Diversity, Inclusion, and Equality – It's Just Good Science!

Roxanne McGuire and Chris Ward, Veramed
PHUSE EU Connect, 19-Nov-2025

Veramed

Agenda

01 **Introduction**
Why Equality Matters To Us

02 **Representation**
The Current State

03 **DEI**
Is “Just Good Science”

04 **Deep Dive**
Where Are We Now?

05 **Regulatory Context**

06 **Path Forward**
What Can We Do?

07 **Conclusion**

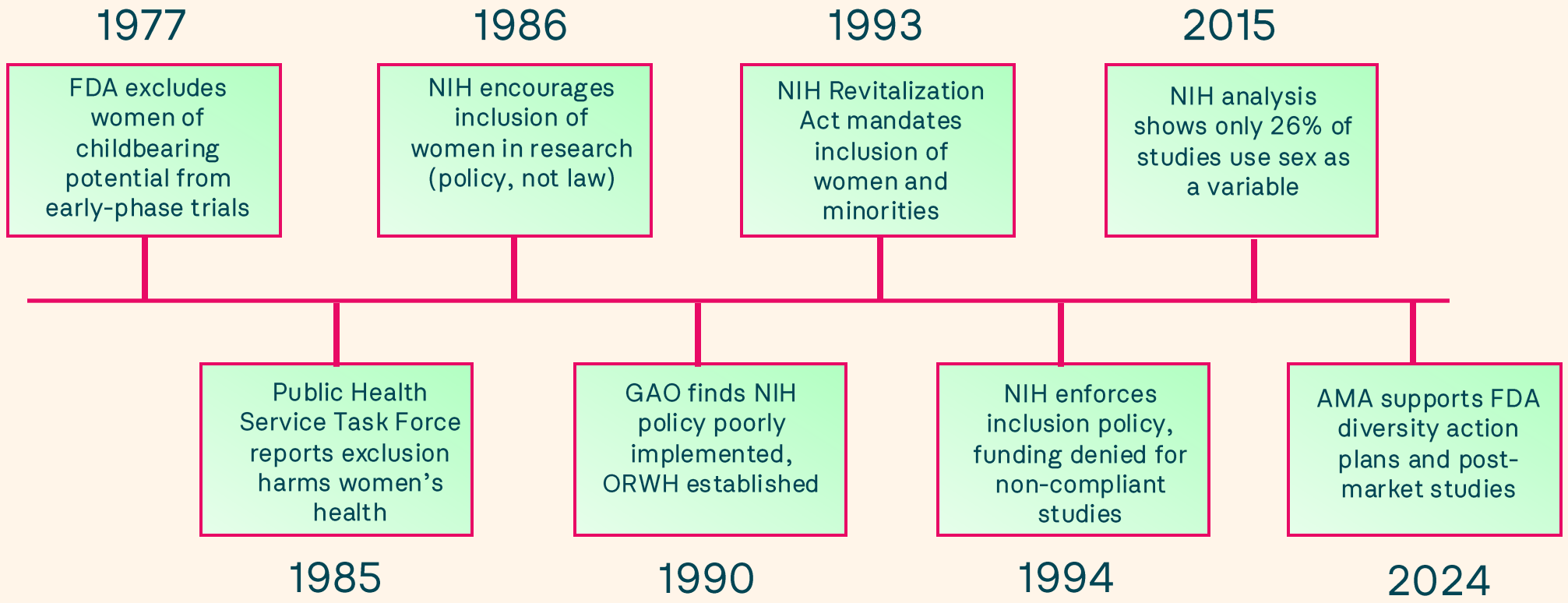
08 **Acknowledgements**

09 **Questions?**

Introduction

Why Equality Matters To Us.

Introduction
Why Equality Matters To Us.



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.

Current State of Representation

Current State of Representation

In 2020, CDER approved 53 novel drugs

Overall, 32,000 patients participated

	WOMEN	WHITE	BLACK or AFRICAN AMERICAN	ASIAN	HISPANIC	AGE 65 AND OLDER	UNITED STATES
AVERAGE	56%	75%	8%	6%	11%	30%	54%

* The percentage of all other races combined (American Indian or Alaska Native, Native Hawaiian or other Pacific islander, Other, Unknown/Unreported) makes up to 100% of race category.
* The percentage of Non-Hispanic and Unknown/Unreported ethnicity makes up to 100% of ethnicity category.
* The percentage of patients from anywhere else in the world makes up to 100% of geographic category.

Percent Participation in Clinical Trials by Subpopulation* for New Molecular Entities and Therapeutic Biologics Approved in 2020

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

Current State of Representation

Review of DTS data published between 2015 and 2019 focusing on the participation at the U.S. trial sites.
Demographic trial data are compared to data obtained from the U.S. Census Bureau to understand the extent to which such trials represent the diversity of the U.S. population.

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 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.

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.

Current State of Representation

Review of DTS data published between 2015 and 2019 focusing on the participation at the U.S. trial sites.
Demographic trial data are compared to data obtained from the U.S. Census Bureau to understand the extent to which such trials represent the diversity of the U.S. population.

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%)	16,738 (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 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.

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.

Current State of Representation

Review of DTS data published between 2015 and 2019 focusing on the participation at the U.S. trial sites.
Demographic trial data are compared to data obtained from the U.S. Census Bureau to understand the extent to which such trials represent the diversity of the U.S. population.

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 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.

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.

Current State of Representation

Review of DTS data published between 2015 and 2019 focusing on the participation at the U.S. trial sites.
Demographic trial data are compared to data obtained from the U.S. Census Bureau to understand the extent to which such trials represent the diversity of the U.S. population.

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 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.

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.

Current State of Representation

Review of DTS data published between 2015 and 2019 focusing on the participation at the U.S. trial sites.
Demographic trial data are compared to data obtained from the U.S. Census Bureau to understand the extent to which such trials represent the diversity of the U.S. population.

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 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.

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.

Current State of Representation

Review of DTS data published between 2015 and 2019 focusing on the participation at the U.S. trial sites.
Demographic trial data are compared to data obtained from the U.S. Census Bureau to understand the extent to which such trials represent the diversity of the U.S. population.

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 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.

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.

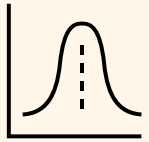
DEI

Is “Just Good Science”

Why Representation Matters

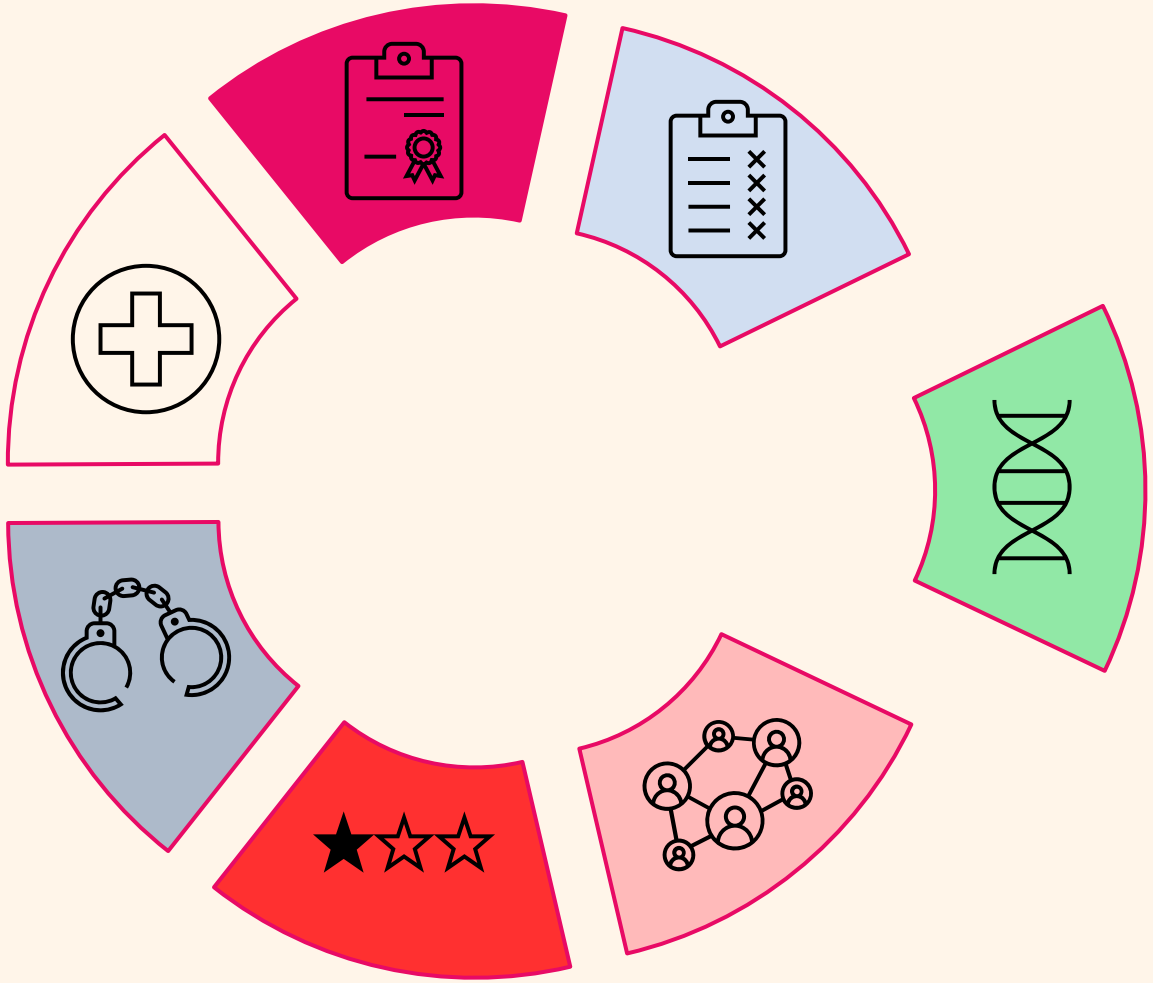
Affects accuracy, fairness and usefulness of medical research

