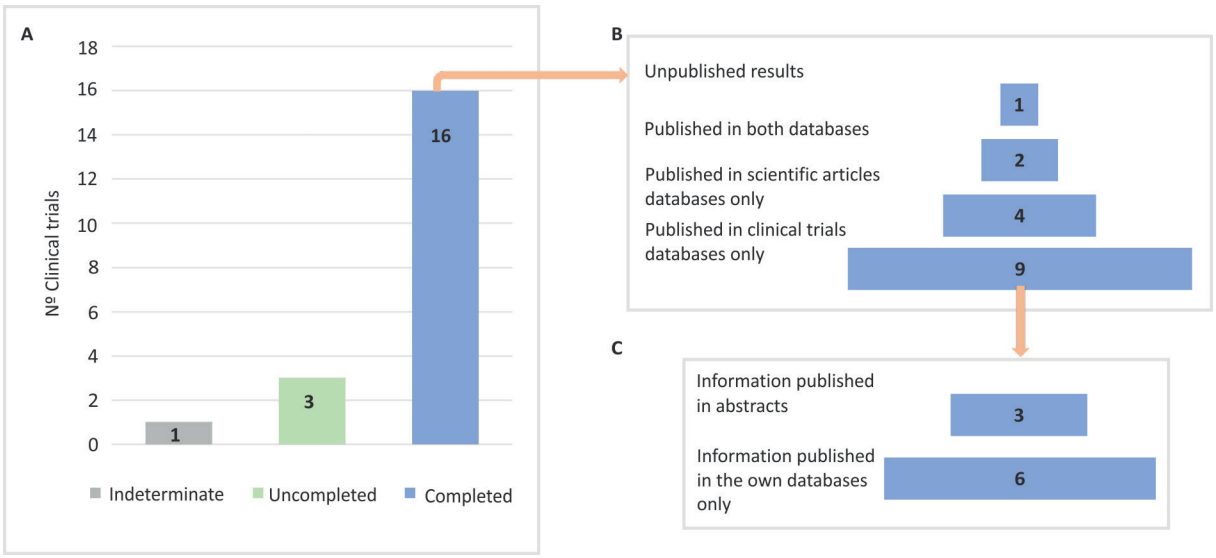

Figure 1 - Evaluation of clinical trials results publications.

Note: A: State of completion of the clinical trial; B: Publication means of clinical trials results; C: Publication of clinical trails information (no results); The slice represented as Indeterminate means lack of information in the metadata indicating the completion of the study.

Source: Elaborated by the authors (2021)

Our data indicate an average time lapse of 40.8 months between the completion of a CT and its results becoming available in clinical databases or the scientific literature, with figures ranging from 20 to 71 months (Figure 2).

