

Data flow within global clinical trials: a scoping review

Kaitlyn Kwok,¹ Neha Sati,² Louis Dron,² Srinivas Murthy¹

To cite: Kwok K, Sati N, Dron L, *et al*. Data flow within global clinical trials: a scoping review. *BMJ Global Health* 2022;**7**:e008128. doi:10.1136/bmjgh-2021-008128

Handling editor Seye Abimbola

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjgh-2021-008128>).

Received 27 November 2021
Accepted 27 March 2022

ABSTRACT

Objective To document clinical trial data flow in global clinical trials published in major journals between 2013 and 2021 from Global South to Global North.

Design Scoping analysis

Methods We performed a search in Cochrane Central Register of Controlled Trials (CENTRAL) to retrieve randomised clinical trials published between 2013 and 2021 from *The BMJ*, *BMJ Global Health*, *the Journal of the American Medical Association*, *the Lancet*, *Lancet Global Health* and *the New England Journal of Medicine*. Studies were included if they involved recruitment and author affiliation across different country income groupings using World Bank definitions. The direction of data flow was extracted with a data collection tool using sites of trial recruitment as the starting point and the location of authors conducting statistical analysis as the ending point.

Results Of 1993 records initially retrieved, 517 studies underwent abstract screening, 348 studies underwent full-text screening and 305 studies were included. Funders from high-income countries were the sole funders of the majority (82%) of clinical trials that recruited across income groupings. In 224 (73.4%) of all assessable studies, data flowed exclusively to authors affiliated with high-income countries or to a majority of authors affiliated with high-income countries for statistical analysis. Only six (3.2%) studies demonstrated data flow to lower middle-income countries and upper middle-income countries for analysis, with only one with data flow to a lower middle-income country.

Conclusions Global clinical trial data flow demonstrates a Global South to Global North trajectory. Policies should be re-examined to assess how data sharing across country income groupings can move towards a more equitable model.

INTRODUCTION

There are increasing calls in the global health community for a reframing of the current model of global health research, with more Global South-led leadership.¹ For greater generalisability, clinical trials will often need to be global in scope, as the COVID-19 pandemic has highlighted.²

Clinical trials that enrol patients across different income groupings of the world are typically funded by high-income country (HIC) funders and are led by HIC

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Global clinical trials are predominantly led by researchers and institutions from the Global North.

WHAT THIS STUDY ADDS

⇒ There is an unsurprising pattern of consistent data movement in global randomised clinical trials from patients in the Global South to analysis in the Global North.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ Efforts should be made to improve research infrastructure for the conduct of data analysis in the Global South for trials that recruit globally to avoid a continued imbalance of data movement towards high-income countries.

⇒ Researchers in the Global North should advocate for more equitable structures of global randomised clinical trials.

investigators, but recruit patients in lower income regions.¹ This asymmetry of clinical trial leadership has led to increasing concerns about equity, data ownership, post-trial accessibility and capacity building, particularly for problems that affect individuals around the planet.^{1 3 4} This trajectory coincides with clinical trial sponsors and trial coordinators primarily being based in HICs without equal data sharing or capacity building in recruiting regions.⁵ Often, recruitment for these global trials occurs solely in low-income countries (LICs) and lower middle-income countries (LMICs), yet analysis is conducted solely in HICs.⁶⁻⁸ This appears to follow a common, extractive pattern from Global South to benefit the Global North, which may be attributed to traditional power imbalances, and the ongoing colonial model of global health research.⁹

Despite this, there remains a lack of quantitative estimates on the scale of extractive data flows in global clinical trials, and the subdivisions in how this data are handled. This scoping review seeks to systematically document clinical trial data flow involving authors



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Faculty of Medicine, The University of British Columbia, Vancouver, British Columbia, Canada

²Cytel Inc, Vancouver, British Columbia, Canada

Correspondence to

Dr Srinivas Murthy;
srinivas.murthy@cw.bc.ca

affiliated with countries of different income categorisations across major medical journals.

METHODS

Search strategy and selection criteria

A search was performed on 29 April 2021 using the Cochrane Central Register of Controlled Trials (CENTRAL) for randomised clinical trials published between 2013 and 2021. Publications were retrieved from *The BMJ*, *BMJ Global Health*, *the Journal of the American Medical Association (JAMA)*, *The Lancet*, *Lancet Global Health*, and *the New England Journal of Medicine (NEJM)*, as major journals that would publish high-impact work with global relevance. Abstract screening for eligible articles was first performed by one data collector (KK), followed by full-text screening conducted by two independent data collectors (KK and NS). The study fell outside the scope of requiring ethics committee review according to Canada's Tri-Council Policy Statement on the ethical conduct of research involving humans.