Data may not show how treatments work across different populations



Diverse participants help researchers to understand drug safety and efficacy for everyone

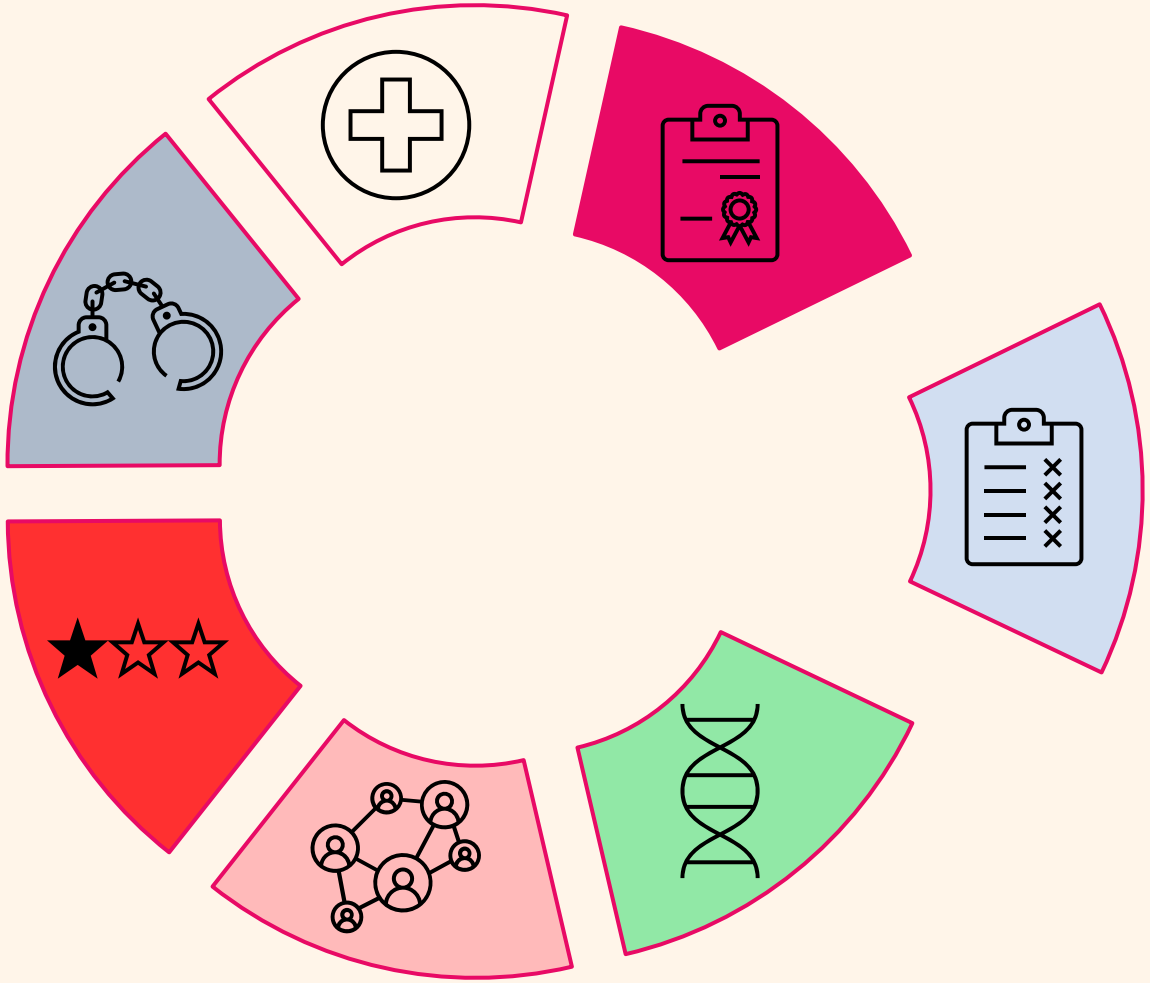
DEI – It's Just Good Science



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

DEI – It's Just Good Science

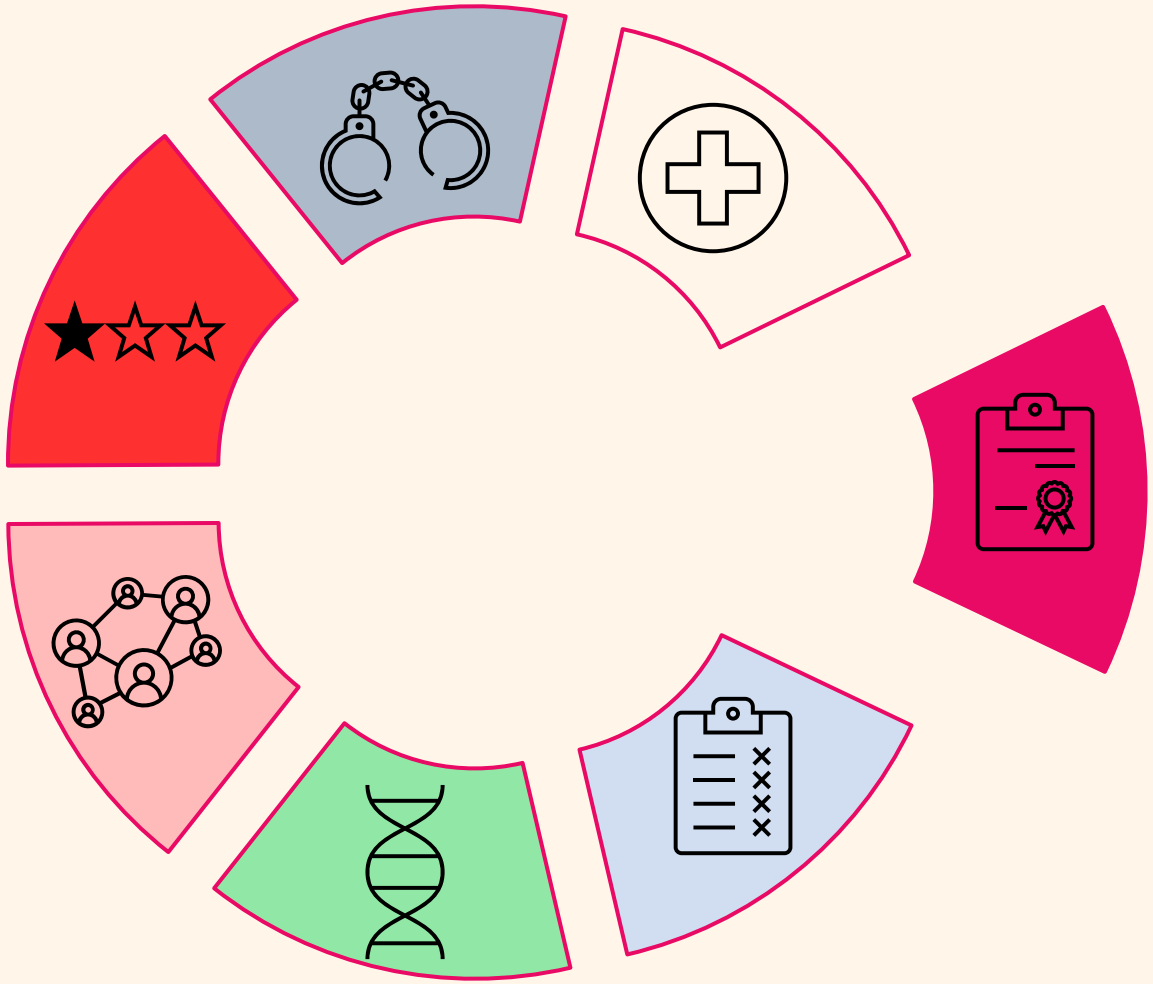


Historical Failures

The Tuskegee Syphilis Studies and Henrietta Lacks

https://historycollection.com/20-photos-tuskegee-syphilis-study/
https://www.sciencealert.com/henrietta-lacks-family-reach-a-settlement-over-her-stolen-immortal-cell-line

