Paper SA06

Diversity, Inclusion, and Equality - It's just good science!

Roxanne McGuire, Veramed, Edinburgh, Scotland
Chris Ward, Veramed, London, England

ABSTRACT

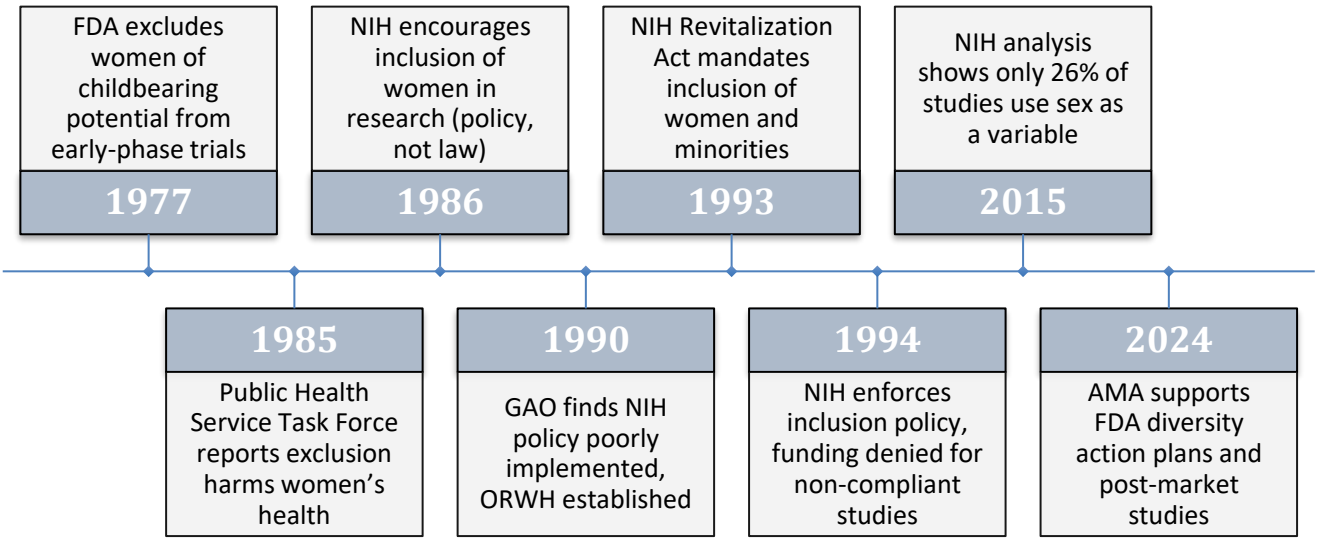
Diversity, inclusion, and equality in clinical trials are critical to advancing equitable healthcare and ensuring scientific integrity. Historically, many clinical studies have underrepresented populations based on race, ethnicity, gender, age, and socioeconomic background, leading to gaps in knowledge and disparities in healthcare outcomes. The National Institutes of Health (NIH) Revitalisation Act 1993 was the first legal requirement pushing for the inclusion of women in federally funded medical research and a turning point of clinical trials.

In this paper I will discuss that incorporating diverse populations in clinical research allows for a more accurate understanding of how treatments affect various groups, enhances the generalisability of results, and promotes equitable access to innovative therapies. As regulatory bodies increasingly emphasise the need for representative data, prioritising diversity and equality in clinical trials is both a scientific and moral imperative.

INTRODUCTION

In September 1977, the Food and Drug Administration (FDA) issued guidelines recommending the exclusion of all women of childbearing potential from Phase I and II clinical trials, including single women, those using contraception, and even those whose partners had vasectomies. The intent was to protect potential pregnancies from unknown risks posed by experimental drugs, a cautionary response to the thalidomide tragedy of the 1960s, when thousands of children were born with severe birth defects.

Whilst well-intentioned, this policy had unintended consequences that reverberated across decades of medical research. It created a significant data gap in understanding how drugs affect women differently from men, delayed progress in women's health, and overlooked the autonomy of women to make informed decisions about participating in research. In effect, protecting hypothetical pregnancies came at the cost of real and urgent health needs, compromising the quality and safety of care for half the population [1].



FDA = Food and Drug Administration, NIH = National Institutes of Health, GAO = US General Accounting Office, renamed in 2004 to US General Accountability Office, ORWH = Office of Research on Women's Health, AMA = American Medical Association.

Driven by advocacy, scientific inquiry, and policy reform, the FDA reversed this guidance in 1993, issuing new recommendations that formally required the inclusion of women in clinical trials. That same year, Congress passed the NIH Revitalization Act, mandating that NIH-funded research include women and minorities, and be designed to detect sex-based differences in outcomes. This marked a turning point in clinical research, one that continues to shape how we build safer, more effective treatments for everyone.

NIH began rejecting funding for studies that did not comply with the NIH Revitalization Act. The Office of Research on Women's Health (ORWH) was established to monitor compliance, and annual reporting on sex, race, and ethnicity in trials became mandatory. The American Medical Association (AMA) and FDA have since pushed for diversity action plans and post-marketing studies to ensure drugs are safe and effective for underrepresented groups.

"**In 1993, recognizing the impact that excluding women from clinical trials had on health equity, the FDA repealed their 1977 rule and Congress passed a mandate that all studies funded by NIH include women and assess differences amongst sexes**." *American Medical Association*

CURRENT STATE OF REPRESENTATION

Over the past decade, the industry has made meaningful strides to review and support diversity, equality, and inclusion (DEI) initiatives - aiming to foster greater transparency and communication. However, as efforts to analyse clinical trial data have shown, significant gaps remain. The data still falls short of representing real-world patient populations, limiting our ability to deliver truly equitable healthcare solutions.

In February 2021, the FDA published the *2020 Drug Trials Snapshot Summary Report*[2], offering a comprehensive overview of demographic representation in clinical trials for newly approved drugs and biologics.
The initiative offers clear, accessible information about who participated in the clinical trials that supported the approval of new drugs. Designed for both consumers and healthcare professionals, these snapshots aim to promote transparency and accountability in trial design by sharing demographic data (including sex, race, age, and ethnicity) and highlighting where trials were conducted. They also reveal whether different groups experienced varying benefits or side effects, helping to identify gaps and improve equity in clinical research.
It showed that in 2020, over 32,000 patients participated in clinical trials supporting FDA approval of 53 new drugs and biologics. Whilst women made up 56% of participants overall, representation across race and ethnicity remained uneven; 75% were white, only 8% were Black or African American, 6% Asian, and 11% Hispanic. These disparities underscore the ongoing need for inclusive research practices. As the FDA notes, "*We hope this information is helpful to promote dialogue on the appropriate representation of different subgroups in clinical trials*." The data also reveal that therapeutic areas like neurology had 74% female participation, whilst infectious disease trials included only 36% women and just 32% white participants, highlighting variability across conditions.

