Racial and Ethnic Disparities among Participants in US-Based Phase 3 Randomized Cancer Clinical Trials

Stephen R. Grant, M.D.1, Timothy A. Lin, M.D.2, Austin B. Miller, B.S.3, Walker Mainwaring, B.A4, Andres F. Espinoza, M.D.4, Amit Jethanandani, M.D. M.P.H.5, Gary V. Walker, M.D. M.P.H. M.S.6, Benjamin D. Smith, M.D.1, B. Ashleigh Guadagnolo, M.D. M.P.H.1, Reshma Jagsi M.D. D.Phil.7, C. David Fuller, M.D. Ph.D.1,8, Charles R. Thomas, Jr., M.D.8, Ethan B. Ludmir, M.D.1*

Affiliations: 1The University of Texas MD Anderson Cancer Center, Houston, TX, USA; 2The Johns Hopkins University School of Medicine, Baltimore, MD, USA; 3The University of Texas Health Science Center McGovern Medical School, Houston, TX, USA; 4Baylor College of Medicine, Houston, TX, USA; 5The University of Tennessee Health Science Center College of Medicine, Memphis, TN, USA; 6Banner MD Anderson Cancer Center, Gilbert, AZ, USA; 7The University of Michigan, Ann Arbor, MI, USA; 8Oregon Health & Science University, Portland, OR, USA

*Corresponding Author: Ethan B. Ludmir, M.D., 1400 Pressler St, Unit 1422, Houston TX, USA 77030. Email: ebludmir@mdanderson.org.

© The Author(s) 2020. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Abstract

While improving representation of racial and ethnic groups in United States (US) clinical trials has been a focus of federal initiatives for nearly three decades; the status of racial/ethnic minority enrollment on cancer trials is largely unknown. We utilized a broad collection of phase 3 cancer trials derived from ClinicalTrials.gov to evaluate racial/ethnic enrollment among US cancer trials. The difference in incidence by race/ethnicity (D-IRE) was the median absolute difference between trial and corresponding Surveillance, Epidemiology, and End Results (SEER) data. All statistical tests were two-sided. Using a cohort of 168 eligible trials, median D-IRE was +6.8% for Whites (IQR +1.8%, +10.1%, p<.001 by Wilcoxon signed-rank test comparing median D-IRE to a value of zero), −2.6% for Blacks (IQR −5.1%, +1.2%, p=.004), −4.7% for Hispanics (IQR −7.5%, −0.3%, P<.001), and −4.7% for Asians (IQR −5.7%, −3.3%, P<.001). These data demonstrate overrepresentation of Whites, with continued underrepresentation of racial/ethnic minority subgroups.
In 1993, the National Institute of Health (NIH) Revitalization Act [1-3] detailed a plan for the inclusion of racial/ethnic minority groups in clinical research in the United States (US). Numerous studies have since examined racial/ethnic minority representation on cancer trials, with the large majority demonstrating overrepresentation of Whites with underrepresentation of racial/ethnic minorities [4-6]. There is limited data, however, on recent enrollment trends in US-only cancer trials, few reports on Native American/Hawaiian groups, and little known of the association between racial/ethnic enrollment and trial type. To that end, we utilized a broad clinical trials database to determine racial/ethnic enrollment disparities in recent US clinical trials with the hypothesis that historically underrepresented minorities will be under enrolled on recent cancer clinical trials.

ClinicalTrials.gov was queried by the following search parameters: Terms: “cancer”; Study Type: “All Studies”; Status: excluded “Not yet recruiting”; Phase: Phase 3; and Study Results: “With Results.” Of 1,239 identified trials, 168 addressed a therapeutic intervention with exclusive US enrollment. The ClinicalTrials.gov database and manuscript publications were reviewed for race/ethnicity reporting. Race/ethnicity was self-reported, with exclusion of multiple or unknown race. All current US cooperative groups (SWOG, ECOG-ACRIN, NRG, Alliance, and COG) were represented in this dataset. Incidence of race/ethnicity in the US cancer population was estimated using the “Race recode” and “Origin recode” of the Surveillance, Epidemiology, and End Results (SEER) database, extracted in five-year increments with race and ethnicity considered separately.