DEI – It's Just Good Science

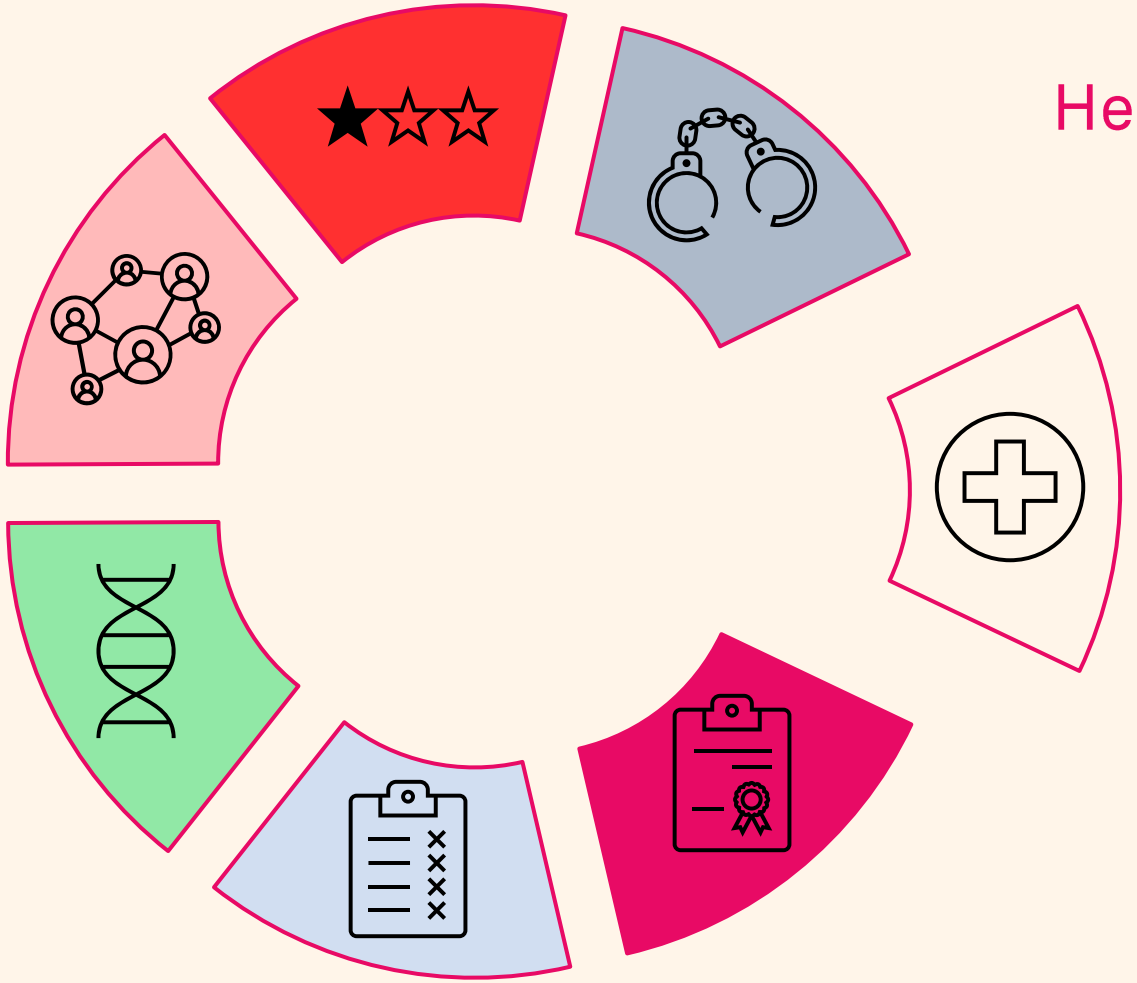


Data Integrity and Data Quality

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.

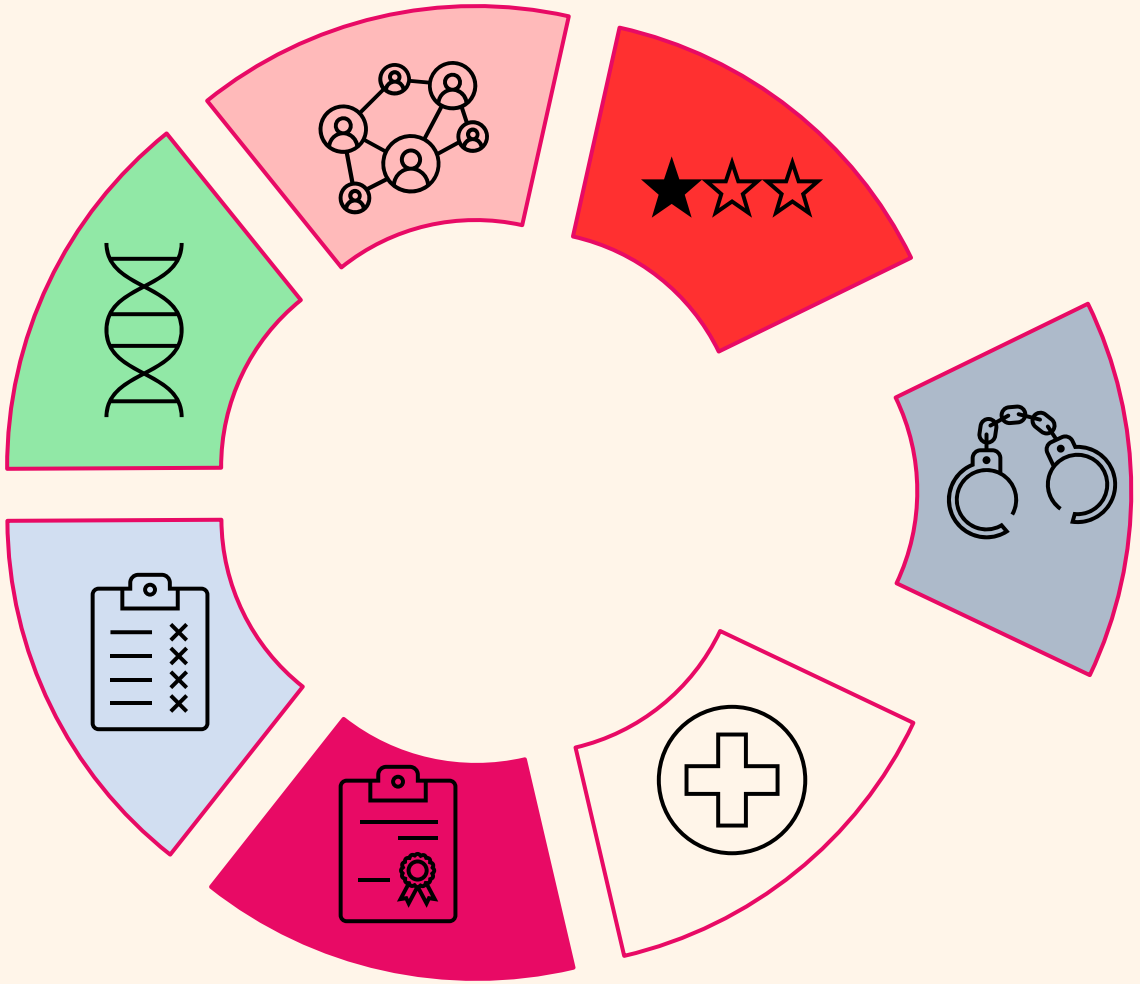
DEI – It's Just Good Science



Healthcare Access and Engagement

TESTS	RESULT	FLAG	UNITS	REFERENCE INTERVAL	LAB
BUN+Creat					
BUN	10		mg/dL	6 - 24	01
Creatinine, Serum	0.67		mg/dL	0.57 - 1.00	01
eGFR If NonAfricn Am	104		mL/min/1.73	>59	
eGFR If Africn Am	120		mL/min/1.73	>59	
BUN/Creatinine Ratio	15			9 - 23	

DEI – It's Just Good Science

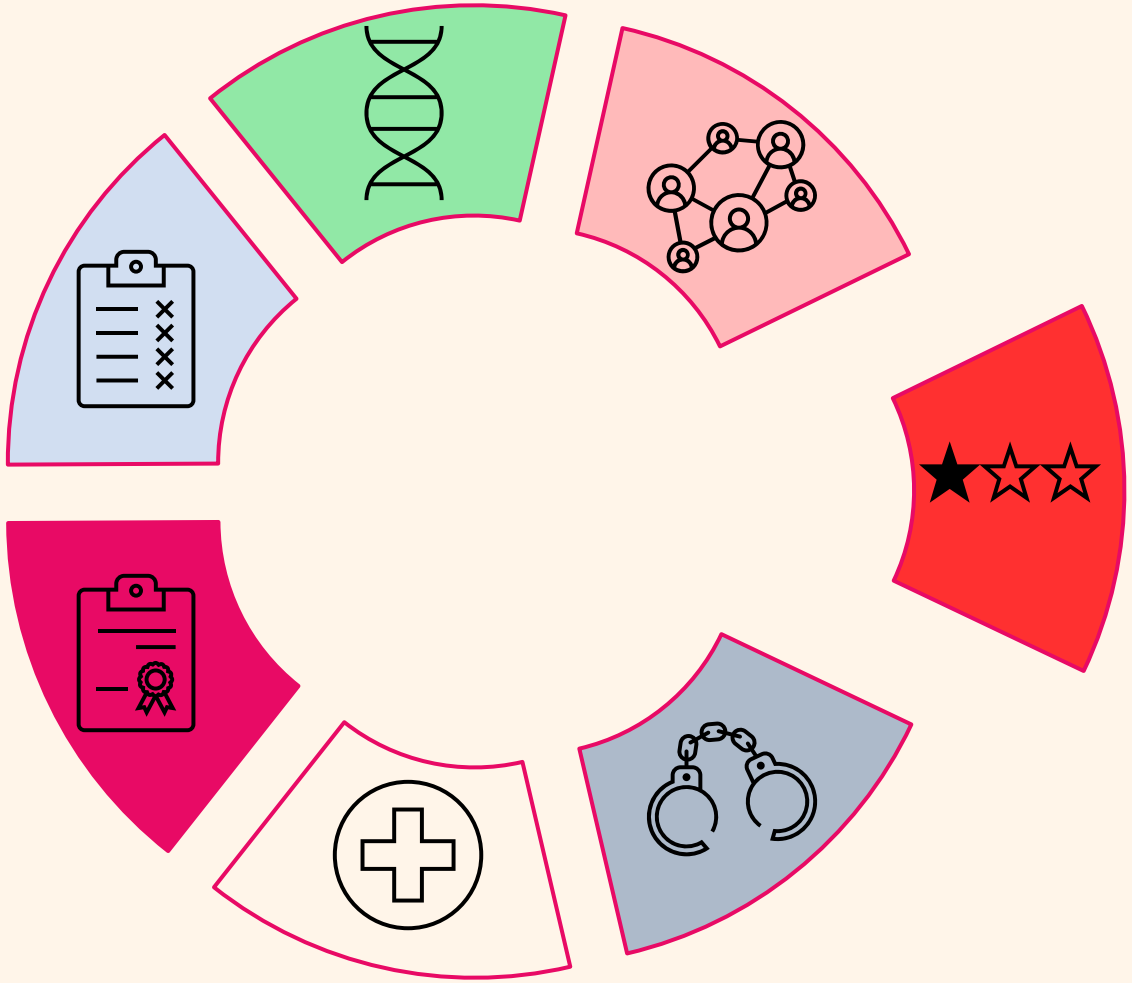


Research Fraud

Publication bias highlights the critical need for transparency in clinical research to ensure that treatment decisions are based on complete and accurate evidence