Building on the FDA's commitment to transparency, the 2021 paper titled *"U.S. Racial and Ethnic Participation in Global Clinical Trials by Therapeutic Areas"*[3] offers a deeper analysis by comparing U.S.-based trial participation to national census data. Despite growing efforts to improve health equity, racial and ethnic representation in U.S.-based clinical trials, inconsistency remains across therapeutic areas. An analysis of over 100,000 U.S. participants from 2015 to 2019 revealed that while Black or African American participation was at or above census estimates in most therapeutic areas (reaching 45% in psychiatry and 22% in infectious diseases) Asian and Hispanic populations were consistently underrepresented. For example, Asian participation ranged from just 0.75% in cardiovascular trials to 4% in dermatology, and Hispanic or Latino representation was below census estimates in most areas, with only 7% in oncology and cardiovascular trials.

TABLE 1 Clinical trials participation at the U.S. Sites by Race (*N* = 102,596)

White	Black or African American	Asian	American Indian or Alaska Native	Other*
80,310 (78.27%)	16,733 (16.31%)	2,139 (2.08%)	531 (0.52%)	2,883 (2.81%)

*combined categories 'Native Hawaiian or Other Pacific Islander', 'Other race', 'Mixed race' and 'Unknown/Unreported/Missing'.

Table Screenshots from [3]*U.S. racial and ethnic participation in global clinical trials by therapeutic areas*.

TABLE 2 Clinical Trials Participation at the U.S. Sites by Ethnicity (*N* = 102,596)

Hispanic or Latino	Not Hispanic or Latino	Missing
15,691 (15.29%)	77,353 (75.39%)	9,552 (9.31%)

TABLE 3 US Census Bureau Estimate* of the Resident Population in 2015 (*N* = 320,635,163)

Demographic Category	U.S. Population
Race	
White	247,382,690 (77.15%)
Black or African American	42,532,491 (13.26)
Asian	17,752,744 (5.53%)
American Indian or Alaska Native	4,004,358 (1.24%)
Ethnicity	
Hispanic or Latino	56,254,742 (17.54%)
Not Hispanic or Latino	129,427,521 (40.36%)

*Adapted from US Census Bureau Population Division, Annual Estimates of the Resident Population.

Table Screenshots from [3] *U.S. racial and ethnic participation in global clinical trials by therapeutic areas*.

The authors conclude that "*some racial and ethnic minorities are still not well represented in most drug development programmes, and this remains an area where improvement is needed*." These findings underscore the importance of designing trials that reflect the diversity of real-world populations to ensure equitable access to safe and effective treatments.
Together, these reports reveal persistent gaps in representation among racial and ethnic minorities and underscore the need for more inclusive trial designs that reflect the diversity of real-world populations.

DEI IS "JUST GOOD SCIENCE"

WHY REPRESENTATION MATTERS

Underrepresentation in clinical trials affects the accuracy, fairness, and usefulness of medical research. When racial and ethnic groups are left out or underrepresented, the data may not show how treatments work across different populations, especially for diseases that impact some communities more than others due to genetic or environmental factors. Including a diverse range of participants helps researchers understand how safe and effective a drug is for everyone.
Underrepresentation also contributes to ongoing health disparities by limiting access to new therapies and excluding communities from the benefits of research. It makes it harder to see who is being served and who is being left behind, which is why transparent and inclusive trial practices are essential. Fixing these gaps isn't just about better science - it's about making healthcare more equitable and trustworthy.

"It is a capital mistake to theorise before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."
A Scandal in Bohemia, Arthur Conan Doyle

PHARMACOGENOMIC DIFFERENCES

Understanding pharmacogenomic differences across racial and ethnic groups is essential for delivering fair and effective healthcare especially in certain treatments, where genetic variation significantly influences drug response.
The HIV study by Ford et al. (2022) [4] underscores this by finding specific genetic variants that were more common in African populations, some of which were linked to known disease risks and only appeared in these groups.
This is crucial because widely used HIV medications like Efavirenz and Nevirapine are metabolized by enzymes such as CYP2A6, CYP2B6, and UGT2B7, which show unique genetic patterns in different populations, influencing how well the drugs work and "*contributes to inter-individual variations in patient responses*."
The authors emphasise that "*many variants identified in this study have also been linked to previously documented variants*" with known disease associations, and some were "*only observed in African populations*," highlighting the need for population-specific data.
The study highlights the need for better, more representative genetic data to guide treatment decisions and dosage adjustments. Ultimately, these findings support the case for routine pharmacogenetic testing and reinforce why DEI in research is essential, so that all populations benefit equally from advances in precision medicine.

HISTORICAL FAILURES

The Tuskegee Syphilis Study

The Tuskegee Syphilis Study, conducted by the U.S. Public Health Service from 1932 to 1972, involved 600 African American men in Tuskegee, Alabama, 399 with syphilis and 201 without. The participants were misled into believing they were receiving free healthcare for “bad blood,” a colloquial term used to describe various ailments, and no informed consent was obtained. Even after penicillin became the standard treatment for syphilis in 1947, the men were intentionally denied access to it so researchers could observe the natural progression of the disease.

As a result, many suffered or died from syphilis-related complications, and the study inflicted lasting harm on families and communities. It stands as a profound violation of medical ethics, marked by deception, exploitation, and systemic racism. The legacy of Tuskegee continues to influence medical mistrust among Black Americans, contributing to hesitancy in participating in clinical trials and engaging with healthcare systems, even when modern safeguards are in place. This history underscores the critical importance of informed consent, transparency, and community engagement in ethical research practices.

Henrietta Lacks – HeLa Cells

Henrietta Lacks’ story remains a powerful example of how systemic racism and exploitation in medicine have contributed to deep mistrust in clinical research, particularly among communities of colour. In 1951, cells were taken from Lacks without her knowledge or consent (later known as HeLa cells) and became one of the most important tools in biomedical research. Her family was neither informed nor compensated for decades. In fact, they did not receive any financial settlement until 2023, when they reached a confidential agreement with Thermo Fisher Scientific after a historic lawsuit. As her son Lawrence Lacks poignantly stated, “If our mother is so important to science, why can’t we get health insurance?.” This quote encapsulates the ethical and racial injustices at the heart of her story, highlighting the disparity between the scientific value of her cells and the lack of benefit to her descendants.

Today, her legacy continues to influence perceptions of clinical trials, especially in marginalised communities, where historical injustices have led to scepticism and hesitancy. In the context of DEI goals, her story reinforces the need for diversity in trial populations to reflect genetic and social variation, equality in access and protection, and inclusion through culturally sensitive outreach and ethical engagement.

DATA INTEGRITY AND DATA QUALITY

The presentation titled "Data Integrity and Data Quality in Application Submissions[5]" by Dr. Minglei Cui (FDA Office of Generic Drugs) underscores how failures in data integrity can significantly undermine the credibility of clinical research and hinder DEI efforts. Data integrity (defined as the completeness, consistency, and accuracy of data) is essential for ensuring that clinical trial results are reliable and representative.

The presentation details multiple case studies where poor documentation, falsification, sample manipulation, and biased data exclusion led to regulatory actions, including study repetition and audits. These breaches not only delay drug approvals but also compromise the inclusion of diverse populations by distorting subgroup analyses and masking safety or efficacy signals relevant to underrepresented groups. As the FDA notes, “lack of data reliability would have a negative impact on the acceptability of data submitted in support of a marketing application,” making robust data governance and early issue reporting critical to both scientific integrity and equitable healthcare outcomes.