Studies were included if they were multisite studies with at least one high (or upper middle) income country site and at least one lower middle (or low) income site, involved authors primarily affiliated with institutions from countries with various income groupings, or both. Income group classification is based on World Bank definitions, classifying countries as LICs, LMICs, upper middle-income countries (UMICs) and HICs. There were no restrictions on disease area or intervention. Studies were only included if they were designed as a randomised clinical trial. We excluded studies if the country of recruitment and all author affiliations stemmed from one income-group classification exclusively. Studies that were not full publications were also excluded. In cases of disagreement for inclusion, both data collectors reviewed the paper together and came to a consensus.

Data extraction

A data collection instrument was created and used to extract information on the following key variables: trial sample size, disease area of interest, trial site locations, trial funding, author contributions and affiliations, procedures for data analysis, housing and management and data sharing availability.

Funding statements were used to determine which country acted as the main source of funding for each study. We also looked at the types of organisations funding each study and categorised them as industry-related, such as pharmaceutical companies, or non-industry-related, such as national funders, foundations or academic institutions.

Data analysis

Data movement was assessed using the site(s) of recruitment as the starting point and the location of authors conducting statistical analyses as the ending point. Author institution affiliations or trial sponsor details from the report were used as indicators for the site of statistical analysis. If information on the authors involved

Figure 1 – Study Flow Diagram

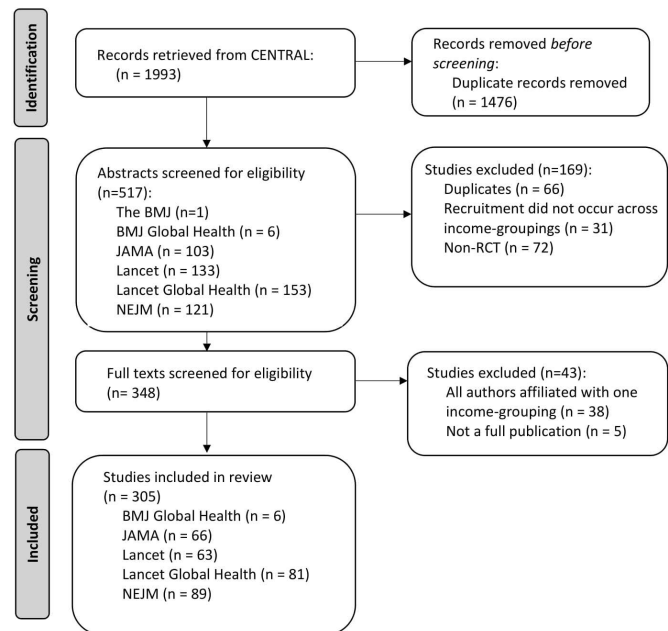


Figure 1 Study flow diagram. CENTRAL, Cochrane Central Register of Controlled Trials; *NEJM*, *New England Journal of Medicine*.

in statistical analysis or statements of where the analysis took place could not be found in the manuscript, the online supplemental material, or the protocol, the sponsoring country, was extracted from clinical trial databases to determine final location of data flow.

Study design was defined using randomisation flow diagrams or statements used in the methods section of papers. When multiple funders were listed, the primary funder was determined by the funding body acknowledged in the abstract section of the paper or the first organisation listed in the funding acknowledgement. Large philanthropic funders were considered to be based at the centres of their headquarters.

Patient and public involvement

There was no patient or public involvement in this study.

RESULTS

A total of 1993 records were retrieved, 517 studies underwent abstract screening, 348 studies underwent full-text screening and 305 studies were included for data extraction and analysis (see [figure 1](#)). A total of 67 (22.0%) records were reviewed in duplicate to adjust and tailor the data collection instrument. Of 238 (69.5%) records were subsequently extracted independently, of which 26 studies demonstrated a disagreement threshold of <20%.

The largest proportion of randomized clinical trials (RCTs) was focused on infectious diseases (40.7%), followed by chronic diseases (22%) and acute diseases (14%) ([table 1](#)). The dominant funding sources were from high and upper MICs, making up 91.8% of RCTs ([table 2](#)).