Trial and SEER data were compared, correlating with the median year of patient enrollment. Comparative SEER data was filtered for the disease site of interest, except for trials where enrollment included patients with more than three primary disease sites. The difference in incidence by race/ethnicity (D-IRE) was defined as the median absolute difference in race/ethnicity incidence between trial and corresponding disease-specific SEER data, with a negative value indicating underrepresentation (Table 1). The ratio of incidence by race/ethnicity (R-IRE) was defined as the median ratio of trial (numerator) and SEER (denominator) incidence, with a value less than 1 indicating underrepresentation. D-IRE and R-IRE values for each race/ethnicity were calculated only for trials that reported on that particular
subgroup. Statistical tests were non-parametric with an a priori threshold of \( \alpha=0.05 \) for statistical significance. Mann-Whitney U and Kruskall-Wallis ANOVA were used to compare subgroups. The Wilcoxon signed rank test was used to compare median D-IRE for each race/ethnicity to a value of 0 and R-IRE to a value of 1 (null hypotheses).

Of 168 eligible trials, 96 (57.1\%) reported the proportion of at least one race/ethnicity, representing 34,329 patients. Of these 96 trials, 97.9\% reported proportion of White enrollees, compared to 84.4\% reporting Black, 52.1\% Hispanic, 67.7\% Asian, 61.4\% Native American, and 56.3\% Native Hawaiian/Pacific Islander. The median proportion of White enrollees on the included trials was 88.7\%, compared to 8.6\% Black, 4.0\% Hispanic, 1.4\% Asian, 0.1\% Native American, and 0.0\% Native Hawaiian/Pacific Islander. Cooperative-group-sponsored trials were more likely to report race/ethnicity (66.0\% vs. 45.9\% for non-cooperative-group-supported trials, \( p=.02 \)).

The median D-IRE was +6.8\% for Whites (IQR +1.8\%, +10.1\%, \( p<.001 \) by Wilcoxon signed-rank test comparing median D-IRE to a value of zero), −2.6\% for Blacks (IQR −5.1\%, +1.2\%, \( p=.004 \)), −4.7\% for Hispanics (IQR −7.5\%, −0.3\%, \( P<.001 \)), and −4.7\% for Asians (IQR −5.7\%, −3.3\%, \( P<.001 \)) (Table 1, Figure 1). The median R-IRE was 1.08 for Whites (IQR 1.02, 1.12, \( p<.001 \)), 0.76 for Blacks (IQR 0.47, 1.09, \( p=.003 \)), 0.50 for Hispanics (IQR 0.16, 0.96, \( p<.001 \)), and 0.19 for Asians (IQR 0.08, 0.45, \( p<.001 \)) (Figure 1). There was no difference in the D-IRE over time (median trial enrollment date <2005 vs. 2005-2007 vs. 2008-2010 vs. >2010) for any race/ethnicity. A sensitivity analysis excluding trials enrolling multiple disease sites (22 of 96 trials) showed similar findings by racial/ethnic groups (D-IRE of +6.3\% for Whites [IQR +1.8\%, +10.2\%, \( p<.001 \)], −2.5\% for Blacks [IQR −4.6\%, +1.5\%, \( p=.03 \)], −4.5\% for Hispanics [IQR −6.8\%, −0.4\%, \( P=.001 \)], and −4.6\% for Asians [IQR −5.8\%, −3.1\%, \( P<0.001 \)])).

Equitable representation on US cancer clinical trials is necessary to ensure generalizable results and allow for equal access to new treatment advances, and has been an explicit priority for NIH-supported clinical trials for nearly three decades [1-3]. Potential drivers of underrepresentation are complex and include narrow eligibility criteria, lack of access to participating centers, patient preference, fear/mistrust,
as well as socioeconomic, language, and cultural factors [7-9]. Seminal work published over 15 years ago examining NIH cooperative group trials demonstrated similar underrepresentation of racial/ethnic minority subgroups [4]. Further work is thus needed to identify and address continued barriers of enrollment faced by minority groups, and to more fully understand the impact of NIH efforts on their enrollment.

The primary limitation of this study is the large number of trials (43%) not reporting any race or ethnicity and inconsistent race/ethnicity groupings across studies, and the associated bias this may introduce. Results should thus be interpreted with caution. In addition, although other covariates such as age, gender, and socioeconomic factors are known to correlate with race/ethnicity disparities [10, 11], these could not be fully evaluated given use of aggregate race data. Ongoing studies with granular patient-level data are thus needed to elucidate specific underrepresented patient populations. There are also limitations with the comparator SEER dataset which, while largely representative of the general US population, captures only 25-28% of incident cancer cases nationwide, and may underrepresent Black and Hispanic patients [12]. Nonetheless, to our knowledge this analysis represents the index aggregate utilization of federal public access datasets to assess domestic disparities in racial/ethnic phase 3 clinical trial enrollment. Although others have investigated FDA-approval trials [13], our use of US-limited large-scale public access datasets affords a benchmark for future serial assessments and US policy decisions.