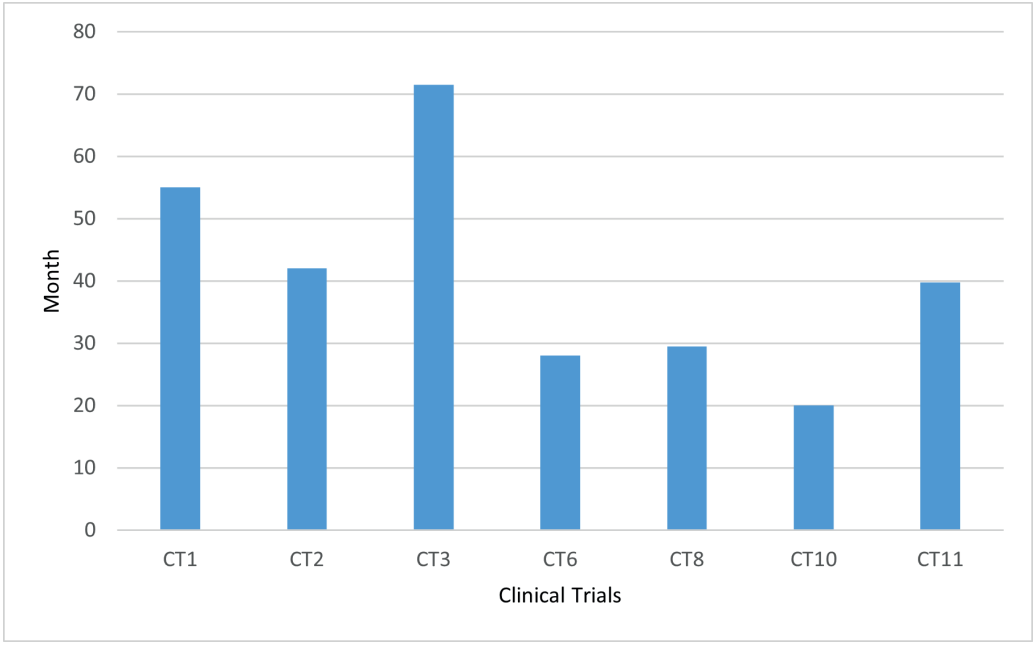

Figure 2 - Time between completion of the clinical trial and publication of the results.

Source: Prepared by the authors (2021).

Discussion

The development of a healthcare solution is initially related to scientific discoveries that constitute the first phase of the innovation process and lead to a large volume of publications in scientific journals (Targino, 2001). As innovations come to technological maturity, research in the medical field can be protected by patents. In addition, before products come to market, research results can be disclosed on platforms designed to record CTs and the pipelines of sponsoring companies (Pimenta, 2017; Ellwood; Williams; Egan, 2020).

Although scientific journals are the primary means for publishing CTs, there are gaps in the information available in SA databases (Shrivastava; Shrivastava; Ramasamy, 2018). Similarly, the evaluated CT databases are incomplete. While the information contained in SA titles is useful for identifying CTs, our data indicate that the metadata related to the sponsor's name, study phase, and intervention's name are the identifiers that are most reliably reported in the scientific literature. The latter metadata therefore proved to be the most effective way to track CTs published in the scientific literature.

We highlight the challenge of linking scientific publications to the corresponding CT. This difficulty is due to identifying information of the CT being spread throughout the article, hindering the search for a non-specialist reader. In addition, the citation of the "Main ID" is sometimes completely lacking in scientific publications. Often (55.5%), a secondary identifier provided by the company sponsoring the CT is used instead (Table 2).

The WHO recommends the use of a "Secondary ID" for studies with the same purpose that must be registered on several databases to comply with specific local legislation (World Health Organization, 2018). However, despite this guidance, the mere presence of a "Secondary ID" does not ensure that a CT is accurately cited in a SA if there is no "Main ID" citation. The absence of the "Main ID" must always be questioned in scientific publications since it jeopardizes the mission of the clinical record, which is to allow comprehensive access to information.

The non-availability of results for most of the completed CTs in the clinical databases consulted is a significant bottleneck for the publication and transparency of information, especially given the phenomenon of selective publication of results or non-publication of negative results (Sayão; Sales; Felipe, 2021). Many authors have highlighted the problem of selective publication, when only studies with positive or statistically significant results are published, while those with negative or non-significant results remain undisclosed (Page; Mckenzie; Higgins, 2018; Borysowski; Wnukiewicz-Kozłowska; Górski, 2020).

In addition, it should be noted that in several cases, results were published an average of three years from the end of the CT. This indicates a considerable time lag for communication. This data is consistent with studies showing that the results of up to 50% of CTs may not be available using any means (Hopewell *et al*., 2001; Chan *et al*., 2014; Schmucker *et al*., 2014). Although legal regulations in the European Union and the US impose an obligation to include summarized CT results in their respective registry, compliance with these regulations has generally been lacking (Prayle; Hurley; Smyth, 2012; Chen *et al*., 2016; Goldacre *et al*., 2018). Our data reveal failures in the publication of results from CTs that have already been completed. However, it is difficult to determine whether these failures result from selective publication or merely an excessive delay in disclosure.

