

Ongoing and planned Randomized Controlled Trials of AI in medicine: An analysis of Clinicaltrials.gov registration data

1. Mattia Andreoletti, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland. Corresponding: mattia.andreoletti@hest.ethz.ch
2. Berkay Senkalfa, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland
3. Alessandro Blasimme, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

Abstract

The integration of Artificial Intelligence (AI) technologies into clinical practice holds significant promise for revolutionizing healthcare. However, the realization of this potential requires rigorous evaluation and validation of AI applications to ensure their safety, efficacy, and clinical significance. Despite increasing awareness of the need for robust testing, the majority of AI-related Randomized Controlled Trials (RCTs) so far have exhibited notable limitations, impeding the generalizability and proper integration of their findings into clinical settings. To understand whether the field is progressing towards more robust testing, we conducted an analysis of the registration data of ongoing and planned RCTs of AI in medicine available in the Clinicaltrials.gov database. Our analysis highlights several key trends and challenges. Effectively addressing these challenges is essential for advancing the field of medical AI and ensuring its successful integration into clinical practice.

Introduction

Machine learning (ML), Large Language Models (LLMs), and other cutting-edge technologies at the core of Artificial Intelligence (AI) hold promise to revolutionize clinical care. The potential of AI applications in the clinical space depends on their capacity to streamline patient care, to reduce medical errors, to improve diagnostic and prognostic accuracy, and to predict which therapeutic options are best suited to the specific characteristics of each patient.

While AI-specific challenges, such as limited interpretability and bias risks have been widely discussed, especially in relation to its clinical use¹, a lesser degree of attention has been paid to issues related to the evidence base supporting the implementation of AI-based applications in the clinical context. Recently, the importance of rigorous testing and real-world validation, including well-designed Randomized Controlled Trials (RCTs), has been emphasized before AI can be

integrated into medical practice. However, the vast majority (85.9%) of FDA-approved clinical AI applications (both medical devices and medical software) received 510(k) clearance^{2,3}, meaning they were cleared for medical use based on tests demonstrating substantial equivalence to a predicate, that is, an already licensed product. Predicate equivalence mostly relies on non-clinical testing, especially in cases of devices that regulators classify as low-risk⁴.

Moreover, so far, a significant portion of the evidence surrounding medical AI tools has focused on their performance metrics, such as diagnostic accuracy, with dubious quality⁵. However, improved performance does not necessarily lead to a net clinical benefit for patients⁶⁻⁸. Lacking adequate clinical validation, there remains a significant gap in our understanding of the actual clinical significance of such new tools.

In recent years, the medical and scientific communities have increasingly emphasized the need for rigorous evaluation and validation of AI technologies before their integration into clinical practice. Numerous studies and commentaries have highlighted the potential risks of prematurely adopting AI-driven solutions in healthcare⁹⁻¹³. Concerns about reliability, safety, and ethical implications have dominated discussions, leading to calls for systematic evaluation frameworks and robust testing protocols¹⁴.

This paper aims to examine the ongoing and planned RCTs assessing medical use of AI, based on Clinicaltrials.gov registration data, and focusing on their reported characteristics. Our analysis will provide valuable insights into the current state of clinical research in medical AI.

Methods

For our study, we queried the ClinicalTrials.gov database for studies listing "Artificial Intelligence" as an intervention or treatment. The data was downloaded as an CSV file and imported into Excel to facilitate review and analysis. All available information was included in the downloaded dataset. We then filtered the results by study type to select only interventional studies, and by date to identify studies with completion dates extending beyond the search date (February 27, 2024).

Our inclusion criteria encompassed only randomized studies enabling the integration of AI into clinical practice, thereby influencing patient health management by clinical teams, as well as

those permitting the application of AI-assisted tools in clinical settings, encompassing diagnosis, treatment, and prognostication.

Then, from an initial pool of 253 trials, we excluded 24 due to non-randomized status and 76 for not being applicable to our criteria. Furthermore, 21 studies lacked sufficient relevance to the medical field, and 5 failed to demonstrate clear utilization or impact assessment of AI. The remaining 127 trials comprised our final dataset for analysis [Figure 3].

We systematically extracted data from all included trials, capturing details such as whether they were multi- or single- center, conducted on a national or multinational scale, participant demographics including age and sex, medical specialties and oncological subspecialties, types of interventions employed, countries involved, enrollment status and numbers, study duration, sponsor types, and characteristics of primary outcome measures. To enhance reproducibility and maintain control over results, we minimized interpretation to a significant extent. Our analysis and subsequent visualization of the data were conducted using a combination of GraphPad Prism, MS Excel, and Adobe Photoshop.