HEALTHCARE ACCESS AND ENGAGEMENT

For decades, the estimated glomerular filtration rate (eGFR) (a key measure of kidney function) used a race-based adjustment that falsely inflated kidney function scores for Black patients. This adjustment was based on the flawed assumption that race correlates directly with muscle mass or kidney function[6]. Instead, the focus has now shifted to more accurate, individualised measures that don’t rely on broad racial generalisations. As a result of the old equation, many Black patients were told their kidneys were functioning better than they were, delaying referrals to specialists, eligibility for transplants, and access to life-saving treatments.

In 2021, the National Kidney Foundation and the American Society of Nephrology recommended removing race from the eGFR equation, acknowledging that while race is a social construct, not a biological one, it still has real-world consequences that must be addressed through equitable healthcare practices. The change has led to more accurate assessments and helped thousands of Black patients move up on transplant waiting lists. The removal of race from the eGFR equation is a critical step towards correcting systemic bias in medicine and improving equity in kidney care.

RESEARCH FRAUD

Reboxetine, an antidepressant developed for treating major depressive disorder, has been approved since 1997 in parts of Europe but not in the U.S. due to concerns about its efficacy and safety. The drug became the centre of a major research controversy in 2010 when an investigation[7] by the independent German Institute for Quality and Efficiency in Health Care (IQWiG). The investigation revealed 74% of the trial data was withheld by its manufacturer,

Pfizer. The published studies exaggerated the benefits and downplayed the risks, while the unpublished data showed that Reboxetine was no more effective than placebo and had a higher rate of adverse effects compared to other antidepressants. This case is a prominent example of publication bias and highlights the critical need for transparency in clinical research to ensure that treatment decisions are based on complete and accurate evidence.
Although not directly a DEI issue, this case has significant implications for equity in healthcare. Selective reporting and clinical trial data misrepresentation can result in the prescription of ineffective treatments and can disproportionately affect underserved populations who may have limited access to alternative care or information. Additionally, such practices erode public trust (especially among communities historically marginalised in research) making it harder to engage diverse populations in future studies. The case underscores the importance of transparency, accountability, and inclusive evidence generation as foundational to achieving equity in health research.

DATA INTEGRITY: MISSED SAFETY

The UK's Medicines and Healthcare products Regulatory Agency (MHRA) recently issued a safety alert concerning an increased risk of unintended pregnancies among women using GLP-1 receptor agonist weight loss drugs, including Ozempic, Wegovy, Mounjaro, Saxenda, and Victoza. The MHRA received over 40 pregnancy reports through its Yellow Card scheme, some of which occurred despite the use of birth control. These medications are known to cause gastrointestinal side effects such as vomiting, diarrhoea, and delayed gastric emptying, which may impair the absorption of oral contraceptives and reduce their effectiveness. Additionally, weight loss itself may increase fertility, further raising the risk of unintended conception.
The agency emphasised the lack of safety data on GLP-1 exposure during pregnancy, conception, and breastfeeding, and advised immediate discontinuation of these drugs if pregnancy occurs. To mitigate risk, the MHRA recommends considering non-oral contraceptive options such as intrauterine devices (IUDs) or implants, and continuing contraception for up to two months after stopping Semaglutide and one month after stopping Tirzepatide before attempting to conceive.
Unintended pregnancies among women prescribed GLP-1 receptor agonists have also exposed significant gaps in clinical trial design and patient education. While patients are expected to follow medical advice regarding contraception, recent studies[8] indicate that only 21% of women initiating GLP-1 therapy reported using any form of contraception, and many were unaware of the increased fertility associated with weight loss or the potential for reduced effectiveness of oral contraceptives due to drug-induced gastrointestinal effects.
Further compounding these concerns, research[9] from the University of Amsterdam has shown that GLP-1 receptor agonist exposure during pregnancy in animal studies led to reduced foetal weight and growth, delayed ossification and skeletal variants, and reduced maternal weight gain. These findings underscore the urgent need for more robust reproductive safety data and clearer guidance for women of childbearing potential who are prescribed these medications.
Regulatory bodies have responded with updated guidance, but the initial clinical trials for these medications appear to lack alignment with DEI principles. Critical reproductive health factors such as pregnancy risk, contraceptive efficacy, and fertility changes were insufficiently considered or omitted altogether. This oversight reflects a systemic failure in trial design and risk communication, placing substantial responsibility on pharmaceutical companies and regulators to ensure that safety data is inclusive, comprehensive, and clearly conveyed to all patient populations.

Further compounding safety concerns, a BMJ report [10] published in 2025 revealed that 82 deaths in the UK have been linked to adverse reactions from GLP-1 receptor agonists. These deaths, also reported through the Yellow Card scheme, raised alarms about unsafe prescribing practices (particularly via online platforms) and highlighted the risks associated with limited access to regulated care. Notably, the published data did not include demographic details such as age, sex, or race, making it impossible to assess whether certain populations were disproportionately affected. This lack of transparency is a DEI concern, as it hinders efforts to identify disparities in safety outcomes and ensure that post-market surveillance reflects the experiences of diverse patient groups.
Inclusive data collection and reporting are essential not only for scientific integrity but also for equitable healthcare. These findings reinforce the need for clinical trials and regulatory systems to prioritize diversity in both design and monitoring. GLP-1 receptor agonists may be the "*drugs of the moment*", but they are currently lacking in both inclusive trial design and robust post-market monitoring and currently fall short in ensuring safety and accountability across the full spectrum of patient populations.

SOCIAL FACTORS

Mistrust in Medical Institutions

The Pfizer Trovan clinical trial conducted in Nigeria in 1996[11] is a widely cited example of how unethical research practices can lead to deep mistrust in medical institutions. During a meningitis epidemic, Pfizer tested an experimental antibiotic, trovafloxacin, on 200 children without proper ethical approval or informed consent. Allegations later surfaced that the approval letter was backdated and possibly forged, and that families were unaware their children were part of a clinical trial, believing they were receiving standard treatment instead. The trial resulted in

multiple deaths and long-term disabilities among participants, and Pfizer was accused of administering reduced doses of the standard treatment to skew results in favour of its drug. The lack of transparency, exploitation of vulnerable populations, and disregard for ethical standards contributed to widespread mistrust, particularly in Northern Nigeria, where the incident sparked protests and long-term vaccine hesitancy. This case underscores the critical importance of ethical conduct, community engagement, and culturally sensitive communication in clinical research, especially in regions with historical experiences of exploitation

Socioeconomic Barriers

Social and economic barriers can significantly limit participation in clinical trials, especially among under-represented groups. Financial constraints such as travel costs, unpaid time off work, and lack of childcare support can make participation impractical. Many trials are conducted in urban or academic centers, creating access issues for people in rural or underserved areas. Limited health literacy and awareness, combined with historical mistrust in medical institutions, can further discourage involvement.

Language barriers and the absence of translated materials or interpreters exclude non-English speakers, while the growing reliance on digital tools can disadvantage those without internet access or digital literacy. Additionally, strict eligibility criteria may unintentionally exclude individuals with comorbidities, disabilities, or older adults. Lack of diversity among clinical staff can also contribute to hesitancy in participation; people often feel more comfortable and respected when they see themselves represented in the research environment. It fosters trust, makes participants feel seen and heard, and can improve engagement and retention.