The failure to publish CT results is underscored by the fact that the results of six studies (67% of all studies with published results) are only available in CT databases, with no published results in the scientific literature. As such, these databases represent a potential source of information, sometimes unique, about technological innovations in healthcare.

No scientific publications reporting clinical data for GSK3186899 and DNDI-0690 were found, which clearly shows that CT databases are the only way to obtain information on these technologies. Records in CT databases generally provide information on their experimental design and development, providing essential scientific data on the corresponding technology. As such, data from CT records play a fundamental role in providing information

to the scientific community on new healthcare solutions since these records often hold information that can only be found in these databases (Viergever *et al*., 2014). The information in these databases can be used to indicate technologies in development that may be directed to overcome current scientific issues.

It is clear that researchers are generally influenced almost exclusively by SAs and only rarely seek other information sources (Pimenta, 2017). This leads to a tendency to ignore other useful sources when planning and conducting a scientific study, leading to duplicated research and investments. Meanwhile, encouraging the scientific community to seek other information sources can help increase the quality and efficiency of scientific efforts.

Conclusion

This study demonstrates the potential of CT databases as a source of important information for the development of technological solutions. However, our results highlight the challenge of accurately identifying links between scientific publications and the corresponding CT records since the metadata are not identical and can require manual analysis by an expert professional.

This study sheds light on three major obstacles facing the scientific community and the public with respect to access to information and transparency in clinical studies: (i) the lack of correlation between CT records and scientific publications, (ii) the considerable time lag between the conclusion of a CT and the publication of results, and (iii) the non-availability of results from completed CTs in any of the databases consulted.

Acknowledgments

We thank Dr. Cicera Henrique da Silva, full professor at Fiocruz, for a critical review of this article.

Contributors

B.A.F.S. TRIGUEIROS was responsible for the conceptualization, data collection, methodology, analysis and interpretation of data, writing and discussion. A.R. ÁVILA was responsible for analysis, discussion, and review. F.P. PIMENTA was responsible for the conceptualization, methodology, analysis and interpretation of data, writing, discussion and review. All authors are responsible for the approval of the final version.

References

Borysowski, J.; Wnukiewicz-Kozłowska, A.; Górski, A. Legal regulations, ethical guidelines and recent policies to increase transparency of clinical trials. *British Journal of Clinical Pharmacology*, v. 86, n. 4, p. 679-686, 2020. Doi: https://doi.org/10.1111/bcp.14223.

Brasil. *Resolução RDC n° 39, de 5 de junho de 2008. Regulamento para a realização de pesquisa clínica*. [*S.l.*]: Anvisa, 2008. Disponivel em: https://bvsms.saude.gov.br/bvs/saudelegis/anvisa/2008/res0039_05_06_2008.html. Acesso: 20 maio, 2021.

Brasil. *Resolução RDC nº 36, de 27 junho de 2012. Altera a RDC nº 39, de 05 de junho de 2008*. [*S.l.*]: Anvisa, 2012. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/anvisa/2012/rdc0036_27_06_2012.html. Acesso: 20 maio, 2021.

Chan, A.-W. *et al*. Increasing value and reducing waste: addressing inaccessible research. *The Lancet*, v. 383, n. 9913, p. 257-266, 2014. Doi: https://doi.org/10.1016/S0140-6736(13)62296-5.

Chen, R. *et al*. Publication and reporting of clinical trial results: cross sectional analysis across academic medical centers. *BMJ*, v. 352, n. i.637, p.1-8, 2016. Doi: https://doi.org/10.1136/bmj.i637.

Ellwood, P.; Williams, C.; Egan, J. Crossing the valley of death: Five underlying innovation processes. *Technovation*, v. 109,

n. 102162, p. 1-11, 2020. Doi: https://doi.org/10.1016/j.technovation.2020.102162.

Goldacre, B. *et al*. Compliance with requirement to report results on the EU Clinical Trials Register: cohort study and web resource. *BMJ*, v. 362, n. k3218, p. 1-10, 2018. Doi: https://doi.org/10.1136/bmj.k3218.