Results

Study Centers and Sponsors

Of the included trials, 56.69% (n=72 of 127) were conducted in a single-center setting, while the share of multi-center clinical trials stood at 26.77% (n=34 of 127). The remaining 16.54% (n=21 of 127) did not specify the single or multi-center character of the study in the database. Only 7.87% of the included trials (n=10) were conducted in a multinational setting, while 75.59% were carried out in a single country (n=96). The remaining 16.54% (n=21) did not report country information. Beside the 21 trials that do not report study locations, the country with the highest number of trials (n=29) was the USA; followed by China with 11 trials, and Italy with 10 trials. Furthermore, Taiwan was involved in 8 studies, while France, Germany, and Spain were each involved in 7, followed closely by the UK with 5 studies and Hong Kong with 4 studies. For the rest, 24 separate countries across different continents were involved with study numbers ranging from 1 to 3.

Tangentially, looking at the sponsors of individual studies, the highest share was found to be sponsored by universities, public institutions that include hospitals and local governments, or

foundations typically focusing on advancement of health, with 85.04% of the analyzed trials (n=108) falling into this category. Another 6.30% (n=8) were sponsored by industry entities like private companies and incorporations. 5.51% of the studies (n=7) were registered as sponsored by individuals, which were often observed to be medical doctors. Finally, 3.15% (n=4) of the included studies were sponsored by what could be defined as private clinics and health institutions that function under nonprofit organizational definitions [Figure 1].

Demographics and Enrollment

Among the included 127 studies, 86.61% (n=110) enrolled only adults, while 9.45% (n=12) carried out the study in both children and adults, whereas only 3.94% (n=5) enrolled children only. A vast majority of the studies did not have a single-sex character, with 88.98% (n=113) enrolling both males and females. The remaining 11.02% (n=14) of the studies were sex-specific, with 9.45% (n=12) of the assessed RCTs enrolling only females, and 1.57% (n=2) enrolling males only [Figure 2a].

Focusing on the enrollment characteristics of the 127 evaluated studies and looking at the enrollment status, 46.46% (n=59) of the included studies were readily recruiting participants at the time of data extraction. On the other hand, 38.58% (n=49) of the RCTs were not yet recruiting study participants. Among the remaining studies, 7.87% (n=10) had “active, not recruiting” status, while 6.30% (n=8) were enrolling by invitation. Finally, 1 study corresponding to 0.79% was withdrawn before completion, due to lack of funding. As for estimated enrollment numbers, excluding the withdrawn study which reported 0 participants, the trial with the highest enrollment reported 180000 participants while the study with the least enrollment reported only 12 participants. Again, excluding the withdrawn study from the 127 trials included, the mean estimated enrollment number was found to be 6002.437, while the median value stood at 374 [Figure 2b].

Medical Specialties

The most represented medical specialty in our sample is gastroenterology (n=27), including neoplasms and cancers. This was followed by oncology excluding cancers of the gastrointestinal tract with 19 studies, and cardiovascular diseases, also with 19 studies. Neurology also had a significant presence among studies assessing AI use with 15 RCTs, followed by endocrinology

with 11 studies. The remaining specialties included psychology/psychiatry/behavioral science (6 studies), infectious diseases and infectiology (5 studies), pulmonology (5 studies), nephrology/urology (5 studies), dentistry (4 studies), orthopedy (3 studies), reproductive medicine (3 studies), ophthalmology (2 studies), and pharmacology, critical care, and dietetics with 1 study for each category.

Among all 41 RCTs with an oncological focus across specialties, 21 studies were reserved for colorectal neoplasms and cancers, followed by 8 breast cancer studies, 3 studies focusing on gynecological cancers and 3 studies on hepatobiliary cancers. Among remaining studies, 2 RCTs tackled AI use in lung cancers, 2 in skin cancer, followed by 1 in gastric cancer and 1 looking at chemotherapy effects in different tumor types [Figure 2c].

Intervention Classes

Regarding the types of interventions under testing, 46 studies focused on medical devices incorporating AI, followed by 18 studies on AI-powered behavioral interventions. Additionally, 11 studies each explored AI-assisted medical procedures and AI-based diagnostic tests. Notably, the remaining 36 studies reported various "other" interventions. These diverse applications of AI in medical practice include adopting existing AI tools for improving medical practice, including for example, large language model (LLM) chatbots for diagnosis assistance, AI-based screenings, and AI apps to drive public health interventions. Due to their variety, these interventions cannot be neatly classified into predefined categories [Figure 2d]

Primary Outcome Measures

Following established definitions, clinical outcomes are those endpoints that represent objective measurements of direct patient benefit, including metrics related to mortality, survival, and disease-associated quality of life. On the other hand, operational outcomes encompass measures such as diagnostic accuracy, procedure satisfaction, and patient anxiety. In our sample, 55.91% of the RCTs (n=71) measured operational primary outcomes, while the remaining 44.09% (n=56) focused on primary clinical outcomes [Figure 2e].