These barriers collectively hinder equitable representation in research and reinforce the need for inclusive trial design and outreach strategies.

DEEP DIVE – WHERE ARE WE NOW?

The systematic review[12] by Martinez et al. (2022) critically examines how race and ethnicity have been conceptualised, measured, and utilised in U.S.-based epidemiologic research published in five leading journals over a 23-year period. Despite decades of calls for improved practices, the field has made limited progress in treating race and ethnicity as meaningful constructs in health research. Of the 1,050 articles reviewed, 414 met inclusion criteria, yet only four studies explicitly defined race and/or ethnicity. Most failed to distinguish between the two, often collapsing them into a single “ethnoracial” construct, undermining analytical clarity and rigour.

Measurement practices were similarly lacking, with over 75% of studies not stating how race or ethnicity was assessed. Coding schemes frequently centred Whiteness (using “White” as the default reference category or grouping others into ambiguous “nonwhite” or “other” categories) reflecting problematic norms that marginalise diverse populations. These practices contribute to the underrepresentation of racial and ethnic minorities and perpetuate “ritualistic inclusion,” where race and ethnicity are included without thoughtful justification. The authors call for greater accountability in scientific publishing, emphasising that existing guidelines are sound but poorly enforced.

“What is not needed is more of the same”

Thomas A. LaVeist, *Why We Should Continue to Study Race... But Do a Better Job: An Essay on Race, Racism and Health*, 1996.

Data mining from platforms like ClinicalTrials.gov is a powerful tool for evaluating DEI in clinical research. ClinicalTrials.gov is a publicly accessible registry and results database of clinical studies conducted globally, maintained by the U.S. National Library of Medicine at the National Institutes of Health (NIH). Its mission is to promote transparency and scientific integrity, offering researchers, clinicians, and the public access to detailed information about ongoing and completed trials to support informed decision-making.

By mining this data, researchers can extract demographic details such as race, ethnicity, gender, and age from thousands of studies. This enables the identification of representation gaps, revealing whether trials are inclusive of diverse populations, including women and older adults.

Analysing metadata over time allows for tracking trends in DEI efforts, highlighting which therapeutic areas (e.g., oncology, cardiology) may be lagging in inclusive recruitment and how different funding sources (e.g., industry vs. academia) correlate with participant diversity. Linking demographic data to trial outcomes also helps assess the relevance of results, uncovering how diversity influences efficacy, safety signals, regulatory approvals, and post-market performance.

ClinicalTrials.gov includes fields for race and ethnicity reporting, which can be used to audit compliance with DEI guidelines (such as those set by the NIH or FDA) and identify trials that fail to report mandated demographic data. Ultimately, leveraging data mining insights can inform DEI policy recommendations, guide funding priorities, and shape community engagement strategies to improve equitable participation in clinical research.

Whilst I attempted to mine data directly from ClinicalTrials.gov to assess demographic representation, it is evident that some pharmaceutical companies are proactively presenting themselves as committed to DEI in clinical research. For instance, Pfizer's analysis[13] of 213 U.S.-based trials (2011–2020) states that "*overall trial participation of Black or African American individuals was at the US census level (14.3% vs 13.4%)*," while Hispanic or Latino participation "*was below US census (15.9% vs 18.5%)*". Pfizer has publicly committed to "*achieving racially and ethnically diverse participation at or above US census or disease prevalence levels (as appropriate) in all of our trials.*"
In contrast, GSK's review[14] of 495 trials (2002–2019) critiques the use of census data as a benchmark, asserting that "*US Census data are an inappropriate universal benchmark*" and advocating for epidemiologic data to better reflect disease burden and health equity. Their HIV trial data, for example, showed Black or African American representation at 35.1%, which exceeded census levels but fell short of the epidemiologic prevalence of 55.3%. GSK concludes that "*benchmarking in line with epidemiologic data will allow us to set better trial enrollment goals*", signaling a shift toward more disease-specific and equitable recruitment strategies.
However, both papers acknowledge limitations in the underlying data, including inconsistent reporting of race and ethnicity across trials and the exclusion of participants who did not self-identify, which may obscure the full picture of representation.

NIH DATA MINING

Within the NIH US National Library of Medicine, the National Center for Biotechnology information maintains the clinicatrials.gov repository. This online database of clinical research lists many studies, but not Protocols, SAPs, CRFs and Clinical Study Reports submitted to the FDA since its inception in 2000[15]. In addition, a selection of historic trials from the 1980s and 1990's has been included in the repository for a range of reasons often related to sponsor led transparency initiatives and ethical drives by intergovernmental trans-national organisations such as the WHO. There are no set criteria for historic trials to be included in the database and the level of detail available for these trials can vary significantly[16].
The clinicatrials.gov repository provides two search interfaces. A web-based query form and a recently updated API with a number of endpoints. This is undoubtedly a valuable resource. However, many researchers and data scientists have highlighted challenges[17] with analysing the available data primarily due to its unstructured form and lack of standard terminologies. Indeed, much of the available data is in free text form and whilst there is, for certain use cases, great utility in having full texts of Protocols, Investigator Brochures and SAP's available, using such for analysis and inference is challenging.
Appendix I lists the 'enumeration types' available – these are the only variables across the database with standard coded values, and which can be used for queries. It is clear from this list that, for a data scientist searching for Trial Documents pertaining to specific trials, based on queries such as Sponsor, Location, Date, Intervention, etc – this tool is perfectly adequate. However, for broader data mining activities, the database leaves much to be desired.

DATA MINING: DIVERSITY AND INCLUSION

To illustrate the limitations in mining data from the repository, it was decided to search for demographic information for clinical trials related to Psychiatric indications submitted between the years of 2000 and 2019, with a view to determining whether the demographics of trial participants had dramatically changed during this time. Our query returned 10431 records (corresponding to 10431 clinical trials submitted for Psychiatric indications during this time-period).
Firstly, given the difficulty of querying the API (given this is not a task that usually sits within the Statistical Programmer sphere of experience), a Jupyter Notebook query tool was constructed – facilitating querying of the API by search string, date range. The tool extracted a list of available data fields (totalling 381), for example, to list three:

- *protocolSection.eligibilityModule.minimumAge,*
- *protocolSection.resultsModule.briefDescription,*
- *resultsSection.outcomeMeasuresModule.outcomeMeasures.analyses.pValue*

Many of these query fields return not a single value of text string but a Python Dictionary object that can be unpacked into sometimes up to eight separate fields. A full query of all data fields yields over 480 variables and over 300 of these contain free text including the full set of submitted clinical trial documentation.

Extracting analysable data from such datasets is challenging. Limited tabulated demographic data was available, none pertaining to race and whilst minimum and maximum ages of trial participants was available in tabulated form, this data was still inconsistently coded. For example, the participant minimum and maximum age data was variously recorded as (to name a few):

- "Xx Years"
- "XX: No older than"
- "Years XX to XX – inclusive"
- "XX" (with values suggesting weeks or months rather than years (e.g. >200)

A brief attempt to construct a Langchain based AI Agent was made, with some success – this being a data extraction task well suited to the latest generation of Large Language Models. However budgetary and time constraints prevented further exploration of this.