Underscores the importance of transparency, accountability, and inclusive evidence generation as foundational to achieving equity in health research

DEI – It's Just Good Science

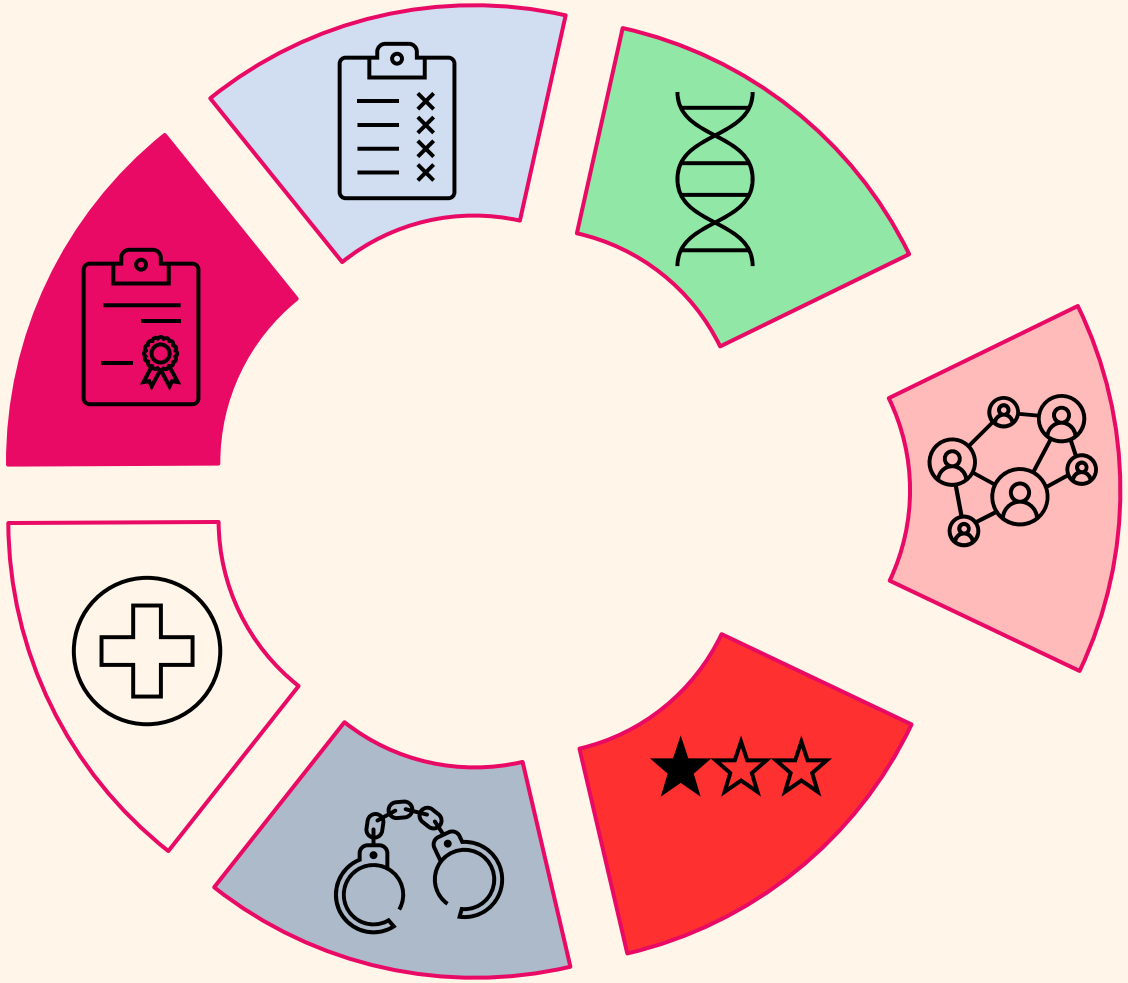


Data Integrity: Missed Safety

Yellow Card
Making medicines safer

DEI – It’s Just Good Science

Social Factors



Mistrust in Medical Institutions: underscores the critical importance of ethical conduct, community engagement, and culturally sensitive communication in clinical research, especially in regions with historical experiences of exploitation

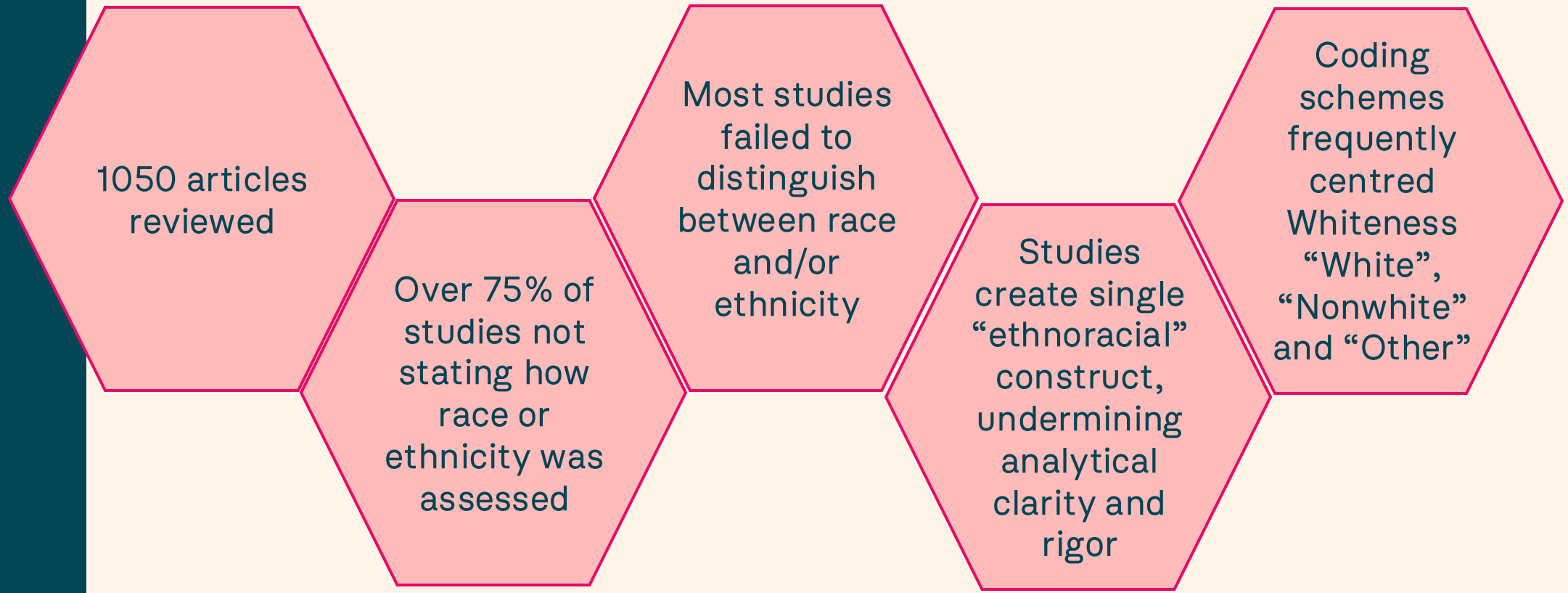
Socioeconomical Barriers: barriers collectively hinder equitable representation in research and reinforce the need for inclusive trial design and outreach strategies.

Deep Dive

Where Are We Now?

Deep Dive
Where are we now?

Systematic review in 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



Martinez et al. (2022)

NIH DATA MINING

The NIH Clinical Trials Database

Whilst the clinicaltrails.gov database has its uses, data mining is challenging given the quantity of unstructured, free-text data and tabulated data non-standard coding

482 Variables – only 42 are standardised

Here min_age has variously been recorded as

- "Min xx Years"
- "xx Yrs"
- "xx Years"
- "Years: xx + exclusive"
- "xx Weeks"

```
[1]: import pandas as pd
```

```
[33]: df = pd.read_csv("extract_20251029_185257.csv")
df = df.drop(columns = ["protocolSection.identificationModule.nctId",
                        "protocolSection.identificationModule.briefTitle",
                        "protocolSection.eligibilityModule.genderBased"])
df.columns = ['_date','sex','min_age','agecats','max_age']

df['minage'] = df["min_age"].str.extract(r"(\d+)", expand=False).astype("float")
df['maxage'] = df["max_age"].str.extract(r"(\d+)", expand=False).astype("float")
df['date'] = pd.to_datetime(df["_date"], format="%Y-%m-%d", errors="coerce").dt.year
```

```
[34]: df
```

[34]:

	_date	sex	min_age	agecats	max_age	minage	maxage	date
0	2010-06-17	ALL	18 Years	['ADULT', 'OLDER_ADULT']	NaN	18.0	NaN	2010.0
1	2023-11-01	ALL	18 Years	['ADULT', 'OLDER_ADULT']	NaN	18.0	NaN	2023.0
2	2022-12-01	ALL	23 Weeks	['CHILD']	12 Months	23.0	12.0	2022.0
3	2023-12	ALL	60 Years	['ADULT', 'OLDER_ADULT']	NaN	60.0	NaN	NaN
4	2006-01	FEMALE	18 Years	['ADULT']	45 Years	18.0	45.0	NaN
...	...	...	...	...	...	...	...	...
34038	2022-11-18	ALL	18 Years	['ADULT', 'OLDER_ADULT']	75 Years	18.0	75.0	2022.0
34039	2027-10-30	ALL	18 Years	['ADULT', 'OLDER_ADULT']	NaN	18.0	NaN	2027.0
34040	2025-07-31	ALL	18 Years	['ADULT', 'OLDER_ADULT']	NaN	18.0	NaN	2025.0
34041	2023-06-26	ALL	21 Years	['ADULT']	50 Years	21.0	50.0	2023.0
34042	2021-12-31	ALL	18 Years	['ADULT', 'OLDER_ADULT']	NaN	18.0	NaN	2021.0

NIH DATA MINING

The NIH Clinical Trials Database

```
pd.set_option("display.max_colwidth", None)
txtdf = df = pd.read_csv("extract_20251029_194435.csv")
txtdf.iloc[:,4]
```

```
0    All infants will be enrolled when in stable clinical condition within two weeks after birth. Therapy sessions will take place minimum three times
per week on three different days of the week directly at the unit during the entire hospitalization until discharge. After discharge music therapy trea
tment will be performed once a week until twelve months of corrected age.\n\nThe infant's wellbeing and relaxation during each session of music therapy
will be measured using the following parameters:\n\nHRV (Heart Rate Variability) analysis Oxygen's saturation Stress level of the child\n\nInfants will
be followed till 12 months of corrected age. Investigators will assess:\n\nNeuro-behavioral and neurological development of the child Stress level of t
he parents
Name: protocolSection.descriptionModule.detailedDescription, dtype: object
```

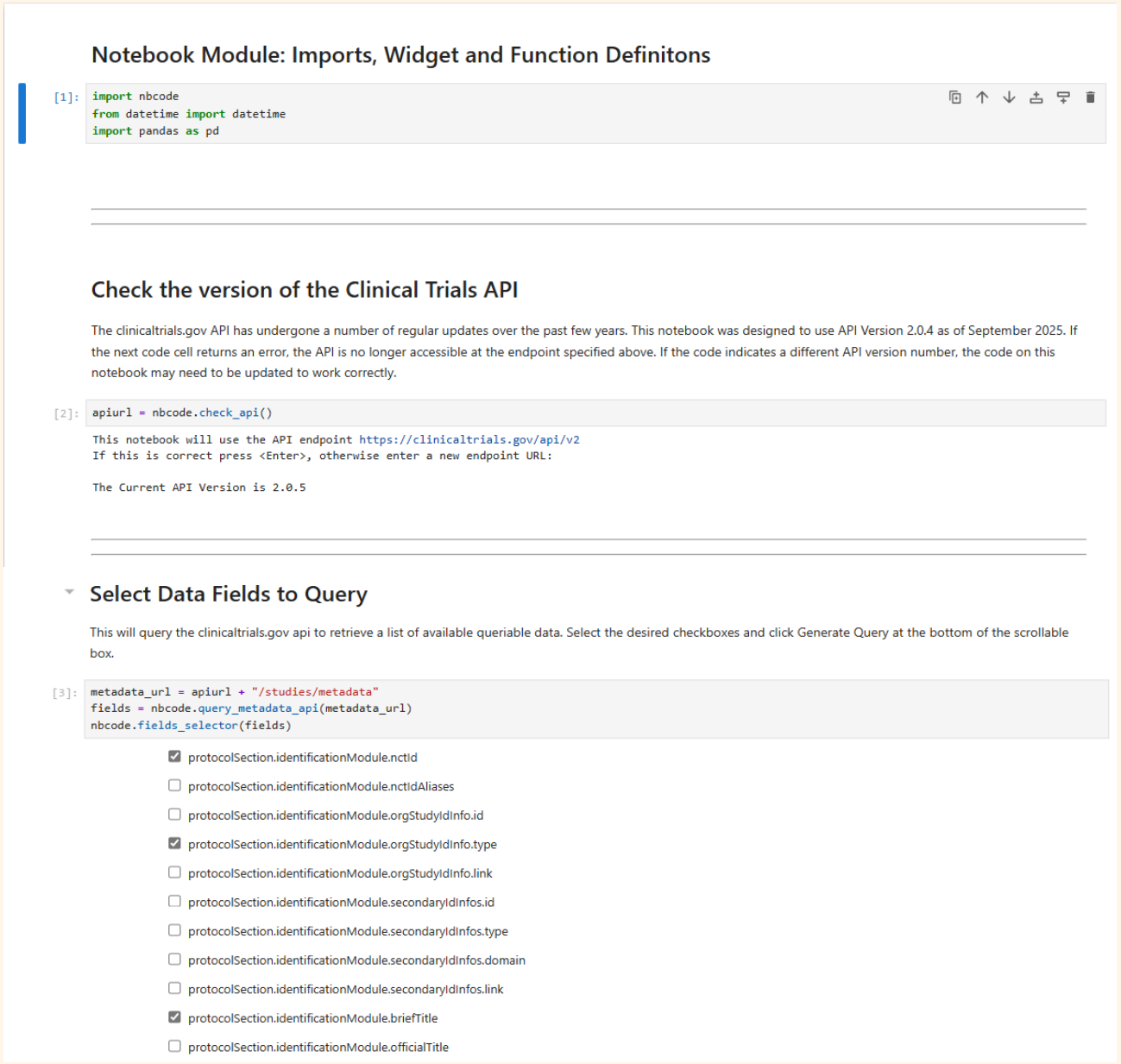
Much of the relevant information is available in free text form.
Given the lack of standard layout and variations in terminology extracting this data programmatically was challenging.

A task for an AI Agent!
Unfortunately, that was beyond the scope and budget available for this project

NIH DATA MINING

A Simple Graphical Interface to the API

The API provides a far richer data source that the web-query tool. We can simplify the process of accessing the data but problems with the data standardisation still remain.



NIH DATA MINING

A Simple Graphical Interface to the API

Query Terms and Date Range

Select a study start and end date range and one or more idications to query

```
[4]: nbcode.selector_form()
```

Pick a Start ... 01/01/1990 Pick an End ... 31/12/1993

Indication: Add Word

```
Added: psychiatric
Added: psychological
Added: psychology

Start Date: 1990-01-01
End Date: 1993-12-31
Indications: ['psychiatric', 'psychological', 'psychology']
```

Submit

Run The Specified Query

```
[11]: queryterm = nbcode.build_query(nbcode.get_form_result())
queryurl = apiurl + "/studies"
params = { "filter.advanced": queryterm, "pageSize": 100, "fields": nbcode.sections['fields'], "countTotal": "true" }
data = nbcode.query_api(params,queryurl)
df = pd.json_normalize(data, sep=".")

for col in df.columns:
    if df[col].apply(lambda x: isinstance(x, list) and all(isinstance(i, dict) for i in x)).any():
        df = df.explode(col).reset_index(drop=True)
        normalized = pd.json_normalize(df[col], sep=".")
        normalized.columns = [f"{col}.{subcol}" for subcol in normalized.columns]
        df = df.drop(columns=[col]).join(normalized)

timestamp = datetime.now().strftime("%Y%m%d_%H%M%S")
filename = f"extract_{timestamp}.csv"
df.to_csv(filename, index=False)
df.head()
```

Fetching pages: 0/? [00:00<?, ?page/s]

[11]:

	protocolSection.identificationModule.nctId	protocolSection.identificationModule.briefTitle	protocolSection.descriptionModule.briefSummary	protocolSection.eligibilityM
0	NCT00004758	Phase II Randomized Study of Early Surgery Vs ...	OBJECTIVES: I. Evaluate the efficacy of surgic...	
1	NCT01653535	Multisite Prevention of Conduct Problems (Fast...	The primary aim of this project is to evaluate...	
2	NCT00001320	Neuroimaging of Dopamine Metabolism in Normal ...	Brain cells communicate with each other by rel...	
3	NCT00004805	Study of the Effect of Four Methods of Cardiop...	OBJECTIVES: I. Describe the psychosocial respo...	
4	NCT00285636	Long Term Follow-Up of Burn Injuries	The purpose of the study is to investigate the...	

```
[ ]:
```

Regulatory Context

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.

Regulatory Context

Enhancing the Diversity of
Clinical Trial Popu
Eligibility Cr
Enrollment Prac
Trial Desi
Guidance for I

U.S. Department of Health and Hu
Food and Drug Administra
Center for Drug Evaluation and Res
Center for Biologics Evaluation and R

November 2020
Clinical/Medical

– Inclusion of children and adolescents in confirmatory clinical trials involving adults when appropriate[12, 13]

– Inclusion of women[14] in clinical trials in adequate numbers to allow for analysis by sex,[15] for example, by avoiding unjustified exclusion based on sex and taking other actions to promote inclusion.[16] For most drugs, representatives of both sexes should be inclu
of clinically significa

– Inclusion of racial an
clinical trial data by r
products (e.g., pharm
racially and ethnically
differences may be at
elimination), extrinsic

[12] For considerations regarding the inclusion
industry *Considerations for the Inclusion of*

[13] For considerations regarding the inclusion
for industry *Cancer Clinical Trial Eligibility*
Patients (July 2020).

[14] For the purposes of this guidance, the term

B. Trial Design and Methodological Approaches

Sponsors should consider various trial design and methodological approaches that will facilitate enrollment of a broader population. The following are examples of approaches to consider:

- Consider characterizing — in early clinical development — drug metabolism and clearance across populations that may metabolize or clear the drug differently (e.g., older adults and participants with liver or kidney dysfunction). Early characterization of drug metabolism and clearance across groups will help avoid later exclusions and, more generally, will allow dose adjustment to optimize effectiveness and safety across different populations.