Study Duration

Among the 127 studies assessed, the earliest study start date was recorded as March 1, 2017, while the latest study start date was set to be February 15, 2025. On the other hand, the earliest

study completion date was March 1, 2024, whereas the latest study completion date was set to be June 30, 2050. Looking at study duration for all the included studies, the shortest duration recorded was 1 day (0.03 months), followed by a study of 90 days (3 months), while the longest study was set to last 9816 days (327.2 months). The average study duration for all included studies was calculated as 1054.2 days, corresponding to 35.14 months [Figure 2f].

Discussion

Analyzing ongoing and planned RCTs of AI in medicine can offer valuable insights into whether the field is advancing towards the adoption of rigorous clinical testing standards. Our analysis provides key indications of the trajectory of clinical evaluation within this domain.

Despite repeated calls for rigorous evaluation through multicenter and multinational studies¹⁵, our findings suggest that such study designs are still lacking. The majority of trials in our dataset were conducted in single-center settings, with only a small percentage being multinational. Additionally, it is noteworthy that most of these studies are concentrated in the United States and China. While not unexpected, this geographical skew may limit the generalizability of findings to other regions across the globe. These findings highlight the need for more collaborative efforts across institutions and countries to ensure the generalizability and applicability of AI-driven solutions across diverse demographics, clinical settings, and cultures.

Moreover, the studies exhibited a broad spectrum of sample sizes. This significant variance may reflect a lack of standardized approaches to sample size calculation in RCTs involving AI. The study duration also expressed high variability, nonetheless the average duration seems to be aligned with RCTs in other fields. RCTs are indeed notoriously long experiments with the median time from starting until completing follow-up reported as 2.6 years¹⁶.

The dominance of trials sponsored by universities and public institutions raises questions about the balance between profit-driven research and public interest. Publicly funded trials provide an opportunity for critical product optimization before entry into clinical practice by limiting commercial and market pressures in RCTs, which may otherwise bias private sponsors towards overlooking underperforming aspects or potentially harmful safety considerations of products being assessed¹⁷. In contrast, profit-driven RCTs may provide resources for larger-scale trials with more extensive participant recruitment and longer follow-up periods. This can enhance the

statistical power and robustness of study findings, thereby improving the reliability of research outcomes. Additionally, private sponsors may have the infrastructure and expertise to conduct complex multicenter or multinational trials, which can provide more diverse and representative findings. One potential rationale for this phenomenon could be the dynamic and intricate nature of the regulatory framework surrounding AI in healthcare, which is still in the process of evolution. Uncertainty about regulatory requirements, including approval processes and compliance standards, may deter companies from investing in RCTs for AI-driven medical interventions. The risk of regulatory hurdles or delays could discourage private-sector sponsors from undertaking clinical trials. Moreover, companies may be hesitant to invest in profit-driven RCTs for AI applications if there is uncertainty about reimbursement policies, or healthcare provider adoption rates. On the one hand, the lack of clear pathways to commercialization and monetization could further discourage companies from allocating resources to conducting well-designed trials. On the other hand, competition in a fast-growing market may create incentive to favor the most rapid way to regulatory clearance – that is by showing predicate equivalence (510k clearance) instead of seeking clinical validation via RCTs.

The analysis of medical specialties targeted by AI-focused trials provides insight into areas where AI is expected to have the greatest impact. Gastroenterology, oncology, and cardiovascular diseases emerged as the primary areas of focus, indicating a forthcoming trend in these domains. An interesting dissection can be made that a significant portion of the RCTs assessing medical devices, with 21 of the 46 studies falling under the gastroenterology specialty, showing an apparent trend for AI use in general, and medical devices incorporating AI in particular for gastroenterology practice. Among all the observed types of interventions being tested, medical devices being the most prevalent intervention class highlights a narrow focus within the realm of AI applications, despite the potential for a much broader range of uses¹⁸.

The predominance of operational primary outcome measures, particularly those related to diagnostic accuracy and precision, underscores the emphasis on improving clinical workflows and decision-making processes. While operational outcomes are essential for assessing the practical implications of AI tools, it is crucial to also prioritize clinical endpoints that directly impact patient outcomes, such as mortality, survival rates, and quality of life, to better capture the actual effectiveness and clinical significance of AI interventions in healthcare.