REGULATORY CONTEXT

The requirement to embed DEI into clinical research has gained substantial traction, driven by both ethical considerations and the need for scientifically robust data that reflects real-world populations. Regulatory agencies have responded with a suite of guidance documents that underscore the necessity of inclusive trial design, recruitment, and reporting practices. From the FDA's foundational 2020 guidance on enhancing diversity in clinical trials to the more recent directives on Race and Ethnicity Diversity Plans and standardised demographic data collection, these documents collectively aim to improve the application, safety, and efficacy of medical products. In the UK, complementary strategies such as the National Institute for Health and Care Research (NIHR) Inclusion Strategy and the Health Research Authority's (HRA) Inclusion and Diversity Plan further reinforce the commitment to equitable research practices. Together, these regulatory frameworks provide essential tools for sponsors and investigators to operationalize DEI principles, ensuring that clinical trials are not only scientifically sound but also socially responsible and representative of the populations they intend to serve.

"Enhancing the Diversity of Clinical Trial Populations, Eligibility Criteria, Enrolment Practices, and Trial Designs - Guidance for Industry[18]", published November 2020.

The FDA's 2020 guidance on enhancing diversity in clinical trials emphasises the importance of broadening eligibility criteria and adopting inclusive enrolment practices to ensure trial populations better reflect those who will use the drug post-approval. It recommends minimising unnecessary exclusions, incorporating diverse demographic and clinical characteristics, and using adaptive and enrichment designs for trials. The guidance also highlights strategies to reduce participant burden, improve outreach and retention (especially among underrepresented groups) and encourages the use of real-world data and community engagement. Special considerations are provided for trials involving rare diseases, with a focus on maximising participation and ensuring representativeness. These efforts aim to improve the applicability, safety, and efficacy assessments of investigational drugs.

"Collection of Race and Ethnicity Data in Clinical Trials Guidance for Industry and Food and Drug Administration Staff[19]", Issued on 26 October 2016, updated January 2025.

The FDA's guidance outlines a standardised approach for collecting and reporting race and ethnicity data in clinical trials to improve the quality, consistency, and transparency of demographic subgroup analyses. It emphasises that race and ethnicity categories are social-political constructs, not scientific classifications, and should be self-reported by participants using a two-question format - first asking about ethnicity (Hispanic/Latino or not), followed by race (with a minimum of five categories: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White).

The guidance supports the FDA Safety and Innovation Act (FDASIA) Section 907, which calls for improved inclusion and analysis of demographic subgroups in clinical research. It also highlights the importance of collecting data that reflect the diversity of the population expected to use the medical product, noting that "*differences in response to medical products have already been observed in racially and ethnically distinct subgroups of the U.S. population*." The FDA encourages sponsors to submit plans early in development to ensure adequate representation and warns that inadequate subgroup data may result in insufficient labeling information or even refusal to file an application.

“**Diversity Plans to Improve Enrollment of Participants from Underrepresented Racial and Ethnic Populations in Clinical Trials Guidance for Industry[20]” - Draft Guidance, published in the Federal Register on April 14, 2022.**

The FDA’s draft guidance emphasises the critical need for sponsors to develop and submit “Race and Ethnicity Diversity Plans” to improve the enrolment of underrepresented populations in clinical trials. These plans aim to ensure that trial data reflects the diversity of the population expected to use the medical product, therefore enhancing the applicability and equity of clinical research outcomes.

The guidance acknowledges that individuals from racial and ethnic minorities often bear a disproportionate disease burden yet remain underrepresented in biomedical research. It also highlights broader issues such as differential access to healthcare and historical mistrust, noting that “*mistrust of the clinical research system may stem from historical events that adversely impacted racial and ethnic minorities, such as the unethical Tuskegee experiments*.” Sponsors are encouraged to go beyond race and ethnicity, considering other underrepresented groups defined by sex, gender identity, age, disability, and socioeconomic status.

The FDA recommends early and proactive planning, stating that “*consistent implementation of actions to improve racial and ethnic diversity in clinical trials… can support early access to medical discoveries and innovations*.” Operational strategies include community engagement, language access, and financial support for participants, with the goal of fostering inclusive and equitable research practices.

“NIHR Research Inclusion Strategy 2022 – 2027[21]” - Issued 26 September 2022, Version: 3.0 - July 2025

The NIHR Research Inclusion Strategy 2022 - 2027, launched in September 2022 by the National Institute for Health and Care Research (NIHR), outlines a strategic commitment to embedding inclusive practices across the UK’s health and care research system. The strategy defines research inclusion broadly, addressing inequalities not only related to protected characteristics under the Equality Act 2010, but also those linked to socioeconomic status, geography, and access to care.

By promoting inclusive study design, tackling structural barriers, and investing in infrastructure and workforce development, the strategy aims to ensure that research is representative, accessible, and relevant to all communities. This initiative significantly advances DEI by fostering diverse participation, tailoring support to meet different needs, and making research a routine and equitable part of health and care experiences.

"Guidance for developing and submitting an Inclusion and Diversity Plan – second draft[22]", **published 13 May 2025**

The Health Research Authority (HRA), in partnership with the Medicines and Healthcare products Regulatory Agency (MHRA), released the second draft of its guidance on developing and submitting an Inclusion and Diversity Plan (IDP) on 13 May 2025. This guidance is part of a broader initiative to improve the inclusivity of clinical trials in the UK by encouraging researchers and sponsors to proactively consider how their studies can better reflect the diversity of the population. Although submission of an IDP is not mandatory, it is strongly recommended for trials involving investigational medicinal products, medical devices, or novel interventions.

The guidance aims to ensure that clinical research includes individuals who may be impacted by the findings, particularly those from under-served or under-represented communities. By fostering more inclusive recruitment and study design, the IDP supports the principles of DEI, enhances the relevance and safety of research outcomes, and promotes ethical and socially responsible research practices.

PATH FORWARD – WHAT CAN WE DO?

In early 2025, the FDA quietly removed its draft guidance on enhancing diversity in clinical trials[23], following an executive Presidential order that curtailed federal DEI initiatives. The guidance had outlined requirements for Diversity Action Plans (DAPs), aiming to improve representation of underrepresented populations in clinical research. Its removal has created uncertainty around compliance with the Food and Drug Omnibus Reform Act (FDORA), which mandates such guidance by mid-2025. This regulatory shift raises concerns about the future of inclusive research practices and their impact on equitable healthcare outcomes.

While the FDA's withdrawal of DEI guidance in 2025 was unexpected, frustrating and disappointing, the pursuit of diversity in clinical trials remains a scientific imperative. Regardless of shifting political landscapes, efforts to embed equity and inclusion in clinical research must continue through intentional design, reporting, and advocacy.

EMBEDDING DIVERSITY IN RESEARCH LEADERSHIP

Diversity within research teams, especially at the leadership level, is essential for driving meaningful DEI initiatives in clinical trials. Diverse leadership fosters broader perspectives and more representative decision-making, helping to identify and address blind spots in trial design and execution. Additionally, teams that reflect the communities they serve are more likely to build trust and credibility, reinforcing the principle that people are more inclined to trust leaders who resemble them. Despite longstanding gaps in workforce diversity, it is encouraging to see institutions implementing targeted programmes to increase investigator diversity and launching broader initiatives to raise awareness and identify areas for improvement.