Table 1 Study characteristics table

	Included studies n=305
Journal	
<i>BMJ Global Health</i>	6 (2.0%)
<i>JAMA</i>	66 (21.6%)
<i>The Lancet</i>	63 (20.7%)
<i>The Lancet Global Health</i>	81 (26.6%)
<i>NEJM</i>	89 (29.2%)
Year of publication	
2013–2015	82 (26.9%)
2016–2018	130 (42.6%)
2019–2021	93 (30.5%)
Topic	
Acute diseases and conditions	43 (14.1%)
Childhood development	17 (5.6%)
Chronic diseases and conditions	67 (22.0%)
Infectious diseases	124 (40.7%)
Maternal and neonatal care	34 (11.1%)
Mental and behavioural disorders	12 (3.9%)
Other	8 (2.6%)

NEJM, New England Journal of Medicine.

Direction of data in RCTs

The direction of data movement in RCTs is described in [table 3](#). Of the 186 studies which had a clear single site of analysis, there were only six studies identified (3.2%), which demonstrated data flow to an LMIC or UMIC,

Table 2 Funding table

	Included studies n=305
Income classification of funding country*	
High-income economies	250 (82.0%)
Upper middle-income economies	30 (9.8%)
Lower middle-income economies	0
Low-income economies	0
UNICEF/WHO	4 (1.3%)
Multi-income classification funding†	15 (4.9%)
None	6 (2.0%)
Funding source type	299
Industry	61 (20.4%)
Non-industry	238 (79.6%)
Education or medical institution	10 (3.3%)
Government	117 (39.1%)
Philanthropy	111 (37.1%)

*Based on World Bank definitions.
†Combined low income and high income or upper middle income and high income.

Table 3 Data flow and management table

	All studies n=305
Clear single site of analysis	186
LIC to HIC	17 (9.14%)
LMIC to HIC	47 (25.27%)
UMIC to HIC	21 (11.29%)
LIC/LMIC/UMIC to LMIC	1 (0.54%)
LIC/LMIC/UMIC/HIC to UMIC	5 (2.69%)
LIC/LMIC/UMIC/HIC to HIC	94 (50.54%)
Trial not registered	1 (0.54%)
Multiple sites of analysis	119
Majority of authors affiliated with HIC	45 (37.82%)
Majority of authors affiliated with LIC, LMIC or UMIC	20 (16.81%)
Equal number of authors affiliated with HIC and with LIC, LMIC or UMIC	26 (21.85%)
Recruited and analysed in the same country	28 (23.53%)
Data sharing statement availability	
Available	86 (28.2%)
Not available	219 (71.8%)

HIC, high-income country; LIC, low-income country; LMIC, lower middle-income country; UMIC, upper middle-income country.

only one of which had an LMIC site for analysis.^{10–15} The remaining studies with a clear single identified site of analysis, data flowed exclusively to HICs (n=179). For the 119 studies with multiple sites of analysis, the bulk of statistical authorship was heavily weighted towards HIC, with 38% having a majority of statistical authors being based in HIC settings. Overall, of the 305 RCTs examined, 224 (73.4%) had authors conducting statistical analysis exclusively in HICs or had most statistical authors located in HICs.

Statements relating to the site of data management and the site of data housing were also extracted from each trial's article, online supplemental material, and protocol. Limited information was found explicitly described about management or housing across all journals ([table 3](#)).

DISCUSSION

In this analysis of data flows within individual global randomised clinical trials that cross income groupings, there is a pattern of data movement from the Global South towards the Global North. Although previous studies have sought to define the regional disparities and barriers that exist in clinical trials, documentation of data movement within global trials has yet to be examined.^{16–18} One potential explanation for our findings is that research experience and resources for conducting statistical analysis are traditionally found among authors

who are trained within and living in the Global North.¹ With the headquarters of most global health organisations and industry sponsors often residing in the Global North, it is unsurprising that the statistical analysis and data stewardship for clinical trials largely occurred in these same countries.⁹ However, consistently sending data for analysis to the Global North will only continue to exacerbate an imbalance in clinical research, often in contradiction of the principles of many global health organisations and national funders.

Creating a framework wherein participant data are managed and analysed in diverse regions, with high-income regions sending data to lower income regions, creates better opportunities to grow research capabilities across the globe, rather than solely in HICs, recognising the need for global collaboration. We found very few examples of this in our review, suggesting that historical barriers to this collaborative environment persist, even across high-quality and contemporary clinical trials.

The ability for local research institutions to synthesise evidence and respond rapidly within their own environmental context is integral to combat emerging and evolving health threats.¹⁹ Research programmes that emphasise locally led research have already been shown to strengthen institutional capabilities in the Global South.²⁰ Without continued and integrated expertise involving researchers from the Global South, the necessary infrastructure to handle these emergent threats may not be sufficient to react accordingly.