Gouy, C. M. L.; Porto, T. F.; Penido, C. Avaliação de ensaios clínicos no Brasil: histórico e atualidades. *Revista Bioética*, v. 26, n. 3, p. 350-359, 2018. Doi: https://doi.org/10.1590/1983-80422018263254.

Hopewell, S. *et al*. Time to publication for results of clinical trials. *In*: The Cochrane Collaboration (ed.). *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd, 2001. p. MR000011. Doi: https://doi.org/10.1002/14651858.MR000011.

Page, M. J.; Mckenzie, J. E.; Higgins, J. P. T. Tools for assessing risk of reporting biases in studies and syntheses of studies: a systematic review. *BMJ Open*, v. 8, n. 3, p. e019703, 2018. Doi: https://doi.org/10.1136/ bmjopen-2017-019703.

Pimenta, F. P. A patente como fonte de informação (des) necessária para a Biotecnologia em Saúde. *Transinformação*, v. 29, n. 3, p. 323-332, 2017. Doi: https://doi.org/10.1590/2318-08892017000300009.

Prayle, A. P.; Hurley, M. N.; Smyth, A. R. Compliance with mandatory reporting of clinical trial results on ClinicalTrials.gov: cross sectional study. *BMJ*, v. 344, p. d7373-d7373, 2012. Doi: https://doi.org/10.1136/bmj.d7373.

Quental, C.; Salles Filho, S. Ensaios clínicos: capacitação nacional para avaliação de medicamentos e vacinas. *Revista Brasileira de Epidemiologia*, v. 9, n. 4, p. 408-424, 2006. Doi: https://doi.org/10.1590/S1415-790X2006000400002.

Sayão, L. F.; Sales, L. F.; Felipe, C. B. M. Invisible science: publication of negative research results. *Transinformação*, v. 33, p. e200009, 2021. Doi: https://doi.org/10.1590/2318-0889202133e200009.

Schmucker, C. *et al*. Extent of Non-Publication in Cohorts of Studies Approved by Research Ethics Committees or Included in Trial Registries. *PLoS ONE*, v. 9, n. 12, p. e114023, 2014. Doi: https://doi.org/10.1371/journal.pone.0114023.

Shrivastava, S.R.; Shrivastava, P. S.; Ramasamy, J. Emphasizing the necessity to register and report the results of all clinical trials: World Health Organization. *Medical Journal of Dr. D.Y. Patil Vidyapeeth*, v. 11, n. 1, p. 81-82, 2018. Doi: https://doi.org/10.4103/MJDRDYPU.MJDRDYPU_105_17.

Targino, M. G. Comunicação científica na sociedade tecnológica: periódicos electrônicos em discussão. *Comunicação e Sociedade*, v. 3, p. 93-112, 2001. Doi: https://doi.org/10.17231/comsoc.3(2001).1317.

Tse, T.; Fain, K. M.; Zarin, D. A. How to avoid common problems when using ClinicalTrials.gov in research: 10 issues to consider. *BMJ*, v. 361, n. k1452, p. 1-8, 2018. Doi: https://doi.org/10.1136/bmj.k1452.

Van Raan, A. F. J. Scientometrics: state-of-the- art. *Scientometrics*, v. 38, n. 1, p. 205-218, 1997.

Vanti, A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ciência da Informação*, v. 31, n. 2, p. 369-379, 2002. Doi: https://doi.org/10.1590/S0100-19652002000200016.

Viergever, R. F. *et al*. The quality of registration of clinical trials: still a problem. *PLoS ONE*, v. 9, n. 1, p. e84727, 2014. Doi: https://doi.org/10.1371/journal.pone.0084727.

World Health Organization. *International standards for clinical trial registries*: the registration of all interventional trials is a scientific, ethical and moral responsibility, version 3.0, 2018. Disponível em: https://apps.who.int/iris/handle/10665/274994. Acesso em: 5 set. 2021.