

Five Data-Informed Principles for Advancing Inclusive Research in Clinical Trials: A Pharma Perspective

Spencer L James, MD^{1*} Shalini Hede, PharmD^{1*} Altovise T Ewing-Crawford, PhD¹ Ruma Bhagat, MD¹ Nicole Richie, PhD¹ Mitchell D'Rozario, PhD¹ Pierre Theodore, MPH¹ Bea Lavery, MSc¹ Sarah Bentouati, PharmD² Assaf P Oron, PhD³ Catherine W Gillespie, PhD³ Cleo A Ryals, PhD⁴ Megan H Bair-Merritt, MD, MSCE⁵ Johanna Chesley, MPH⁵ Evelyn Jiagge, PhD⁶ Bruno Jolain, MD²

* Co-first authors.

Author affiliations:

¹Genentech Inc., 1 DNA Way, South San Francisco, CA, 94080, USA; ²F. Hoffmann-La Roche Ltd, Basel, Switzerland; ³Institute for Health Metrics and Evaluation at the University of Washington, 3980 15th Avenue NE, Seattle, WA 98195 USA; ⁴Flatiron Health, 233 Spring Street, New York, NY 10013, USA; ⁵Boston Medical Center, Office of Research and Sponsored Programs, 960 Massachusetts Avenue, 2nd floor, 2545, Boston, MA 02118, USA; ⁶Henry Ford Health, 2799 W Grand Blvd, Detroit, MI 48202, USA

Corresponding author information:

Shalini Hede

Genentech Inc., 1 DNA Way, South San Francisco, CA, 94080

hede.shalini@gene.com

Phone: 1.908.566.8436

SUPPLEMENT

Supplemental Figure 1. Principle 1: Population science considerations. WHO is the population most affected by the disease or condition?



Biological & Genetics Considerations:

- Biomarkers
- Ancestry
- Risk factors
- Others include:
 - Epidemiology
 - Disparities in diagnosis
 - Biology and progression of disease
 - Treatment patterns
 - Morbidity and mortality
 - Quality of care



Demographics Considerations:

- Race & Ethnicity
- Age & Sex



Population science descriptors (via literature reviews) translated and captured in Diversity Action Plan reports across disease areas

Supplemental Figure 2. Principle 2: Data-informed country and site placement. WHERE are the patients?

Geographical Proportionality Trial Enrollment (Region)	■ Global trial enrollment by region is recommended in International Council for Harmonisation (ICH) Guideline on General Principles for Planning and Design of Multi-Regional Clinical Trials (E17) guidelines on multiple regional clinical trials and should be proportional to the geographical burden of disease, as defined by appropriate reference metrics ■ Data sources such as the US Cancer Statistics (USCS) for cancer in the US and the Institute for Health Metrics and Evaluation (IHME) Global Burden of Disease study globally, can be leveraged based on data quality, representativeness, and disease area coverage to provide epidemiology estimates at the indication or disease area level.
Enable Access and Commercialization (Country)	■ Country-specific enrollment goals should adhere to standardized country-specific requirements and frameworks. Diversity outside the US cannot be captured or measured by US-specific demographic frameworks alone, such as the United States Office of Management and Budget's (US OMB) standards on race and ethnicity
Representative real-world population (Site)	■ Data-informed opportunities to maximize inclusion that leverage real-world evidence and ensure clinical trials are representative of real-world patient populations. Data sources such as the US Census, US Cancer Statistics, Surveillance, Epidemiology, and End Results program, electronic health record-derived real-world data (e.g., TrinetX, IHME, Flatiron), and genomic databases are available to obtain epidemiological estimates for US and global populations. ■ Resources: AIR Calculator, R-index

Supplemental Figure 3. Principle 3. End-to-end inclusive trial design. WHAT is the trial design? ASCO, American Society of Clinical Oncology; FDA, Food and Drug Administration.

Data-driven

Artificial intelligence algorithms combined with real-world data can potentially improve several aspects of clinical trials including:

- Patients screening
- Predict likelihood of patients who will enroll
- Extract features from electronic health records (EHRs)



ASCO evidence-based recommendations that laboratory tests should be used as exclusion criteria only when necessary. ([link](#))

FDA has also developed a series of guidelines for eligibility criteria, however these are not standards and have not been widely implemented by many companies.

User Informed

Include input from providers, patient/patient advocacy groups, and investigators on the study trial design.