In conclusion, over 40% of US cancer clinical trials fail to report race or ethnicity data. Of reporting trials, White patients continue to be overrepresented while racial/ethnic minority subgroups are underrepresented. Mandatory, standard, and granular reporting of racial and ethnic data elements should be considered in future iterations of ClinicalTrials.gov and other US cancer datasets, with continued efforts needed to ensure equitable clinical trial enrollment.

**Funding:** The authors report no funding for this work.
Disclosures: The authors report no financial disclosures or conflicts of interests related to this work.

Data Availability: All available data can be obtained by contacting the corresponding author.

Author Contributions: Conception and design—EL, CDF, CRT, SG; Data acquisition, analysis, and interpretation—EL, SG, TL, AM, WM, AE, AJ; Drafted initial manuscript—SG; Editorial oversight and correspondence—EL; Drafting, revising, editing manuscript—EL, SG, TL, AM, WM, AE, AJ, GW, CDF, CRT, BS, BAG, RJ

References

1. NIH Revitalization Act, Subtitle B, Part 1, Sec. 131-133, June 10, 1993. Available from: https://history.nih.gov/research/downloads/pl103-43.pdf, accessed February 10, 2020.

2. NIH Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research. Federal Register, Part VIII (59 FR 1450–14513), March 28, 1994. Available from: https://grants.nih.gov/grants/guide/rfa-files/rfa-ca-96-004.html, accessed February 10, 2020.

3. Freedman, L.S., et al., Inclusion of women and minorities in clinical trials and the NIH Revitalization Act of 1993—the perspective of NIH clinical trialists. Control Clin Trials, 1995. 16(5): p. 277-85; discussion 286-9, 293-309.

4. Murthy, V.H., H.M. Krumholz, and C.P. Gross, Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA, 2004. 291(22): p. 2720-6.

5. Pang, H.H., et al., Enrollment Trends and Disparity Among Patients With Lung Cancer in National Clinical Trials, 1990 to 2012. J Clin Oncol, 2016. 34(33): p. 3992-3999.

6. Sateren, W.B., et al., How sociodemographics, presence of oncology specialists, and hospital cancer programs affect accrual to cancer treatment trials. J Clin Oncol, 2002. 20(8): p. 2109-17.
7. Unger, J.M., et al., *The Role of Clinical Trial Participation in Cancer Research: Barriers, Evidence, and Strategies*. Am Soc Clin Oncol Educ Book, 2016. 35: p. 185-98.

8. Shavers, V.L. and M.L. Brown, *Racial and ethnic disparities in the receipt of cancer treatment*. J Natl Cancer Inst, 2002. 94(5): p. 334-57.

9. Ford, J.G., et al., *Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review*. Cancer, 2008. 112(2): p. 228-42.

10. Hutchins, L.F., et al., *Underrepresentation of patients 65 years of age or older in cancer-treatment trials*. N Engl J Med, 1999. 341(27): p. 2061-7.

11. Unger, J.M., et al., *Patient Income Level and Cancer Clinical Trial Participation: A Prospective Survey Study*. JAMA Oncol, 2016. 2(1): p. 137-9.

12. Yu, J.B., et al., *NCI SEER public-use data: applications and limitations in oncology research*. Oncology (Williston Park), 2009. 23(3): p. 288-95.

13. Loree, J.M., et al., *Disparity of Race Reporting and Representation in Clinical Trials Leading to Cancer Drug Approvals From 2008 to 2018*. JAMA Oncol, 2019: p. e191870.
**Table 1: Difference in Incidence by Race/Ethnicity (%)**