- Using an adaptive clinical trial design would allow for pre-specified trial design changes during the trial when data become available, including altering the trial population.[22] An adaptive design can start with a narrow population if there are concerns about safety and can expand to a broader population based on interim safety data from the trial that provide support for doing so. Adaptive trials may also provide for broader enrollment when there is uncertainty regarding whether the drug will be safe and effective in certain populations, with an interim analysis that will enable adjustment of future enrollment based on pre-specified criteria regarding response.

- Consider a broader pediatric development program early. The arbitrary sequential enrollment of pediatric subgroups by chronological age for some conditions could unnecessarily delay development of medicines for children by limiting the population for study. Therefore, staggering enrollment in pediatric studies based on chronological age (i.e., enrollment of older pediatric participants first, then younger pediatric participants) should be justified with a clear scientific rationale (e.g., a known or potential developmental safety concern).[23]

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

Regulatory Context

Collection of Race and Ethnicity Data in Clinical Trials

Guidance for Industry and Food and Drug Administration Staff

Document issued on October 26

For questions about this document, contact the FDA Office of Min omh@fda.hhs.gov.

U.S. Department of Health and Human Serv
Food and Drug Administration (FD
Office of the Commissioner (OC
Office of Minority Health (OMH
Office of Women's Health (OWH
Center for Drug Evaluation and Research
Center for Biologics Evaluation and Researc
Center for Devices and Radiologic Health

October 2016
Clinical Medical

IV. COLLECTING RACE AND ETHNICITY DA CLINICAL TRIALS

The classifications discussed below provide a minim presenting data on race and ethnicity for Federal rep Policy Directive 15, the categories in this classifica not be interpreted as being scientific or anthropolog determinants of eligibility for participation in any F developed to provide a common framework for unif use of data on race and ethnicity by Federal agencie

The recommendations in this section reflect the Ag for more consistent demographic subgroup data col inside and outside the United States, the Agency rec is based on the current OMB Directive for collectin recommendations are also consistent with the forma set forth in NIH guidance[47].

A. Two-Question Format

In order to be consistent with OMB and other recommends using the two-question format fo with the ethnicity question preceding the ques

Question 1 (answer first): Do you consider y Hispanic/Latino?

Question 2 (answer second): Which of the fo describes you? More than one choice is accept

B. Self-Reporting

FDA recommends that trial participants self-re those individuals be permitted to designate a m self-reported designations is not feasible (e.g., respond), we recommend that the information other knowledgeable source. Race and ethnicity conducting the trial.

C. Ethnicity

For ethnicity, we recommend the following min

E. Use of More Detailed Racial and Ethnic Categories

In certain situations, as recommended in OMB Policy Directive 15, more detailed race and ethnicity information may be desired. For example, for clinical trials conducted outside the United States, FDA recognizes that the recommended categories for race and ethnicity were developed in the United States and that these categories may not adequately describe racial and ethnic groups in foreign countries. Furthermore, *White* can reflect origins in Europe, the Middle East, or North Africa; *Asian* can reflect origins from areas ranging from India to Japan.

In situations where appropriate, FDA recommends using more detailed categories by geographic region to provide sponsors the flexibility to adequately characterize race and ethnicity. As outlined in the 2011 HHS Implementation Guidance on Data Collection Standards for Race, Ethnicity, Sex, Primary Language, and Disability Status[4], if additional granularity or more detailed characterizations of race or ethnicity are collected to enhance understanding of the trial participants, FDA recommends these characterizations be traceable to the five minimum designations for race and two designations for ethnicity listed in sections D and C above. Example (from the above referenced 2011 HHS Guidance[4]):

Ethnicity Data Standard
Are you Hispanic, Latino/a, or of Spanish origin? (One or more categories may be selected)

a. ___ No, not of Hispanic, Latino/a, or Spanish origin
b. ___ Yes, Mexican, Mexican American, Chicano/a
c. ___ Yes, Puerto Rican
d. ___ Yes, Cuban
e. ___ Yes, Another Hispanic, Latino/a or Spanish origin

→ These categories roll up to the Hispanic or Latino category of the OMB standard

Race Data Standard
What is your race? (One or more categories may be selected)

a. ___ White
b. ___ Black or African American
c. ___ American Indian or Alaska Native

→ These categories are part of the current OMB standard

d. ___ Asian Indian
e. ___ Chinese
f. ___ Filipino
g. ___ Japanese
h. ___ Korean
i. ___ Vietnamese
j. ___ Other Asian

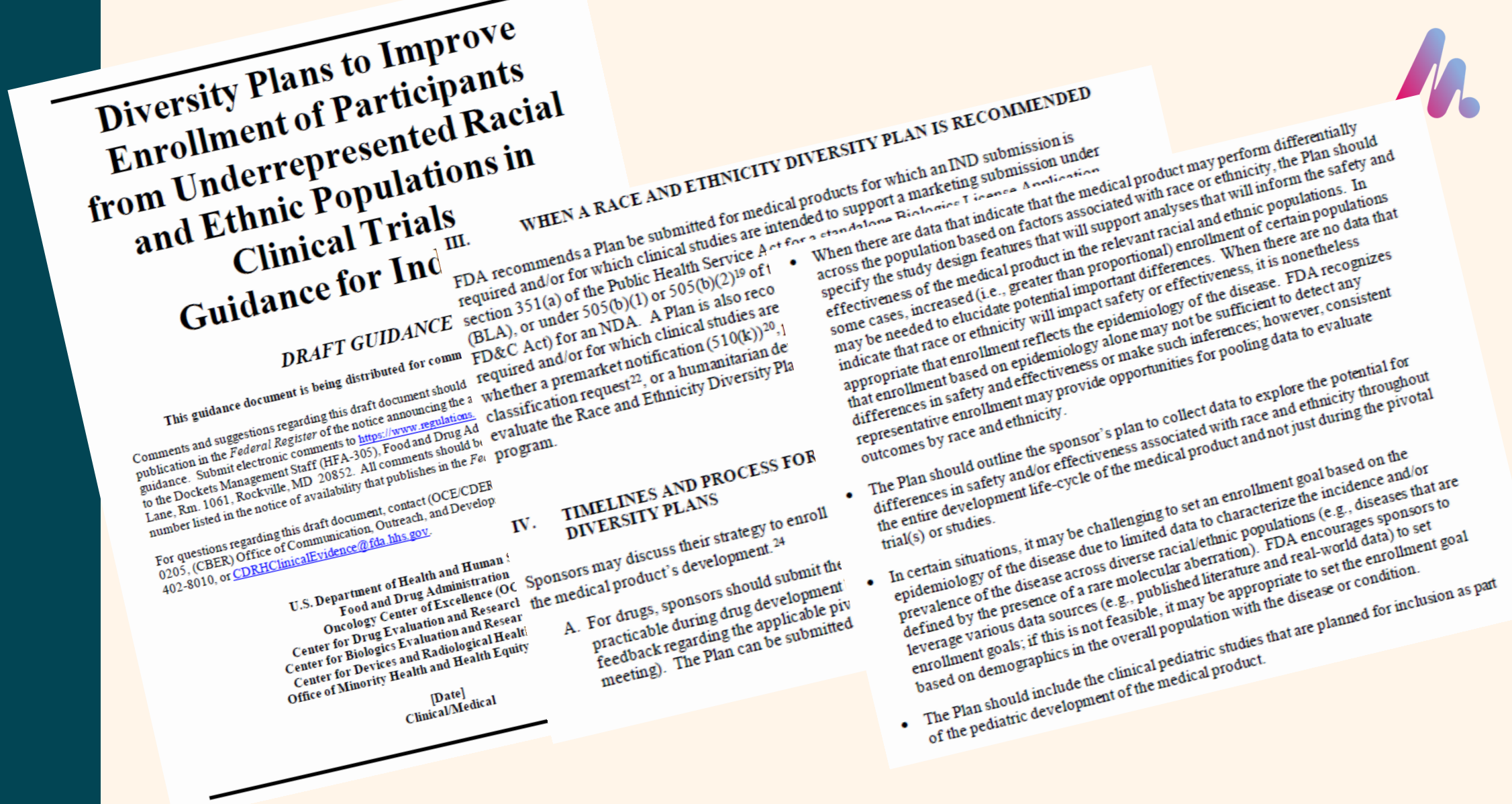
→ These categories roll up to the Asian category of the OMB standard

k. ___ Native Hawaiian
l. ___ Guamanian or Chamorro
m. ___ Samoan
n. ___ Other Pacific Islander

→ These categories roll-up to the Native Hawaiian or Other Pacific Islander category of the OMB standard

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

Regulatory Context



Diversity Plans to Improve Enrollment of Participants from Underrepresented Racial and Ethnic Populations in Clinical Trials

Guidance for Ind

DRAFT GUIDANCE

This guidance document is being distributed for comm

Comments and suggestions regarding this draft document should
publication in the *Federal Register* of the notice announcing the a
guidance. Submit electronic comments to https://www.regulations.
to the Dockets Management Staff (HFA-305), Food and Drug Ad
Lane, Rm. 1061, Rockville, MD 20852. All comments should b
number listed in the notice of availability that publishes in the Fe

For questions regarding this draft document, contact (OCE/CDER
0205, (CBER) Office of Communication, Outreach, and Develop
402-8010, or CDRHClinicalEvidence@fda.hhs.gov.