UPDATE MANDATES AND DOCUMENTATION

Current diversity planning and demographic reporting in clinical trials (as found in "Collection of Race and Ethnicity Data in Clinical Trials Guidance for Industry and Food and Drug Administration Staff" [19]) often rely on a narrow set of self-identified race and ethnicity categories that are shaped by socio-geographic constructs rather than scientific constructs. This approach can oversimplify complex identities and limit the accuracy of data used to assess representation. To address this, more nuanced frameworks are needed, such as incorporating intersectional data (e.g., race combined with socioeconomic status, gender identity, or geographic location), using granular ethnicity classifications (e.g., distinguishing between East Asian, South Asian, and Southeast Asian populations), and applying genomic ancestry markers where appropriate to complement self-reported data. These approaches can improve the precision and relevance of clinical trial findings while promoting more equitable research practices.

EARLIER DIAGNOSIS AND RESEACH INITIATIVES

Earlier diagnosis of disease also plays a vital role in improving DEI in clinical trials. When individuals (especially those from underserved or underrepresented communities) are diagnosed sooner, they are more likely to be identified and recruited for research opportunities. This helps reduce barriers to participation caused by delayed access to care, systemic bias, or lack of awareness. Earlier diagnosis enables trials to better reflect the real-world patient population, leading to more inclusive study designs and more generally applicable results. By ensuring that diverse groups are represented from the outset, clinical research becomes more equitable, more relevant, and ultimately more impactful for all communities.

The CLOCS project (Cancer Loyalty Card Study[24]), led by Imperial College London, is an innovative research initiative exploring whether everyday shopping data can help detect ovarian cancer earlier. By analysing loyalty card records from major UK retailers, researchers identified that women who were later diagnosed with ovarian cancer had purchased over-the-counter medications, such as pain relief and indigestion remedies, up to eight months before their diagnosis. This suggests they were self-managing symptoms like bloating and abdominal discomfort long before seeking medical help. The study found a 3.5-month gap between GP visits and diagnosis, highlighting a critical window for earlier intervention. Importantly, over half of participants were willing to share their data for health research, showing potential for broader application. The project emphasises the value of real-world behavioural data in identifying early symptom patterns and supports the expansion of this approach to other cancers with vague early symptoms, such as pancreatic and colorectal cancer. These findings reinforce the importance of inclusive data strategies and public engagement in advancing early detection and equitable healthcare.

EMBEDDING EQUALITY: INCLUDE ETHINICITY FRAMEWORK

The INCLUDE Ethnicity Framework[25], developed through a multi-stakeholder process led by Trial Forge and supported by the UK's NIHR, addresses a critical gap in clinical research: the underrepresentation of ethnic minority groups in trials. Despite comprising a significant portion of the UK population and often bearing a disproportionate burden of disease, these groups are frequently excluded due to systemic barriers, narrow eligibility criteria, and assumptions about recruitment feasibility.

At its core, the Framework encourages trial designers to move beyond participant numbers and consider who is included and why. It comprises four reflective questions designed to guide inclusive trial design:

1. Who should my trial apply to?
2. Are the groups identified likely to respond in different ways?
3. Will my study intervention make it harder for some groups to engage?
4. Will the way I have designed the study make it harder for some groups to engage?

These questions are supported by worksheets that prompt trial teams to consider factors such as disease prevalence, cultural norms, language, and logistical barriers. The Framework is applicable across all disease areas and stages of trial development, making it a versatile tool for embedding equity into research practice.

The development process involved over 40 contributors, including patient and public representatives, clinicians, funders, and academics. It was informed by real-world application to six UK trials, including COVID-19 and type 2 diabetes studies, which helped identify practical challenges and refine the tool.

From a DEI perspective, the Framework is a significant advancement. It acknowledges that exclusion from trials not only limits scientific validity but also perpetuates health disparities. By encouraging trialists to consider cultural relevance, trust in healthcare systems, and accessibility, the Framework promotes inclusive research that is both ethically and scientifically sound.

However, the authors caution against misuse. The Framework is not intended to enforce quotas but to foster thoughtful inclusion. It may increase workload and costs, but these are justified by the long-term benefits of more representative and impactful research. Funders play a crucial role in supporting this shift by recognising inclusion efforts in grant evaluations and providing resources for community engagement.

In conclusion, the INCLUDE Ethnicity Framework offers a practical, reflective approach to designing trials that better serve diverse communities. It is a model for how DEI principles can be used in clinical research, ensuring that the benefits of innovation reach all segments of society.

THE ACCESSIBILITY AND USABILITY OF PUBLIC CLINICAL TRIAL DATASETS

As discussed above, all the data needed for our DEI study was available on the clinicaltrials.gov repository however the data is largely available in unstructured form. The repository is undoubtedly a valuable resource, especially suited perhaps to the extraction and study of trial documents and statistical outcomes of individual trials – in the way one may check a book out of a library, the repository allows users to search for, find and extract information about individual trials.

However, this information is in a form that is accessible for humans but not suitable for data mining. The extraction of structured, tabulated data from unstructured data such as text and images, has been a challenge for machine learning for a long time. Whilst advances in Convolutional Neural Networks in 2012 opened up image data to analysis. More recent advances in Natural Language Processing (notably the development of the Transformer Neural network in 2017) has opened the door to the statistical analysis of textual data. Application of such AI techniques would stand to make the clinicaltrials.gov repository a truly accessible store of valuable information to advance medical research.

CONCLUSION

More than three decades after the NIH Revitalization Act mandated the inclusion of women and minorities in clinical trials, the journey toward equitable representation remains incomplete. Despite regulatory progress and growing awareness, underrepresentation persists, particularly in industry-sponsored studies, limiting the generalisability and fairness of medical research.

Historical injustices like Tuskegee and the exploitation of Henrietta Lacks continue to cast long shadows, fuelling mistrust and disengagement among marginalised communities. Structural barriers such as restrictive eligibility criteria, socioeconomic challenges, and lack of culturally sensitive outreach further hinder participation. Meanwhile, inconsistent demographic reporting and selective data practices compromise transparency and accountability. These gaps are not just ethical failures; they are scientific ones.

To truly advance equitable healthcare, we must embed DEI principles into every stage of clinical research: from trial design and recruitment to data analysis and leadership representation. This means investing in inclusive frameworks, enforcing meaningful mandates, and elevating diverse voices within research teams. It also means recognising that diversity is not a checkbox, it is the foundation of good science.

If our trials do not reflect the populations they aim to serve, we risk perpetuating disparities rather than healing them. The path forward demands courage, collaboration, and commitment. Because in the end, inclusive research isn't just better science, it's better care.

REFERENCES

[1]*Women in Clinical Trials of New Drugs — A Change in Food and Drug Administration Policy*. Merkatz et al., 1993, New England Journal of Medicine, 329(4), 292–296.

[2]*2020 Drug Trial Snapshot Summary Report. FDA.Gov*, 2021, https://www.fda.gov/drugs/drug-approvals-and-databases/drug-trials-snapshots.

[3]*U.S. racial and ethnic participation in global clinical trials by therapeutic areas*, Lolic M, Araojo R, Okeke M, Temple R. Clin Pharm Ther. 2021; 46:1576–1581.

