Trends in race and sex reporting in lung cancer phase III clinical trials

Faaiq Aslam[1], Rami Manochakian[2], Yanyan Lou[2], Gerardo Colon-Otero[3], and Taimur Sher[4]

[1]Mayo Clinic Alix School of Medicine
[2]Mayo clinic
[3]Mayo Clinic Florida
[4]Mayo Clinic

April 24, 2023

Abstract

Background: To understand the changing trends in reporting of race and sex as a demographic variable in phase III lung cancer clinical trials published over the last 35 years. Methods: A total of 426 articles reporting results of phase 3 lung cancer clinical trials published from 1984 to 2019 were identified in PubMed. Statistical analysis on trends over time on percentage of minority and female participation were performed. Results: Only 137 (32.2%) of the 426 studies analyzed reported race of participants. Among those studies, we found that the mean participation rate of white participants was significantly higher (82.65%) ($p < 0.001$). We found a decrease in African American participants and an increase in Asian participants over time. When looking at sex, we found that although the rate of male participation (69.02%) was significantly higher than that of female participation (30.98%), the female participation has improved with time at a rate of 0.65% per year. Conclusions: We found that the reporting and participation of minority races continues to lag that of other demographic factors like sex in phase III clinical trials in lung cancer. Especially in African Americans, where the participation in lung cancer phase III clinical trials has declined despite the rising incidence in lung cancer.

Title: Trends in race and sex reporting in lung cancer phase III clinical trials

Authors: Faaiq N Aslam [a], Rami Manochakian[b], Yanyan Lou [b], Gerardo Colon-Otero[b], Taimur Sher [b]

[a] Mayo Clinic Alix School of Medicine, 4500 San Pablo Rd S, Jacksonville, FL USA 32224. Email: aslam.faaiq@mayo.edu. ORCID: 0000-0001-9748-7545

[b] Mayo Clinic Florida, 4500 San Pablo Rd S, Jacksonville, FL USA 32224

Emails:

Faaiq N Aslam: aslam.faaiq@mayo.edu

Rami Manochakian: manochakian.rami@mayo.edu

Yanyan Lou: lou.yanyan@mayo.edu

Gerardo Colon-Otero: gcolonotero@mayo.edu

Taimur Sher: sher.taimur@mayo.edu

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Correspondence: Address to Taimur Sher MD, Mayo Clinic Florida, 4500 San Pablo Rd S. Jacksonville, FL 32224 (sher.taimur@mayo.edu)

Statements and Declarations

Ethics Approval and Consent to Participate

Not Applicable

Consent for Publication

Not Applicable

Availability of Data and Materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Author's Contributions

Faaiq N Aslam and Taimur Sher were instrumental in conceptualization, methodology, investigation, data curation, formal analysis, visualization, and writing—original draft.

Rami Manochakian, Yanyan Lou, and Gerardo Colon-Otero were instrumental in conceptualization, methodology, data curation, formal analysis, and writing—review & editing.

All authors read and approved the final manuscript.

Acknowledgements

Not Applicable

Abstract

Background: To understand the changing trends in reporting of race and sex as a demographic variable in phase III lung cancer clinical trials published over the last 35 years.

Methods: A total of 426 articles reporting results of phase 3 lung cancer clinical trials published from 1984 to 2019 were identified in PubMed. Statistical analysis on trends over time on percentage of minority and female participation were performed.

Results: Only 137 (32.2%) of the 426 studies analyzed reported race of participants. Among those studies, we found that the mean participation rate of white participants was significantly higher (82.65%) ($p < 0.001$). We found a decrease in African American participants and an increase in Asian participants over time. When looking at sex, we found that although the rate of male participation (69.02%) was significantly higher than that of female participation (30.98%), the female participation has improved with time at a rate of 0.65% per year.

Conclusions: We found that the reporting and participation of minority races continues to lag that of other demographic factors like sex in phase III clinical trials in lung cancer. Especially in African Americans, where the participation in lung cancer phase III clinical trials has declined despite the rising incidence in lung cancer.