U.S. Department of Health and Human
Food and Drug Administration
Oncology Center of Excellence (OC
Center for Drug Evaluation and Researcl
Center for Biologics Evaluation and Resear
Center for Devices and Radiological Healt
Office of Minority Health and Health Equity

[Date]
Clinical/Medical

III. WHEN A RACE AND ETHNICITY DIVERSITY PLAN IS RECOMMENDED

FDA recommends a Plan be submitted for medical products for which an IND submission is required and/or for which clinical studies are intended to support a marketing submission under section 351(a) of the Public Health Service Act for a standalone Biologics License Application (BLA), or under 505(b)(1) or 505(b)(2)[19] of t FD&C Act) for an NDA. A Plan is also reco required and/or for which clinical studies are whether a premarket notification (510(k))[20],[1] classification request[22], or a humanitarian de evaluate the Race and Ethnicity Diversity Pla program.

- When there are data that indicate that the medical product may perform differentially across the population based on factors associated with race or ethnicity, the Plan should specify the study design features that will support analyses that will inform the safety and effectiveness of the medical product in the relevant racial and ethnic populations. In some cases, increased (i.e., greater than proportional) enrollment of certain populations may be needed to elucidate potential important differences. When there are no data that indicate that race or ethnicity will impact safety or effectiveness, it is nonetheless appropriate that enrollment reflects the epidemiology of the disease. FDA recognizes that enrollment based on epidemiology alone may not be sufficient to detect any differences in safety and effectiveness or make such inferences; however, consistent representative enrollment may provide opportunities for pooling data to evaluate outcomes by race and ethnicity.
- The Plan should outline the sponsor's plan to collect data to explore the potential for differences in safety and/or effectiveness associated with race and ethnicity throughout the entire development life-cycle of the medical product and not just during the pivotal trial(s) or studies.
- In certain situations, it may be challenging to set an enrollment goal based on the epidemiology of the disease due to limited data to characterize the incidence and/or prevalence of the disease across diverse racial/ethnic populations (e.g., diseases that are defined by the presence of a rare molecular aberration). FDA encourages sponsors to leverage various data sources (e.g., published literature and real-world data) to set enrollment goals; if this is not feasible, it may be appropriate to set the enrollment goal based on demographics in the overall population with the disease or condition.
- The Plan should include the clinical pediatric studies that are planned for inclusion as part of the pediatric development of the medical product.

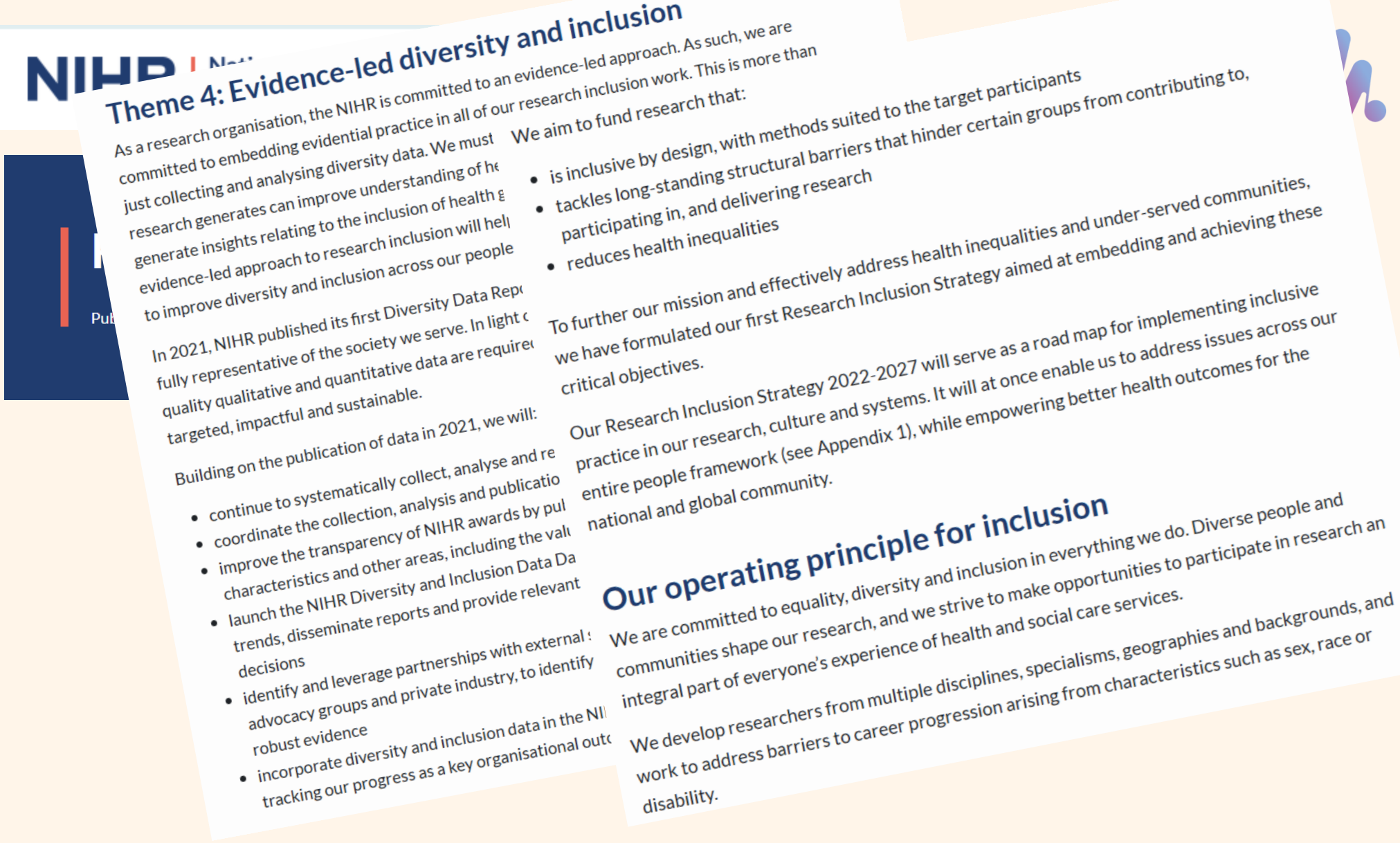
IV. TIMELINES AND PROCESS FOR DIVERSITY PLANS

Sponsors may discuss their strategy to enroll the medical product's development.[24]

A. For drugs, sponsors should submit the practicable during drug development feedback regarding the applicable piv meeting). The Plan can be submitted

"Diversity Plans to Improve Enrolment of Participants from Underrepresented Racial and Ethnic Populations in Clinical Trials Guidance for Industry" - Draft Guidance, published in the Federal Register on April 14, 2022.

Regulatory Context



Theme 4: Evidence-led diversity and inclusion

As a research organisation, the NIHR is committed to an evidence-led approach. As such, we are committed to embedding evidential practice in all of our research inclusion work. This is more than just collecting and analysing diversity data. We must

research generates can improve understanding of he

generate insights relating to the inclusion of health g

evidence-led approach to research inclusion will hel

to improve diversity and inclusion across our people

In 2021, NIHR published its first Diversity Data Repo

fully representative of the society we serve. In light o

quality qualitative and quantitative data are require

targeted, impactful and sustainable.

Building on the publication of data in 2021, we will:

- continue to systematically collect, analyse and re
- coordinate the collection, analysis and publicatio
- improve the transparency of NIHR awards by pul
- characteristics and other areas, including the valu
- launch the NIHR Diversity and Inclusion Data Da
- trends, disseminate reports and provide relevant decisions
- identify and leverage partnerships with external s
- advocacy groups and private industry, to identify robust evidence
- incorporate diversity and inclusion data in the NI
- tracking our progress as a key organisational outc

We aim to fund research that:

- is inclusive by design, with methods suited to the target participants
- tackles long-standing structural barriers that hinder certain groups from contributing to, participating in, and delivering research
- reduces health inequalities

To further our mission and effectively address health inequalities and under-served communities, we have formulated our first Research Inclusion Strategy aimed at embedding and achieving these critical objectives.

Our Research Inclusion Strategy 2022-2027 will serve as a road map for implementing inclusive practice in our research, culture and systems. It will at once enable us to address issues across our entire people framework (see Appendix 1), while empowering better health outcomes for the national and global community.