[4]*Pharmacogenetics of CYP2A6, CYP2B6, and UGT2B7 in the Context of HIV Treatments in African Populations.* Ford, G.R.; Niehaus, A.; Joubert, F.; Pepper M.S. J. Pers. Med. 2022, 12,2013.

[5]*Data Integrity and Data Quality in Application Submissions* by Dr. Minglei Cui (FDA Office of Generic Drugs), 28 April 2022, https://www.fda.gov/media/165536/download.

[6]*Unifying Approach for GFR Estimation: Recommendations of the NKF-ASN Task Force on Reassessing the Inclusion of Race in Diagnosing Kidney Disease,* Delgato et al., AJKD, Volume 79, Issue 2, p268-288.e1, February 2022.

[7]*Finding studies on reboxetine: a tale of hide and seek,* Wieseler et al., BMJ 2010;341:c4942

[8]*Incidence of GLP-1 receptor agonist use by women of reproductive age attending general practices in Australia, 2011–2022: a retrospective open cohort study,* Thapaliya et al., MJA 223 (7), 6 October 2025.

[9]*Effects of GLP-1 agonists and SGLT2 inhibitors during pregnancy and lactation on offspring outcomes: a systematic review of the evidence*. Muller DR, Stenvers DJ, Malekzadeh A, et al. Front Endocrinol (Lausanne) 2023; 14: 215356

[10]*GLP-1 agonists: 82 deaths linked to adverse reactions, UK data show*. BMJ 2025;388:r390.

[11]*Pfizer accused of testing new drug without ethical approval,* BMJ 2001;322:194.

[12]*Conceptualization, Operationalization, and Utilization of Race and Ethnicity in Major Epidemiology Journals, 1995–2018: A Systematic Review,* Matrinaz et al., *Am J Epidemiol.,* 2022;00(00):1–14

[13]*Demographic diversity of participants in Pfizer sponsored clinical trials in the United States.* Rottas et al., 2021, Contemporary Clinical Trials 106, 106421.

[14]*Demographic diversity of US-based participants in GSK-sponsored interventional clinical trials,* Reid et al., Clinical Trials, 2023, Vol. 20(2) 133–144.

[15]*Issues in the registration of clinical trials.* Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DA, JAMA. 2007 May 16;297(19):2112-https://lhncbc.nlm.nih.gov/LHC-publications/PDF/pub2007017.pdf

[16]*Historical Highlights of ClinicalTrials.gov.* Boston University Medical Campus. 2019. https://www.bumc.bu.edu/irb/files/2019/11/Historical-Highlights-of-ClinicalTrials.pdf

[17]*Obstacles to the reuse of study metadata in ClinicalTrials.gov.* Miron et al, (2020) 7:443 | https://doi.org/10.1038/s41597-020-00780-z

[18]*Enhancing the Diversity of Clinical Trial Populations — Eligibility Criteria, Enrollment Practices, and Trial Designs Guidance for Industry*, U.S. Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), November 2020, Clinical/Medical.

[19]*Collection of Race and Ethnicity Data in Clinical Trials Guidance for Industry and Food and Drug Administration Staff*, FDA, 26 October 2016 (updated January 2025).

[20]*Diversity Plans to Improve Enrolment of Participants from Underrepresented Racial and Ethnic Populations in Clinical Trials Guidance for Industry*: Draft Guidance, FDA, 14 April 2022.

[21] *NIHR Research Inclusion Strategy 2022 – 2027*, NIHR, Version: 3.0 - July 2025**,** https://www.nihr.ac.uk/about-us/who-we-are/research-inclusion/strategy-2022-27

[22]*Guidance for developing and submitting an inclusion and diversity plan second draft*, Health Research Authority [MHRA], May 2025, https://www.hra.nhs.uk/about-us/news-updates/guidance-developing-and-submitting-inclusion-and-diversity-plan-second-draft/

[23]*FDA Quietly Removes Draft Guidance on Diversity in Clinical Trials Following Executive Order on DEI*. AJMC 2025, https://www.ajmc.com/view/fda-quietly-removes-draft-guidance-on-diversity-in-clinical-trials-following-executive-order-on-dei.

[24]*Cancer Loyalty Card Study (CLOCS): feasibility outcomes for an observational case–control study focusing on the patient interval in ovarian cancer,* Brewer et al., June 2023 BMJ Open, Volume 13, Issue 6, Article ID e066022.

[25]*Developing the INCLUDE Ethnicity Framework - a tool to help trialists design trials that better reflect the communities they serve,* Treweek et al., Trials (2021) 22:337.

ACKNOWLEDGMENTS

My sincere appreciation to the many, many individuals who contributed to the development of this paper. I am particularly grateful to the members of Veramed's Technical Solutions Group (TSG), Bruce Chambers, Kevin Burges, and Christopher Brown for their thoughtful reviews and steadfast support throughout the process.
I extend special thanks to Kemi Oyeneyin for her expert advice and guidance, and to Tracey Wainwright for her valuable insights into regulatory medical writing. I also would like to thank Tamsin Gammie for their constructive feedback during the review cycles.
I am deeply thankful to my co-author, Chris Ward, for his unwavering support and dedication to this project. His insightful ideas and determination were encouraging during our conversations and research.
Finally, thank you to Chris and Nathan, for their enduring encouragement and support, which have been instrumental throughout this journey.

RECOMMENDED READING

The Immortal Life of Henrietta Lacks by Rebecca Skloot (2010).
Invisible Women by Caroline Criado Perez (2020).
Bad Pharma by Ben Goldacre (2013).

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Roxanne McGuire
Veramed
5th Floor Regal House
Twickenham Town
Twickenham
LONDON
TW1 3QS
UK
(+44) 20 3696 7240
Roxanne.mcguire@veramed.com
www.veramed.co.uk

Chris Ward
Veramed
5th Floor Regal House
Twickenham Town
Twickenham
LONDON
TW1 3QS
UK
(+44) 20 3696 7240
chris.ward@veramed.com
www.veramed.co.uk

APPENDIX I

Enumeration types	Allowable Values
Status	ACTIVE_NOT_RECRUITING
	COMPLETED
	ENROLLING_BY_INVITATION
	NOT_YET_RECRUITING
	RECRUITING
	SUSPENDED
	TERMINATED
	WITHDRAWN
	AVAILABLE
	NO_LONGER_AVAILABLE
	TEMPORARILY_NOT_AVAILABLE
	APPROVED_FOR_MARKETING
	WITHHELD
	UNKNOWN
StudyType	EXPANDED_ACCESS
	INTERVENTIONAL
	OBSERVATIONAL
Phase	NA
	EARLY_PHASE
	PHASE1
	PHASE2
	PHASE3
	PHASE4
Sex	FEMALE
	MALE
	ALL
StandardAge	CHILD
	ADULT
	OLDER_ADULT
SamplingMethod	PROBABILITY_SAMPLE
	NON_PROBABILITY_SAMPLE
IpdSharing	YES
	NO
	UNDECIDED
IpdSharingInfoType	STUDY_PROTOCOL
	SAP
	ICF
	CSR
	ANALYTIC_CODE