Supplemental Figure 4. Principle 4: Patient-reported data collection standards. WHO are the individual patients participating? (A) standardized electronic case report form (eCRF) to Food and Drug Administration/Office of Management and Budget (FDA/OMB) race categories (B) social determinants of health (SDOH) questionnaire (C) sexual orientation gender identity (SOGI) data collection fields according to the Clinical Data Interchange Standards Consortium (<https://www.cdisc.org/kb/ecrf/sexual-orientation-gender-identity-sogi>)

A

Race Grouping	Race Sub-category	Mapping (to FDA guidance)
ALASKA NATIVE OR AMERICAN INDIAN	North American Indian Or Alaska Native	American Indian Or Alaska Native
	Central American Indian	American Indian Or Alaska Native
	South American Indian	American Indian Or Alaska Native
	Other Indigenous Not Listed	American Indian Or Alaska Native
ASIAN	Chinese	Asian
	Filipino	Asian
	Japanese	Asian
	Korean	Asian
	South Asian	Asian
	Vietnamese	Asian
	Other Asian Not Listed	Asian
BLACK	African American	Black Or African American
	African Caribbean	Black Or African American
	Black African	Black Or African American
	Other Black Identity Not Listed	Black Or African American
INDIGENOUS AUSTRALIAN OR PACIFIC ISLANDER	Aboriginal Or Torres Strait Islander	Native Hawaiian Or Other Pacific Islander
	Maori	Native Hawaiian Or Other Pacific Islander
	Native Hawaiian	Native Hawaiian Or Other Pacific Islander
	Other Pacific Islander Not Listed	Native Hawaiian Or Other Pacific Islander
MIDDLE EASTERN, NORTH AFRICAN, OR WHITE	Middle Eastern	White
	North African	White
	White	White
	Unknown	Unknown

B

Education

What is your level of education/schooling?

- ☐ Elementary School
- ☐ Middle School
- ☐ Some High School
- ☐ Graduated from High School
- ☐ Some College
- ☐ Graduated from College
- ☐ Post-Graduate

Employment Status

Before your diagnosis, what was your employment status?

- ☐ Full-time Employed
- ☐ Part-time Employed
- ☐ Self-employed (either full-time or part-time)
- ☐ Retired
- ☐ Not employed

Health Insurance Status

What type of health insurance do you have, if any.

Please select one:

- ☐ Private/Commercial
- ☐ Medicare
- ☐ Medicaid
- ☐ Veteran Affairs (VA) Health Plan
- ☐ Other
- ☐ I do not have a health insurance

Income Status

What is your annual household income?

- ☐ < \$15,000
- ☐ \$15,000 - \$34,999
- ☐ \$35,000 - \$49,999
- ☐ \$50,000 - \$74,999
- ☐ \$75,000 - \$99,999
- ☐ \$100,000 - \$149,999
- ☐ >\$150,000

Marital Status

What is your marital status?

- ☐ Single (never married)
- ☐ Married (including common law)
- ☐ Separated
- ☐ Divorced
- ☐ Widowed
- ☐ Domestic Partnership (same sex, opposite sex or unregistered)

Zip Code of Home Residence

What is the zip code of your home residence?

 Protocol Sexual Orientation and Gender Identity								
	Site Number			Subject Number				

Sexual Orientation and Gender Identity	
1 Sexual Orientation and Gender Identity	
1.1	Was SOGI (Sexual Orientation and Gender Identity) data collected? ☐ Yes SCPERF ☐ No
1.2	What [is/was] the date of the collection? (DD-MM-YYYY) SCDAT SCDTC
1.3	What sex were you assigned at birth on your original birth certificate? (Check one): ☐ Female SEXABRTH_SCORES ☐ Male SCORES where SCTESTCD = "SEXABRTH" ☐ Intersex ☐ Not Reported
1.4	Were you born with a variation in your physical sex characteristics? (this is sometimes called being intersex or having a difference in sex development)? ☐ Yes ISXDIND_SCORES ☐ No SCORES WHERE SCTESTCD = "ISXDIND" ☐ Unknown ☐ Not Reported
1.5	What is your gender identity at this time? (Check all that apply): ☐ Cisgender Woman/Girl ☐ Cisgender Man/Boy ☐ Transgender Woman/Girl ☐ Transgender Man/Boy ☐ Gender Queer ☐ Gender Fluid ☐ Non-Binary GENIDENT_SCORES ☐ Unknown SCORES WHERE SCTESTCD = "GENIDENT" ☐ Not Reported
1.6	Which of the following represents how you think of your sexual orientation at this time? (Check all that apply): ☐ Lesbian SEXORIE_SCORES ☐ Gay SCORES WHERE SCTESTCD = "SEXORIRS" ☐ Straight or Heterosexual ☐ Bisexual ☐ Queer ☐ Pansexual ☐ Asexual ☐ Aromantic ☐ Unknown ☐ Not Reported

Supplemental Figure 5. Enabling access to trials. HOW will patients access trials and treatment beyond trial completion?



A fundamental and ethical aspect of study placement and access to only conduct trials in countries where a sustainable path to regulatory and access approval exists



Partnerships to enable access during trials ie. transportation



Post-trial access/reimbursement programs



Demonstrating trustworthiness