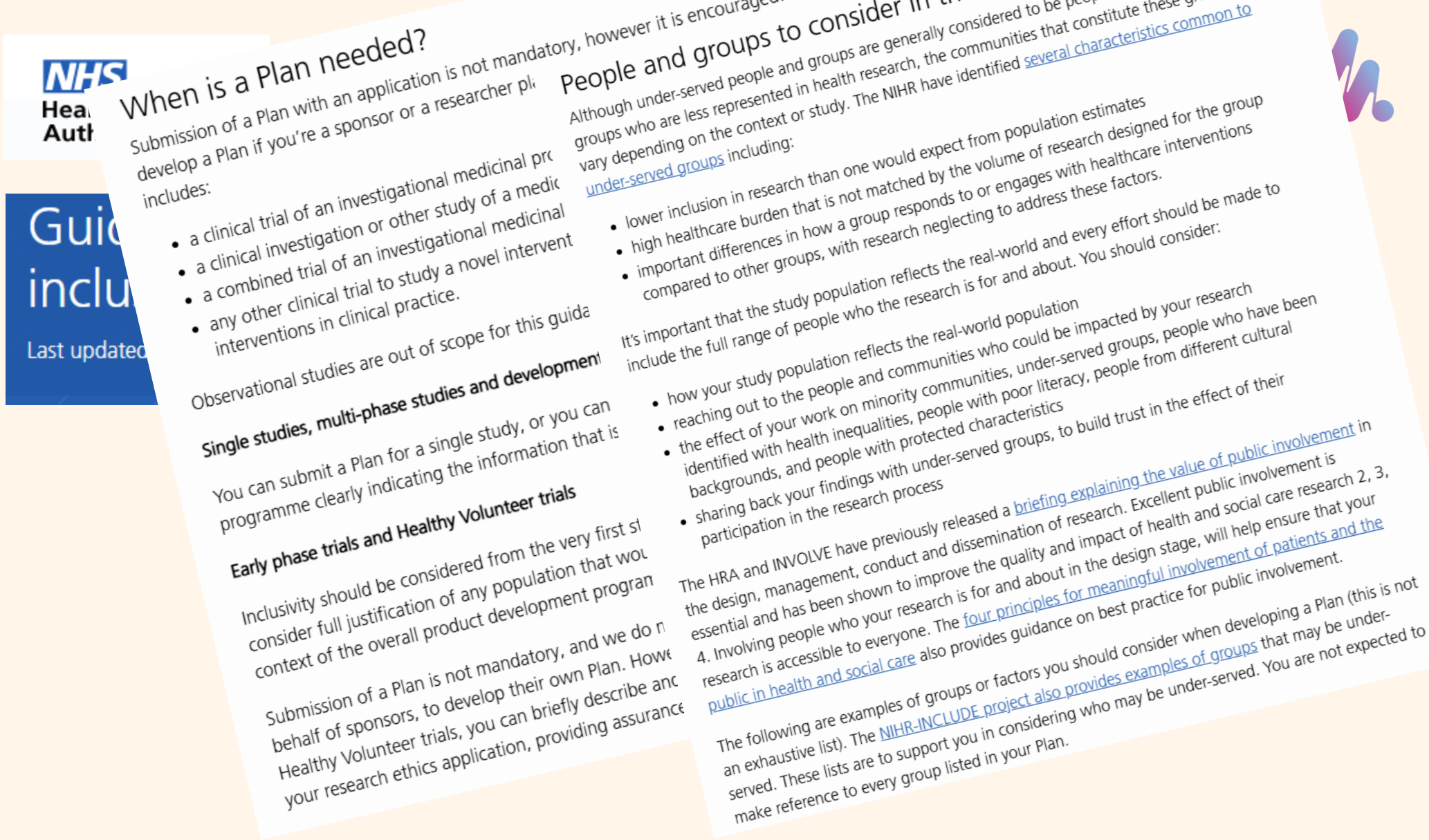
Our operating principle for inclusion

We are committed to equality, diversity and inclusion in everything we do. Diverse people and communities shape our research, and we strive to make opportunities to participate in research an integral part of everyone's experience of health and social care services.

We develop researchers from multiple disciplines, specialisms, geographies and backgrounds, and work to address barriers to career progression arising from characteristics such as sex, race or disability.

"NIHR Research Inclusion Strategy 2022 – 2027" - Issued 26 September 2022, Version: 3.0 - July 2025

Regulatory Context



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

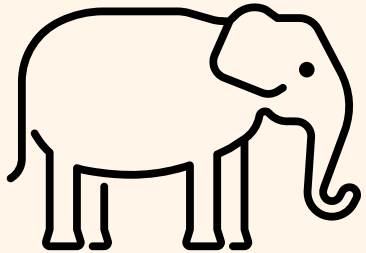
Path Forward

What Can We Do?

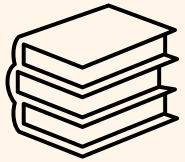
Path Forward
What Can We Do?

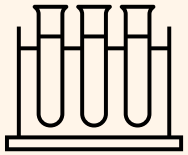
In early 2025, the FDA quietly removed its draft guidance on enhancing diversity in clinical trials. This regulatory shift raises concerns about the future of inclusive research practices and their impact on equitable healthcare outcomes.

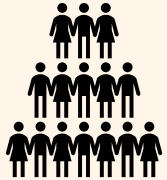
While the FDA's withdrawal of DEI guidance in 2025 was unexpected, the pursuit of diversity in clinical trials remains a scientific imperative, regardless of shifting political landscapes.



Path Forward – What Can We Do?







Path Forward – What Can We Do?

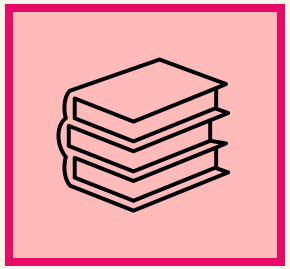
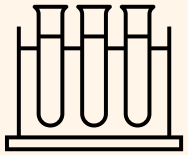
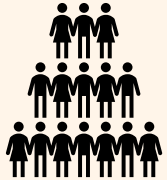
EMBEDDING DIVERSITY IN RESEARCH LEADERSHIP

Diversity within research teams 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.

Build trust and credibility, people are more inclined to trust leaders who resemble them

Path Forward – What Can We Do?

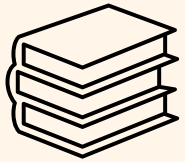
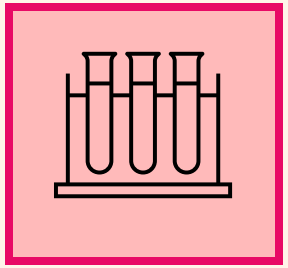
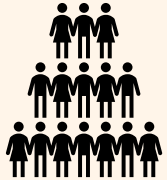
UPDATE MANDATES AND DOCUMENTATION

Current diversity planning and demographic reporting 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.

More nuanced frameworks are needed, such as incorporating intersectional data, using granular ethnicity classifications and applying genomic ancestry markers where appropriate to complement self-reported data

Path Forward – What Can We Do?

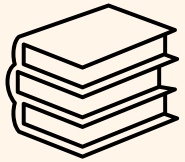
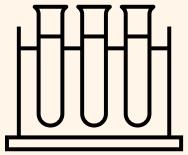
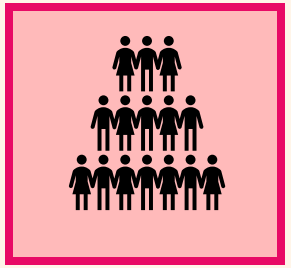
EARLY DIAGNOSIS AND RESEARCH INITIATIVES

Earlier diagnosis of disease also plays a vital role in improving DEI in clinical trials. When individuals are diagnosed sooner, they are more likely to be identified and recruited for research opportunities

Earlier diagnosis enables trials to better reflect the real-world patient population, leading to more inclusive study designs and more generalizable 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.

Path Forward – What Can We Do?

EMBEDDING EQUALITY: INCLUDE ETHINICITY FRAMEWORK

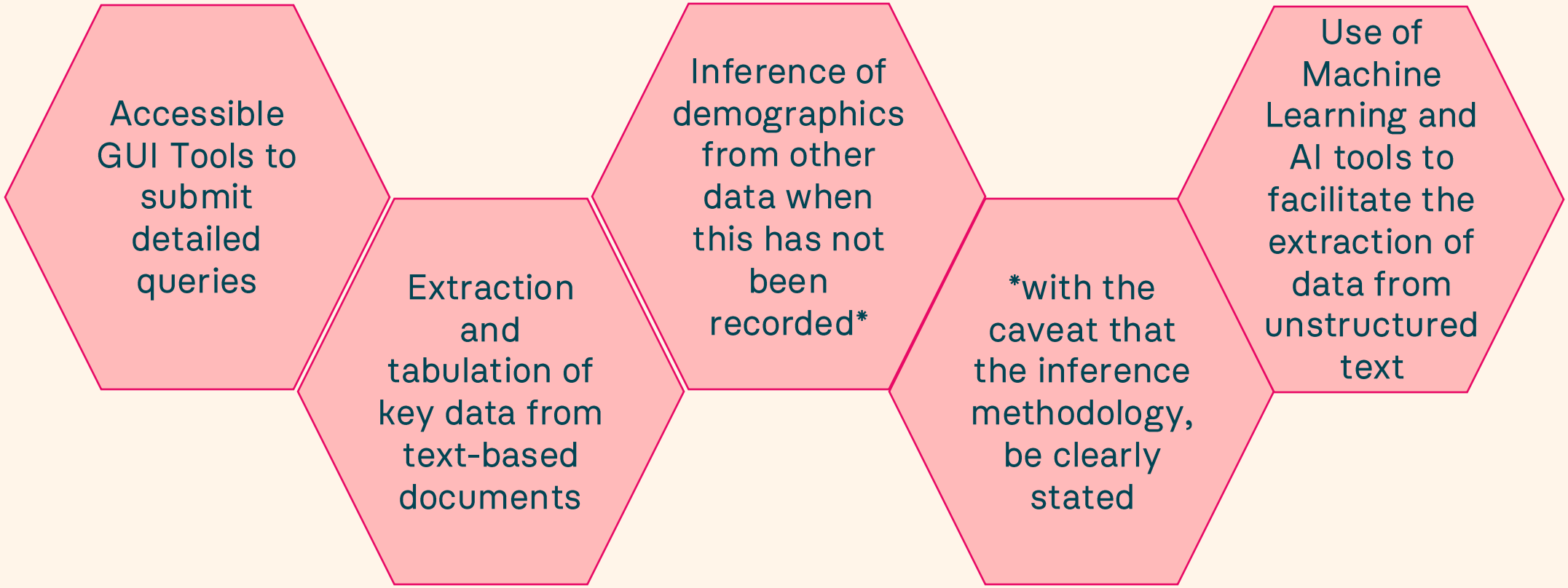
Addresses a critical gap in clinical research: the underrepresentation of ethnic minority groups in trials

The Framework encourages trialists to move beyond participant numbers and consider who is included and why

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

Path Forward – What Can We Do?

THE ACCESSIBILITY AND USABILITY OF PUBLIC CLINICAL TRIAL DATASETS



Conclusion

If our science doesn’t reflect everyone, represent everyone, who are we actually healing?

Acknowledgements

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.

Questions?

roxanne.mcguire@veramed.com
chris.ward@veramed.com