Calls to promote Global South-led research are not new.^{1 5} Historically, the idea of globalisation of health research has primarily focused on partnering countries of different income groupings together with funding and leadership stemming from HICs. However, researchers from sponsoring countries may have imbalanced power in collaborative relationships in aspects of design, conduct and analysis of that research,²¹⁻²³ demonstrated by this work revealing imbalanced data accessibility. In this review, it was observed, unsurprisingly, that the majority of analysis, and its attached expertise, weighted heavily towards the HIC collaborators present in the relationship, necessitating changes in HIC ideas of ownership for change to occur.

The COVID-19 pandemic highlighted a gap in the global health research landscape with a surge of clinical trials being conducted in the Global North and a lack of representation in the Global South, which was not representative of global case burden.²¹ Rather than building a knowledge database that was representative of diverse populations, many HICs were producing and reporting data that were only applicable to their particular income grouping.²¹ One of the benefits of global health research is to discover the differences that exist between populations and to learn whether it is possible to apply a particular intervention to the global population.⁴ In emerging health crises, having the capacity to produce and analyse data across the globe and share data equally between researchers in the Global North and Global South could

facilitate a more efficient collaborative process. Developing a stronger local capacity for research is sustainable in the long run and better takes into account resource accessibility and environmental limitations.^{24 25}

Strengths and limitations

Our analysis included any type of randomised clinical trials, making it possible to look at data flow movement broadly across different fields of global health research. However, there are limitations that must be addressed in our data in both the extraction and analysis processes. It is important to note that author contributions and listed affiliations are not always representative of the entire efforts of an author in the submission of a manuscript. Indeed, within the included studies, it is not possible to definitively identify those authors who coded and manually reviewed data. Instead, surrogacy was used based on the contributory statements of authors. More granular author statements or protocols, or use of CRediT taxonomy being required, would facilitate greater understandings pertinent to this paper, although, to date, these declarations are limited.²⁶

In tracking the movement of data, we used the sites of authors acknowledged for statistical analysis as a surrogate measure for the final location of data transfer; however, there are other methods that this could have been accomplished, such as looking at corresponding author location. In contrast, the use of corresponding author may be more indicative of overall study direction and writing, rather than statistical analysis and data handling and, therefore, less informative than the presently extracted data. When site of statistical analysis was undetermined, sponsor location was assessed instead which also may not be representative of final location of data transfer.

As there is no specific way to track the flow of data, we use several compounding assumptions for analysis. For example, whether data moved back and forth between sites in the analysis process is not possible to determine, and the ability for authors within the Global South to freely access the data generated is not possible to determine. As such, the broad findings that data had a preponderance towards unidirectional flow from the Global South to the Global North may not translate to an inability of collaborators from the Global South to access the analysable data sets generated from their primary research.

We looked exclusively at clinical trials that were published in higher impact or global health-specific journals. This possibly introduces a bias towards researchers and trials that would publish in those journals, compared with specialty-focused or regionally focused journals.

Finally, we looked exclusively at clinical trials that crossed income boundaries in their recruitment regions. We did not look at trials that exclusively recruited within one income grouping and whether these trials still primarily exported their data to higher income groupings, which is a separate, yet equally important question.

CONCLUSION

The movement of data in clinical trials that recruit globally follows a Global South to Global North trajectory, highlighting an unsurprising imbalance in global health research relationships. If this pattern continues, it is unlikely that continued improvement in global clinical trials will be made possible. Research systems must address steps to move towards a more balanced and equitable model for global health research.

Twitter Srinivas Murthy @srimurthy99

Acknowledgements We would like to thank The Global Health Network for their critical review of this manuscript.

Contributors All authors were involved in the conception, design and conduct of this study. All authors had access to the original data, with SM acting as guarantor.