Notably, our findings align closely with the results of reviews on published RCTs of medical AI^{19–23}. These reviews consistently highlight several prevalent challenges, which we also observed in our sample. Most trials were single-center studies, focused on operational endpoints rather than clinically significant outcomes, and were based on small samples. These features significantly hinder the assessment of medical AI and its integration into clinical practice. However, a limitation of these reviews is that the field of AI, including medical AI, is evolving very rapidly. As a result, reviews that rely on published RCTs, which were designed long before their results were published, cannot capture the latest trends. Our analysis of ongoing and planned trials indicates that this trend has not significantly changed. As noted, there is a pressing need for more internationally collaborative, multi-center trials with larger sample sizes, diverse outcome measures, and clinically meaningful endpoints.

However, it is important to also acknowledge that conducting RCTs for AI applications in medicine is inherently complex and faces numerous methodological and practical challenges^{24,25}. These obstacles, coupled with the elevated costs associated with RCTs, the need to swiftly penetrate the market, the absence of clear regulatory incentives, and the lack of a standardized regulatory framework for medical AI applications, are likely forestalling the development of more robust testing standards. RCTs for testing medical AI are still in their early stages. The challenges highlighted in our review, as well as in other studies, should be meticulously addressed in the design of future trials.

It is important to acknowledge the limitations of our study. We queried only one database (Clinicaltrials.gov) to maintain methodological uniformity, as different databases use varying classification parameters. Additionally, the potential for bias due to the lack of pre-registration of trials must be recognized. This omission could unintentionally skew our data analysis and interpretation.

Data availability

All data produced in the present study are available upon request to the authors.

Author contributions

MA*: Designed the study, collected the data, and drafted the manuscript.

SB*: Organized and analyzed the data, designed the figures, and contributed to drafting the manuscript.

AB: Contributed to the study design and revised the manuscript.

*Equally contributed to the work.

Competing Interests

The authors declare no competing interests.

Acknowledgements

This study did not receive any funding.

References

1. Nittas, V. *et al.* Beyond high hopes: A scoping review of the 2019–2021 scientific discourse on machine learning in medical imaging. *PLOS Digit Health* **2**, e0000189 (2023).
2. Benjamens, S., Dhunoo, P. & Meskó, B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *npj Digit. Med.* **3**, 118 (2020).
3. Muehlematter, U. J., Daniore, P. & Vokinger, K. N. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *The Lancet Digital Health* **3**, e195–e203 (2021).
4. Muehlematter, U. J., Bluethgen, C. & Vokinger, K. N. FDA-cleared artificial intelligence and machine learning-based medical devices and their 510(k) predicate networks. *The Lancet Digital Health* **5**, e618–e626 (2023).

5. Andaur Navarro, C. L. *et al.* Systematic review finds “spin” practices and poor reporting standards in studies on machine learning-based prediction models. *Journal of Clinical Epidemiology* **158**, 99–110 (2023).
6. Hassan, C. *et al.* Real-Time Computer-Aided Detection of Colorectal Neoplasia During Colonoscopy: A Systematic Review and Meta-analysis. *Ann Intern Med* **176**, 1209–1220 (2023).
7. Brocklehurst, P. *et al.* Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *The Lancet* **389**, 1719–1729 (2017).
8. Wong, A. *et al.* External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. *JAMA Intern Med* **181**, 1065 (2021).
9. Van Genderen, M. E. *et al.* Charting a new course in healthcare: early-stage AI algorithm registration to enhance trust and transparency. *npj Digit. Med.* **7**, 119 (2024).
10. Rajpurkar, P., Chen, E., Banerjee, O. & Topol, E. J. AI in health and medicine. *Nat Med* **28**, 31–38 (2022).
11. Van De Sande, D. *et al.* To warrant clinical adoption AI models require a multi-faceted implementation evaluation. *npj Digit. Med.* **7**, 58 (2024).
12. Gorenstein, L., Soffer, S., Apter, S., Konen, E. & Klang, E. AI in radiology: is it the time for randomized controlled trials? *Eur Radiol* **33**, 4223–4225 (2023).
13. Byrne, D. W., Domenico, H. J. & Moore, R. P. Artificial Intelligence for Improved Patient Outcomes—The Pragmatic Randomized Controlled Trial Is the Secret Sauce. *Korean J Radiol* **25**, 123 (2024).
14. McCradden, M. D., Joshi, S., Anderson, J. A. & London, A. J. A normative framework for artificial intelligence as a sociotechnical system in healthcare. *Patterns* **4**, 100864 (2023).