Keywords: Lung cancer, clinical trials, disparities, NIH Revitalization Act, race, sex

1. Background

Clinical trials are an essential means of improving outcomes of cancer patients. Racial minorities and females have been underrepresented in clinical trials through time. Historically, such disparities have been a problem of multiracial societies such as US, but with globalization and human migration, health disparities are becoming a global concern. In February 2020 the European Parliament in its report on health inequalities in the European Union identified addressing growing health inequalities a priority [1]. The National Institute of Health (NIH) of the United States was tasked to address these growing challenges through the NIH Revitalization Act of 1993. An important mandate of the NIH Revitalization Act was adequate reporting on minority and vulnerable population in clinical trials. Despite these important efforts, various studies over the ensuing two decades continue to report persistent disparities in clinical trials [2-4].

As we enter the third decade of the 21st century, with increasing awareness and advancements in communications and global connectivity, an expectation would be that these disparities would be resolving. We set out to evaluate the landscape of race and sex reporting in phase III clinical trials in patients with lung cancer. We selected lung cancer for our study because as the leading cause of cancer death in the US, it has represented a significant proportion of phase III cancer studies (14%) and is seen across the entire spectrum of human races and sex [5-10].

The purpose of our study was to determine the historical trends and current landscape of reporting and representation of females and minorities in lung cancer phase III clinical trials. We chose to study lung cancer phase III clinical trials because lung cancer is the most common cause of cancer-related death and manifests differently depending on one's background, thus adequate reporting and representation are essential [5-8]. Specifically, we looked at participation rates of different races and sex over a 35-year time-period (1984-2019).

2. Methods

2.1 Data Collection

To obtain participation data for lung cancer phase III clinical trials, the publically available database, PubMed, was queried using the search terms "((lung[Title] OR pulmonary[Title]) AND (phase 3[Title] OR phase III[Title])) NOT (review) AND (clinicaltrial[Filter] OR randomizedcontrolledtrial[Filter]) AND (fft[Filter])" on April 22nd, 2020. This search yielded 724 publications, from which 426 publications were analyzed as they met the following inclusion criteria:

The study was a Phase III Lung Cancer Clinical Trial

The study did not present data from a previous clinical trial already recorded in our database (to avoid duplicate recording of data)

The 426 publications were subsequently used to collect year of publication and demographic data relating to participants' race and sex from the study's demographic table to construct the database for the analysis (Figure 1). Most of the studies the search yielded were multinational studies, which was determined by reviewing the author affiliations of these publications. The review of publically available data could not classify them into discrete categories of U.S. vs other regions. This study was considered exempt for IRB review per institutional policies.

Figure 1. CONSORT diagram illustrating the study selection process.

For race, the number of White, Asian, African American, and Hispanic participants were recorded. The category "Other" was also used to include any participants that were listed as Other in the original study or did not belong to the White, Asian, African American, or Hispanic groups. The final category of "Unknown" was used to include any participants that were listed as Unknown in the publication. The number of participants in each category were then converted to the percent of the total participant pool each category represented and were then used to compare participation rates across studies. The studies that did not report demographic data on racial background were annotated as such. For sex, the total number of males and females was recorded. These values were subsequently used to calculate the percent of the participants' pool

that was male or female, and the studies that did not report demographic data regarding sex were annotated as such. The data collected was then used to determine the percent of studies that reported participant race or sex, differences in the rate of participation among racial groups and between sex, and how participation rates of have changed over time.

2.2 Statistical Analysis

Statistical analysis was done using the SciPy stats package and figure generation was done using the MatPlotLib package for python [11, 12]. To determine the rate of reporting of race, the percentage of studies that reported race as a demographic and those that did not report race were calculated. Out of the studies that did report race, the percent of those studies that reported on each racial group was also determined. The same method was also used to determine the rate of reporting of sex.

To determine whether disparities in clinical trials' participation exist, the mean participation rate of each race and sex across all the studies that reported race or sex was calculated and the 95% confidence interval for the mean of each demographic was determined. Then, to compare participation rates among different races, the percent of white participation from each study was compared with the percent participation of the remaining races (Asian, African American, Hispanic, Other, and Unknown) using a 2-sample t-test and a p-value < 0.01 was considered statistically significant. A one-way ANOVA test was also used to determine whether a statistically significant difference in participation among the Asian, African American, Hispanic, Other, and Unknown groups exists and a p-value < 0.01 was considered statistically significant. To compare participation rates between sex, a 2-sample t-test was used and a p-value < 0.01 was considered statistically significant.