| Number of trials | White | | | Black | | | Hispanic | | | Asian | | |
|------------------|-------|---|---|-------|---|---|-------|---|---|-------|---|---|
|                  | Median | IQR | p* | Median | IQR | p* | Median | IQR | p* | Median | IQR | p* |
| All included trials | 96 | +6.8 | (+1.8, +10.1) | <.001 | -2.6 | (-5.1, +1.2) | .004 | -4.7 | (-7.5, -0.3) | <.001 | -4.7 | (-5.7, -3.3) | <.001 |
| Median year | | | | | | | | | | | | |
| <2005 | 15 | +6.5 | (+3.2, +10.4) | | -2.7 | (-4.6, +1.7) | | -1.5 | (-5.0, +0.4) | | -5.3 | (-5.7, -3.7) | |
| '05-'07 | 26 | +6.3 | (+3.7, +11.1) | .93 | -3.4 | (-7.5, -0.3) | .58 | -5.3 | (-7.9, -0.9) | .76 | -4.5 | (-5.6, -3.8) | .96 |
| '08-'10 | 35 | +6.1 | (+1.5, +9.6) | | -2.1 | (-4.2, +1.1) | | -4.7 | (-7.4, -0.4) | | -4.7 | (-5.8, -2.3) | |
| >2010 | 20 | +7.6 | (+0.4, +10.1) | | -2.6 | (-5.0, +2.0) | | -4.6 | (-9.2, -2.5) | | -4.6 | (-5.7, -3.3) | |
| Cooperative group | | | | | | | | | | | | |
| No | 34 | +5.0 | (-0.5, +11.0) | .25 | -1.4 | (-3.2, +1.8) | .24 | -1.8 | (-6.7, +0.4) | .23 | -4.7 | (-5.7, -2.2) | .48 |
| Yes | 62 | +7.8 | (+3.0, +10.0) | | -3.1 | (-5.7, +0.5) | | -4.9 | (-7.9, -2.8) | | -4.8 | (-5.7, -3.5) | |
| Industry funded | | | | | | | | | | | | |
| No | 60 | +7.6 | (+1.7, +9.9) | .89 | -2.7 | (-6.1, +0.9) | .69 | -5.2 | (-8.1, -0.4) | .16 | -4.9 | (-5.7, -3.3) | .67 |
| Yes | 36 | +6.3 | (+1.8, +10.4) | | -2.3 | (-4.6, +1.6) | | -3.4 | (-4.7, -0.3) | | -4.5 | (-5.8, -3.3) | |
| Disease Site | | | | | | | | | | | | |
| Breast | 18 | +8.0 | (+1.8, +10.4) | | -2.7 | (-6.7, +1.7) | | -4.7 | (-9.2, -3.4) | | -5.5 | (-7.1, -3.4) | |
| Colorectal | 5 | +7.7 | (+6.2, +11.5) | .35 | -4.1 | (-6.8, -3.0) | .38 | +4.3 | (-0.9, +9.6) | .26 | -5.1 | (-6.7, -4.1) | .27 |
| Lung | 5 | +3.3 | (+1.5, +5.5) | | -1.7 | (-3.0, +1.7) | | -4.7 | (-5.9, -3.4) | | -3.1 | (-4.5, -2.2) | |
| Prostate | 8 | +3.1 | (-0.2, +8.0) | | +1.6 | (-4.2, +5.1) | | -4.4 | (-5.2, +0.0) | | -4.2 | (-4.7, -2.6) | |
| Primary modality | | | | | | | | | | | | |
| Supportive care | 44 | +8.0 | (+1.8, +10.1) | | -3.9 | (-6.7, -1.3) | | -6.5 | (-8.8, -1.7) | | -5.0 | (-5.8, -3.4) | |
| Targeted systemic therapy | 26 | +8.8 | (+1.9, +11.5) | .10 | -2.7 | (-5.2, +1.3) | .09 | -4.1 | (-5.0, +0.0) | .19 | -4.4 | (-5.5, -3.8) | .56 |
| Cytotoxic chemotherapy | 19 | +3.4 | (+1.8, +10.0) | | +0.1 | (-3.0, +6.1) | | -2.7 | (-4.3, +0.1) | | -3.8 | (-5.7, -2.2) | |
| Radiation or surgery | 7 | +4.0 | (-2.3, +6.7) | | -2.2 | (-4.6, +2.0) | | -6.3 | (-9.2, +10.3) | | -4.4 | (-5.7, +1.2) | |

*Median differences in incidence by race/ethnicity (D-IRE) for each group with interquartile ranges (IQR) are displayed. P values for the top line (all included trials without further subgrouping) were calculated by a Wilcoxon signed rank test comparing median D-IRE to a value of zero (null hypothesis). All other p values were calculated by Mann Whitney U (for variables with two subgroups) or Kruskall-Wallis ANOVA (for variables with more than two subgroups).
**Figure title and legend**

**Figure 1:** Median difference in incidence by race/ethnicity (D-IRE) with interquartile range (IQR) shown with error bars (A), median ratio of incidence by race/ethnicity (R-IRE) with IQR shown with error bars (B), and median proportions of race/ethnicity compared to SEER estimates (C).
Figure 1 -- FINAL

A) Difference in incidence by race/ethnicity

B) Ratio of incidence by race/ethnicity

C) Median patient enrollment by race/ethnicity over time

---

Trial enrollment  SEER population