15. Wu, E. *et al.* How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med* **27**, 582–584 (2021).
16. Ioannidis, J. P. A. Effect of the Statistical Significance of Results on the Time to Completion and Publication of Randomized Efficacy Trials. *JAMA* **279**, 281 (1998).
17. Lundh, A., Lexchin, J., Mintzes, B., Schroll, J. B., & Bero, L. (2017). Industry sponsorship and research outcome. *Cochrane database of systematic reviews*, (2).
18. Silcox, C. *et al.* The potential for artificial intelligence to transform healthcare: perspectives from international health leaders. *npj Digit. Med.* **7**, 88 (2024).
19. Wang, J. *et al.* Investigation and evaluation of randomized controlled trials for interventions involving artificial intelligence. *Intelligent Medicine* **1**, 61–69 (2021).
20. Plana, D. *et al.* Randomized Clinical Trials of Machine Learning Interventions in Health Care: A Systematic Review. *JAMA Netw Open* **5**, e2233946 (2022).
21. Lam, T. Y. T. *et al.* Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review. *J Med Internet Res* **24**, e37188 (2022).
22. Zhou, Q., Chen, Z., Cao, Y. & Peng, S. Clinical impact and quality of randomized controlled trials involving interventions evaluating artificial intelligence prediction tools: a systematic review. *npj Digit. Med.* **4**, 154 (2021).
23. Han, R. *et al.* Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review. *The Lancet Digital Health* **6**, e367–e373 (2024).
24. Guo, C. *et al.* Challenges for the evaluation of digital health solutions—A call for innovative evidence generation approaches. *npj Digit. Med.* **3**, 110 (2020).
25. Bates, D. W. How to regulate evolving AI health algorithms. *Nat Med* **29**, 26–26 (2023).

Figures legend

Figure 1: An overview of study centers and sponsor types of included studies.

Included 127 studies were analyzed according to their single- versus multi-center character, single- versus multi-national design, and sponsor types. An overview of countries involved per each unique study is provided on a map of the world. For multinational studies, each country involved is registered as a unique involvement for that country.

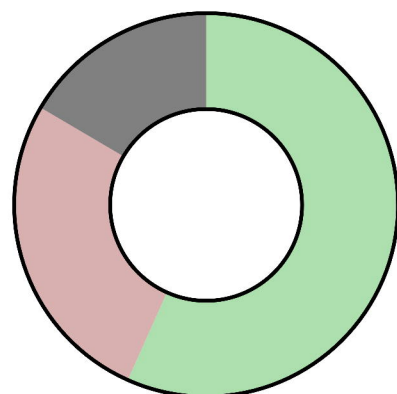
Figure 2: Comprehensive characteristic analyses of included studies.

Comprehensive characteristic analyses of the included studies showing (A) study demographic and sexes targeted, (B) enrollment status and estimated enrollment numbers, (C) medical specialties and subspecialties of oncological studies, (D) interventional input categories, (E) primary outcome groups, and (F) average study duration. One study withdrawn before completion is not included in estimated enrollment numbers but included in the rest of the characteristic analyses. One study with the same start and completion date has been included as 1 day in the calculation of average duration data.

Figure 3: PRISMA flow diagram showing the study design and included studies.

A total of 253 studies were initially identified with no duplicates. During screening, clinical trials that were non-randomized, or randomization status was given as not applicable were excluded. Among the 153 studies sought for retrieval, all were assessed for eligibility, were 26 studies that did not fully suffice the inclusion criteria were excluded, resulting in a total of 127 studies included in the subsequent analyses.