OrgStudyIdType	NIH
	FDA
	VA
	CDC
	AHRQ
	SAMHSA
SecondaryIdType	NIH
	FDA
	VA
	CDC
	AHRQ
	SAMHSA
	OTHER_GRANT
	EUDRACT_NUMBER
	CTIS
	REGISTRY
	OTHER
AgencyClass	NIH
	FED
	OTHER_GOV
	INDIV
	INDUSTRY - INDUSTRY
	NETWORK - NETWORK
	AMBIG - AMBIG
	OTHER - OTHER
	UNKNOWN - UNKNOWN
ExpandedAccessStatus	AVAILABLE - Available
	NO_LONGER_AVAILABLE - No longer available
	TEMPORARILY_NOT_AVAILABLE - Temporarily not available
	APPROVED_FOR_MARKETING - Approved for marketing
DateType	ACTUAL - Actual
	ESTIMATED - Estimated
ResponsiblePartyType	SPONSOR - Sponsor
	PRINCIPAL_INVESTIGATOR - Principal Investigator
	SPONSOR_INVESTIGATOR - Sponsor-Investigator
DesignAllocation	RANDOMIZED - Randomized
	NON_RANDOMIZED - Non-Randomized
	NA - N/A
InterventionalAssignment	SINGLE_GROUP - Single Group Assignment
	PARALLEL - Parallel Assignment
	CROSSOVER - Crossover Assignment
	FACTORIAL - Factorial Assignment

	SEQUENTIAL - Sequential Assignment
PrimaryPurpose	TREATMENT - Treatment
	PREVENTION - Prevention
	DIAGNOSTIC - Diagnostic
	ECT - Educational/Counseling/Training
	SUPPORTIVE_CARE - Supportive Care
	SCREENING - Screening
	HEALTH_SERVICES_RESEARCH - Health Services Research
	BASIC_SCIENCE - Basic Science
	DEVICE_FEASIBILITY - Device Feasibility
	OTHER - Other
ObservationalModel	COHORT - Cohort
	CASE_CONTROL - Case-Control
	CASE_ONLY - Case-Only
	CASE_CROSSOVER - Case-Crossover
	ECOLOGIC_OR_COMMUNITY - Ecologic or Community
	FAMILY_BASED - Family-Based
	DEFINED_POPULATION - Defined Population
	NATURAL_HISTORY - Natural History
	OTHER - Other
DesignTimePerspective	RETROSPECTIVE - Retrospective
	PROSPECTIVE - Prospective
	CROSS_SECTIONAL - Cross-Sectional
	OTHER - Other
BioSpecRetention	NONE_RETAINED - None Retained
	SAMPLES_WITH_DNA - Samples With DNA
	SAMPLES_WITHOUT_DNA - Samples Without DNA
EnrollmentType	ACTUAL - Actual
	ESTIMATED - Estimated
ArmGroupType	EXPERIMENTAL - Experimental
	ACTIVE_COMPARATOR - Active Comparator
	PLACEBO_COMPARATOR - Placebo Comparator
	SHAM_COMPARATOR - Sham Comparator
	NO_INTERVENTION - No Intervention
	OTHER - Other
InterventionType	BEHAVIORAL
	BIOLOGICA
	COMBINATION_PRODUCT
	DEVICE
	DIAGNOSTIC_TEST
	DIETARY_SUPPLEMENT
	DRU

	GENETIC
	PROCEDURE
	RADIATION
	OTHER
ContactRole	STUDY_CHAIR
	STUDY_DIRECTOR
	PRINCIPAL_INVESTIGATOR
	SUB_INVESTIGATOR
	CONTACT
OfficialRole	STUDY_CHAIR
	STUDY_DIRECTOR
	PRINCIPAL_INVESTIGATOR
	SUB_INVESTIGATOR
RecruitmentStatus	ACTIVE_NOT_RECRUITING
	COMPLETED
	ENROLLING_BY_INVITATION
	NOT_YET_RECRUITING
	RECRUITING
	SUSPENDED
	TERMINATED
	WITHDRAWN
	AVAILABLE
ReferenceType	BACKGROUND
	RESULT
	DERIVED
MeasureParam	GEOMETRIC_MEAN
	GEOMETRIC_LEAST_SQUARES_MEAN
	LEAST_SQUARES_MEAN
	LOG_MEAN
	MEAN
	MEDIAN
	NUMBER
	COUNT_OF_PARTICIPANTS
	COUNT_OF_UNITS
MeasureDispersionType	NA
	STANDARD_DEVIATION
	STANDARD_ERROR
	INTER_QUARTILE_RANGE
	FULL_RANGE
	CONFIDENCE_80 - 80% Confidence Interval
	CONFIDENCE_90 - 90% Confidence Interval
	CONFIDENCE_95 - 95% Confidence Interval
	CONFIDENCE_975 - 97.5% Confidence Interval

	CONFIDENCE_99 - 99% Confidence Interval
	CONFIDENCE_OTHER - Other Confidence Interval Level
	GEOMETRIC_COEFFICIENT
OutcomeMeasureType	PRIMARY - Primary
	SECONDARY - Secondary
	OTHER_PRE_SPECIFIED - Other Pre-specified
	POST_HOC - Post-Hoc
ReportingStatus	NOT_POSTED - Not Posted
	POSTED - Posted
EventAssessment	NON_SYSTEMATIC_ASSESSMENT - Non-systematic Assessment
	SYSTEMATIC_ASSESSMENT - Systematic Assessment
AgreementRestrictionType	LTE60 - LTE60
	GT60 - GT60
	OTHER - OTHER
BrowseLeafRelevance	LOW - low
	HIGH - high
DesignMasking	NONE - None (Open Label)
	SINGLE - Single
	DOUBLE - Double
	TRIPLE - Triple
	QUADRUPLE - Quadruple
WhoMasked	PARTICIPANT - Participant
	CARE_PROVIDER - Care Provider
	INVESTIGATOR - Investigator
	OUTCOMES_ASSESSOR - Outcomes Assessor
AnalysisDispersionType	STANDARD_DEVIATION - Standard Deviation
	STANDARD_ERROR_OF_MEAN - Standard Error of the Mean
ConfidenceIntervalNumSides	ONE_SIDED - 1-Sided
	TWO_SIDED - 2-Sided
NonInferiorityType	SUPERIORITY - Superiority
	NON_INFERIORITY - Non-Inferiority
	EQUIVALENCE - Equivalence
	OTHER - Other
	NON_INFERIORITY_OR_EQUIVALENCE - Non-Inferiority or Equivalence
	SUPERIORITY_OR_OTHER - Superiority or Other
	NON_INFERIORITY_OR_EQUIVALENCE_LEGACY - Non-Inferiority or Equivalence (legacy)
	SUPERIORITY_OR_OTHER_LEGACY - Superiority or Other (legacy)
UnpostedEventType	RESET - Reset
	RELEASE - Release
	UNRELEASE - Unrelease

ViolationEventType	VIOLATION_IDENTIFIED - Violation Identified by FDA
	CORRECTION_CONFIRMED - Correction Confirmed by FDA
	PENALTY_IMPOSED - Penalty Imposed by FDA
	ISSUES_IN_LETTER_ADDRESSED_CONFIRMED - Issues in letter addressed; confirmed by FD