Funding This study was funded by BC Children's Hospital Foundation.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Lang T, Siribaddana S. Clinical trials have gone global: is this a good thing? *PLoS Med* 2012;9:e1001228.
- WHO Solidarity Trial Consortium, Pan H, Peto R, et al. Repurposed Antiviral Drugs for Covid-19 - Interim WHO Solidarity Trial Results. *N Engl J Med* 2021;384:497-511.
- Dron L, Dillman A, Zoratti MJ, et al. Clinical trial data sharing for COVID-19-Related research. *J Med Internet Res* 2021;23:e26718.
- Miller JE, Mello MM, Wallach JD, et al. Evaluation of drug trials in high-, middle-, and low-income countries and local commercial availability of newly approved drugs. *JAMA Netw Open* 2021;4:e217075.
- Aguilera B, DeGrazia D, Rid A. Regulating international clinical research: an ethical framework for policy-makers. *BMJ Glob Health* 2020;5:e002287.
- Labhardt ND, Ringera I, Lejone TI, et al. Effect of offering same-day art vs usual health facility referral during home-based HIV testing on linkage to care and viral suppression among adults with HIV in Lesotho: the cascade randomized clinical trial. *JAMA* 2018;319:1103.
- Kennedy SB, Bolay F, Kieh M, et al. Phase 2 placebo-controlled trial of two vaccines to prevent Ebola in Liberia. *N Engl J Med* 2017;377:1438-47.
- Homan T, Hiscox A, Mweresa CK, et al. The effect of mass mosquito trapping on malaria transmission and disease burden (SolarMal): a stepped-wedge cluster-randomised trial. *Lancet* 2016;388:1193-201.
- Abimbola S, Pai M. Will global health survive its decolonisation? *Lancet* 2020;396:1627-8.
- Dawson R, Diacon AH, Everitt D, et al. Efficiency and safety of the combination of moxifloxacin, pretomanid (PA-824), and pyrazinamide during the first 8 weeks of antituberculosis treatment: a phase 2B, open-label, partly randomised trial in patients with drug-susceptible or drug-resistant pulmonary tuberculosis. *Lancet* 2015;385:1738-47.
- Althabe F, Chomba E, Tshetu AK, et al. A multifaceted intervention to improve syphilis screening and treatment in pregnant women in Kinshasa, Democratic Republic of the Congo and in Lusaka, Zambia: a cluster randomised controlled trial. *Lancet Glob Health* 2019;7:e655-63.
- Beardsley J, Wolbers M, Kibengo FM, et al. Adjunctive dexamethasone in HIV-associated cryptococcal meningitis. *N Engl J Med* 2016;374:542-54.
- Widmer M, Piaggio G, Nguyen TMH, et al. Heat-Stable carbetocin versus oxytocin to prevent hemorrhage after vaginal birth. *N Engl J Med* 2018;379:743-52.
- Zhu F-C, Wurie AH, Hou L-H, et al. Safety and immunogenicity of a recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in Sierra Leone: a single-centre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* 2017;389:621-8.
- Venter WDF, Moorhouse M, Sokhela S, et al. Dolutegravir plus two different prodrugs of tenofovir to treat HIV. *N Engl J Med Overseas Ed* 2019;381:803-15.
- Viergever RF, Li K. Trends in global clinical trial registration: an analysis of numbers of registered clinical trials in different parts of the world from 2004 to 2013. *BMJ Open* 2015;5:e008932.
- Marshall IJ, L'Esperance V, Marshall R, et al. State of the evidence: a survey of global disparities in clinical trials. *BMJ Glob Health* 2021;6:e004145.
- Alemayehu C, Mitchell G, Nikles J. Barriers for conducting clinical trials in developing countries- a systematic review. *Int J Equity Health* 2018;17:37.
- Stewart R, El-Harakeh A, Cherian SA, et al. Evidence synthesis communities in low-income and middle-income countries and the COVID-19 response. *Lancet* 2020;396:1539-41.
- Kok MO, Gyaopong JO, Wolffers I, et al. Towards fair and effective north-south collaboration: realising a programme for demand-driven and locally led research. *Health Res Policy Syst* 2017;15:96.
- Park JJH, Mogg R, Smith GE, et al. How COVID-19 has fundamentally changed clinical research in global health. *Lancet Glob Health* 2021;9:e711-20.
- Shiffman J. Global health as a field of power relations: a response to recent commentaries. *Int J Health Policy Manag* 2015;4:497-9.
- McCoy D, Chand S, Sridhar D. Global health funding: how much, where it comes from and where it goes. *Health Policy Plan* 2009;24:407-17.
- Bowsher G, Papamichail A, El Achi N, et al. A narrative review of health research capacity strengthening in low and middle-income countries: lessons for conflict-affected areas. *Global Health* 2019;15:23.
- Vaduganathan M, Samman Tahhan A, Greene SJ, et al. Globalization of heart failure clinical trials: a systematic review of 305 trials conducted over 16 years. *Eur J Heart Fail* 2018;20:1068-71.
- CRediT - Contributor Roles Taxonomy, 2022. Available: <https://casrai.org/credit/> [Accessed April 5, 2022].