Single vs Multicenter

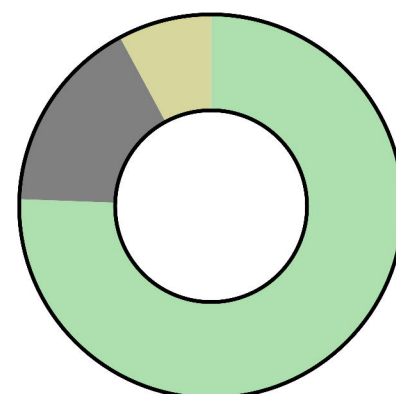


n=127

% of Studies

- 56.69% Singlecenter
- 26.77% Multicenter
- 16.54% Not Reported

Country Involvement

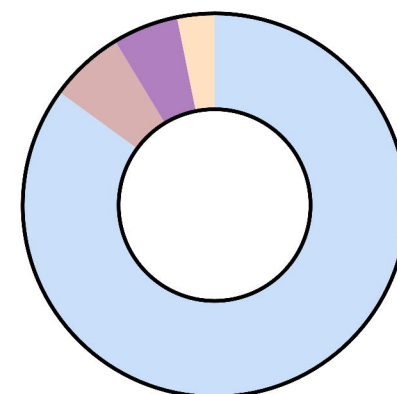


n=127

% of Studies

- 75.59% Single Country
- 16.54% Not Reported
- 07.87% Multinational

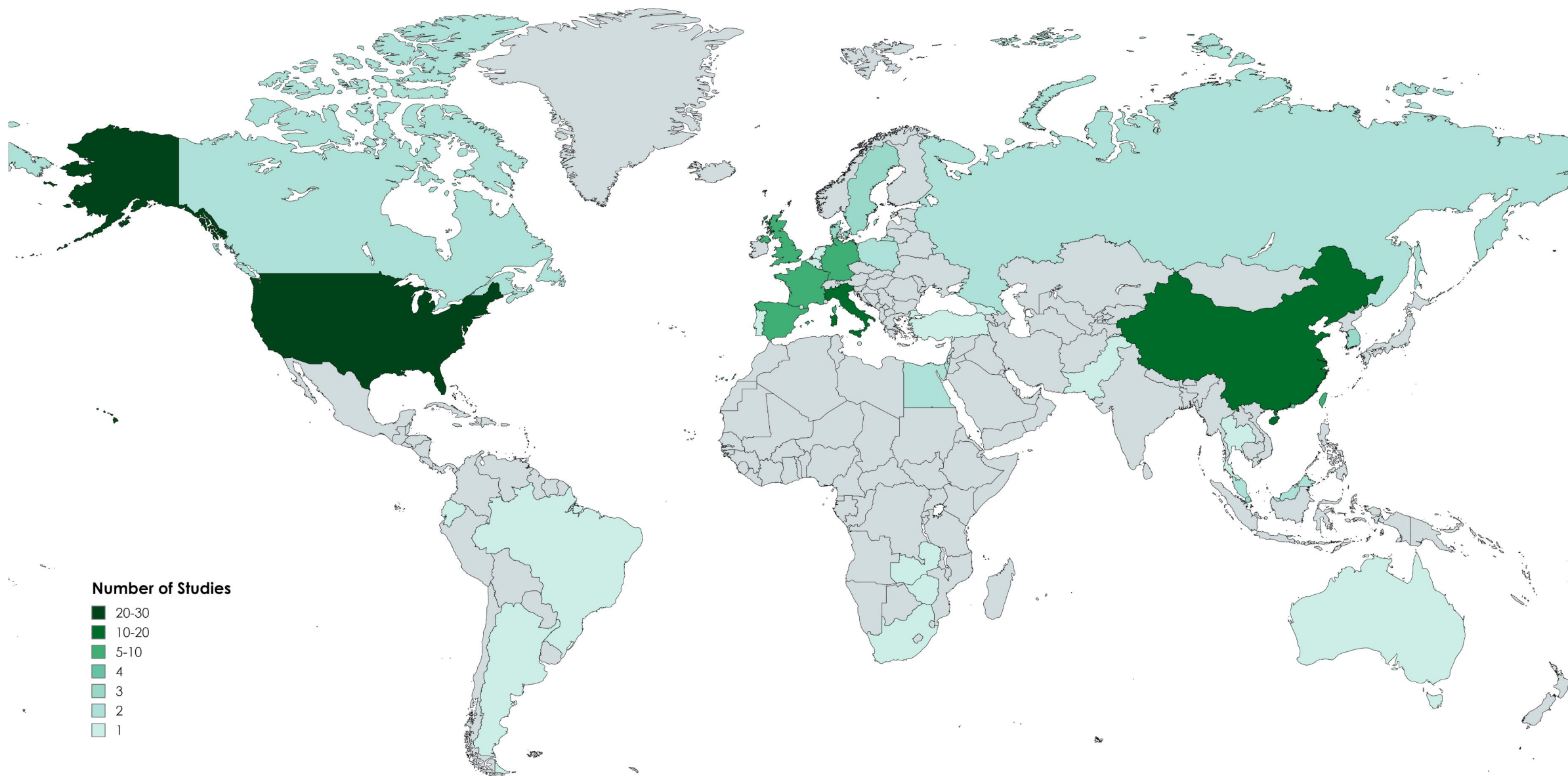
Sponsor Types



n=127

% of Studies

- 85.04% Uni/Public/Foundation
- 06.30% Industry/Private Companies
- 05.51% Individuals
- 03.15% Private Nonprofit Organizations



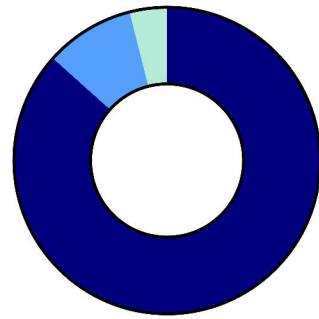
- 29 USA
- 21 Not Reported
- 11 China
- 10 Italy
- 8 Taiwan
- 7 France
- 7 Germany
- 7 Spain
- 5 UK
- 4 Hong Kong
- 3 Sweden
- 3 Denmark
- 3 Singapore
- 3 S. Korea
- 2 Egypt
- 2 Israel
- 2 Poland
- 2 Russia
- 2 Malaysia
- 2 Canada
- 1 Argentina
- 1 Ecuador
- 1 Portugal
- 1 S. Africa
- 1 Zambia
- 1 Zimbabwe
- 1 Türkiye
- 1 Georgia
- 1 Australia
- 1 Pakistan
- 1 Netherlands
- 1 Brazil
- 1 Thailand
- 1 Belgium

Number of Studies

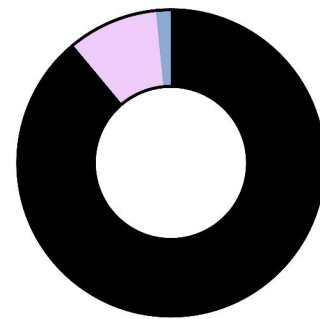
- 20-30
- 10-20
- 5-10
- 4
- 3
- 2
- 1

A

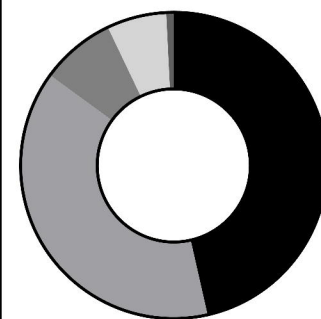
medRxiv preprint doi: <https://doi.org/10.1101/2024.07.09.24310133>; this version posted July 9, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Study Demographic

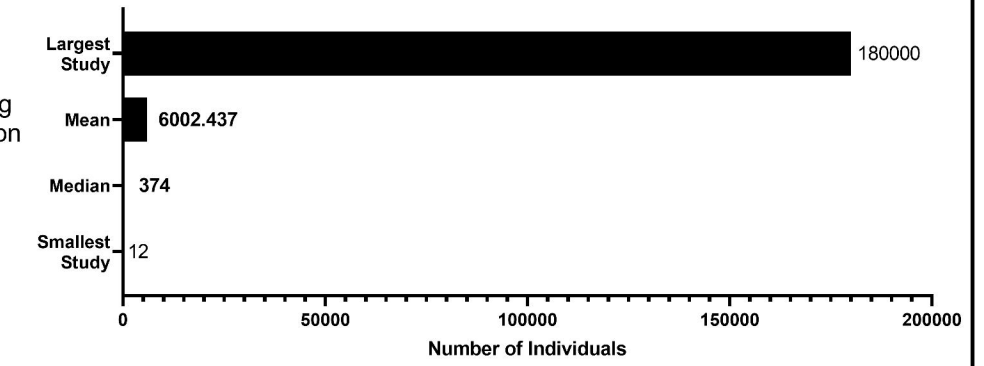
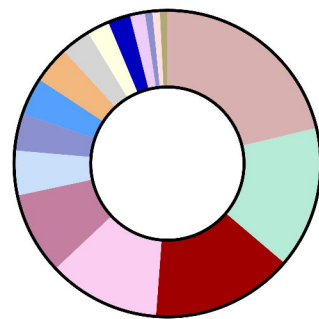
n=127

Sexes

n=127

B**Enrollment Status**

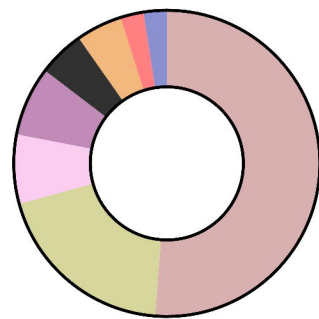
n=127

Enrollment (estimated)**C****Specialty**

n=127

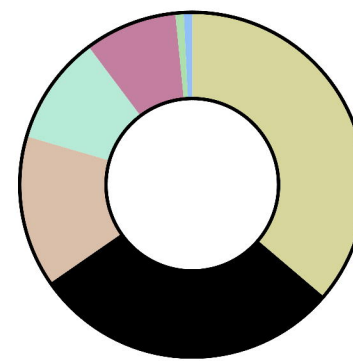
of Studies

- 27 Gastroenterology
- 19 Oncology (excl. GI tract)
- 19 Cardiovascular Diseases
- 15 Neurology
- 11 Endocrinology
- 6 Psychology/Psychiatry/Behavioral Science
- 5 Infectious Diseases
- 5 Pulmonology
- 5 Nephrology/Urology
- 4 Dentistry
- 3 Orthopedy
- 3 Reproductive Medicine
- 2 Ophthalmology
- 1 Pharmacology
- 1 Critical Care
- 1 Dietetics

of Studies Among Oncology Studies

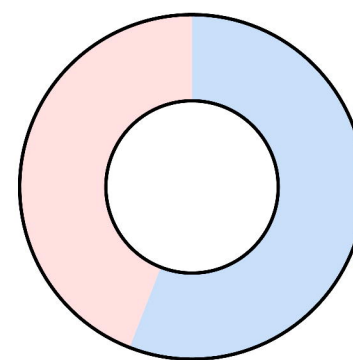
n=41

- 21 Colorectal Neoplasms/Cancers
- 8 Breast Cancers
- 3 Gynecological Cancers
- 3 Hepatobiliary Cancers
- 2 Lung Cancers
- 2 Skin Cancers
- 1 Gastric Cancers
- 1 Overall Chemotherapy

D**Interventional Inputs**

n=127

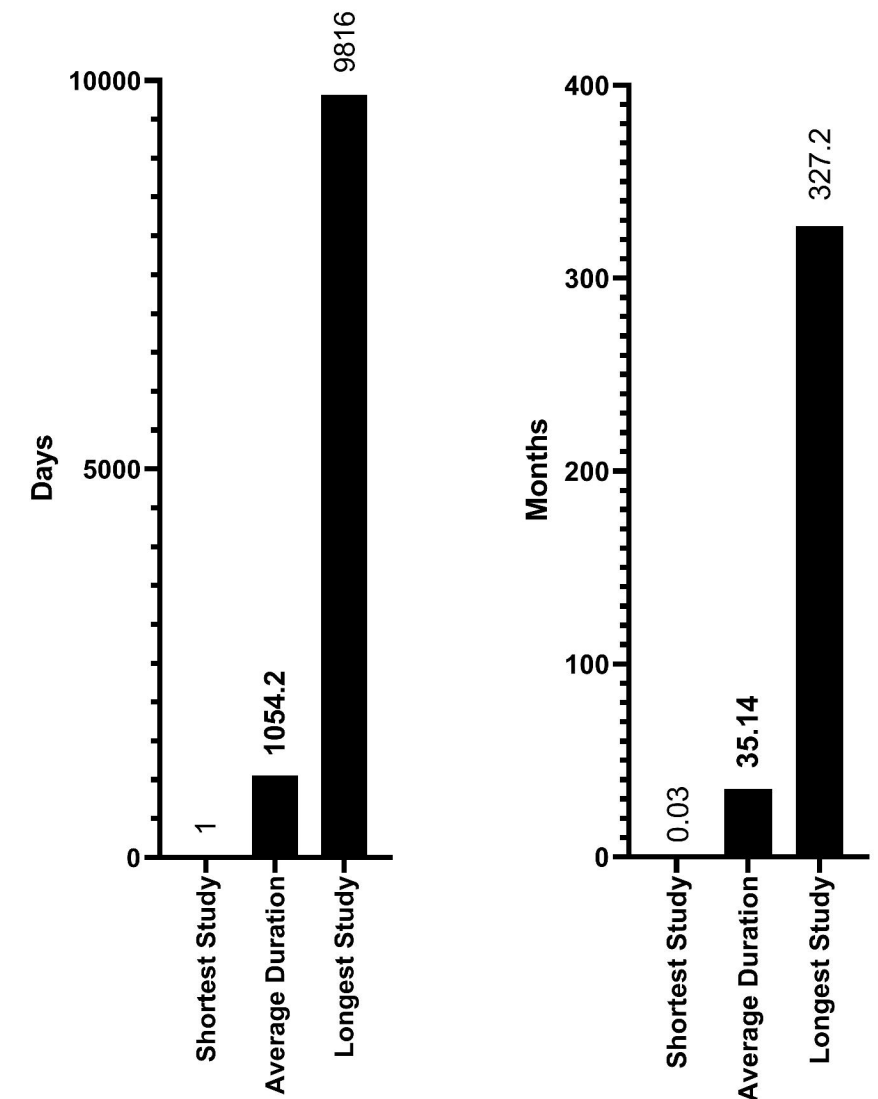
- 46 Device
- 37 Other
- 18 Behavioral
- 13 Diagnostic Test
- 11 Procedure
- 1 Dietary Supplement
- 1 Genetic

E**Primary Outcome Groups**

n=127

% of Studies

- 55.91% Operational
- 44.09% Clinical

F**Avg. Study Duration in days and months**

PRISMA Flow Diagram