Lastly, to determine how participation rate has changed over time, the participation rates of different races from each individual study was plotted on a scatter plot using the MatPlotLib package [12]. Using those datapoints, the SciPy stats package was used to calculate the Pearson correlation coefficient, the p-value (with $p < 0.05$ considered statistically significant), and to draw a line of best fit [11]. The data points for White participation were plotted with each of the remaining races to serve as a comparison as to how participation rates have changed with time.

3. Results

3.1 Rates of race reporting and participation trends

We analyzed the demographic data from a total of 426 lung cancer phase III clinical trials. It was not possible for us to identify US only studies as most studies had authorships from multiple institutions across various regions in the US, Europe, and Asia in various combinations. Only 137 studies (32.2%) reported race as a demographic (Figure 2A) The remaining 289 studies (67.8%) did not report race. Interestingly, we found that from all the studies we analyzed, no study published before 1992 reported demographic data on patient race (Figure 2C). After 1992, more studies started reporting race (Figure 2C). However, race reporting did not appear to improve over time since 1993 (Figure 2C). Out of the 137 studies that reported demographic data on participants' race, 128 studies (93.4%) reported the rate of White participation, 97 studies (70.8%) reported the rate of Asian participation, 92 studies (67.2%) reported the rate of African American participation, 23 studies (16.8%) reported rate of Hispanic participation, 106 studies (77.4%) reported a race that fit into the other category, and 22 studies (16.1%) reported the rate of participants whose race was unknown (Figure 2B).

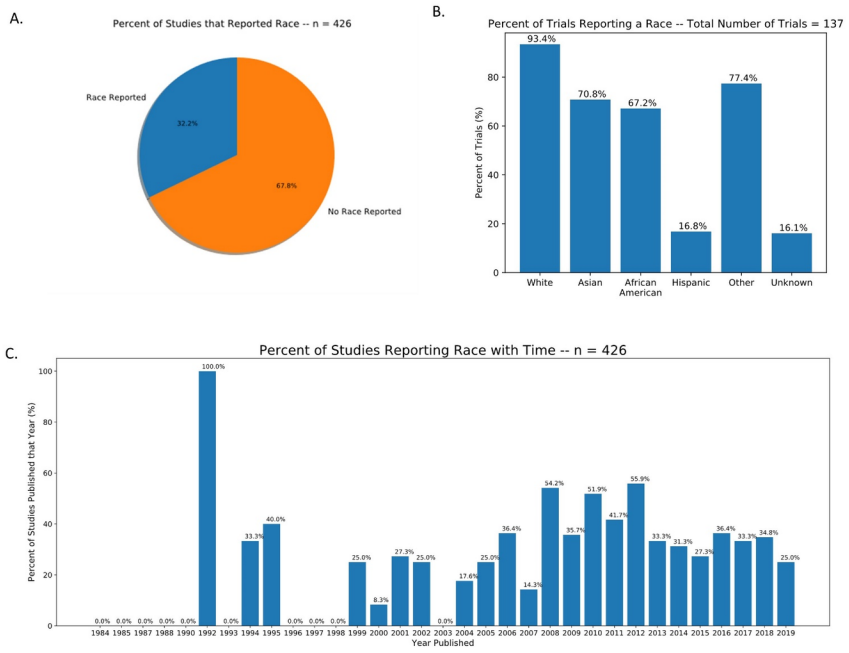

Figure 2. (A) Pie chart illustrating percentage of studies analyzed that reported race as a patient characteristic. (B) Histogram reporting the percentage of studies (of the studies that reported multiple races) that reported participation rates of each race category. (C) Histogram showing percent of studies published each year that reported demographic data on participant race.

The white participants represented 82.55% (95% CI 80.39%-84.71%) of study participants (Table 1) (Figure 3A). The mean rates of participation of Asian, African American, Hispanic, Others, and Unknown were 14.03% (95% CI 11.02%-17.04%), 5.53% (95% CI 4.50%-6.56%), 3.97% (95% CI 2.13%-5.81%), 3.99% (95% CI 3.21%-4.77%), and 2.46% (95% CI 0.85%-4.08%) respectively ($p < 0.001$) (Table 1) (Figure 3A).

Table 1. Mean Rates of Participation of each Race across analyzed Lung Cancer Phase III Clinical Trials.

Race	Mean Participation Rate of Each Race	95% Confidence Interval	Pearson Correlation Coefficient (Change in Participation rate with Time)	p-value
White	82.55%	80.39% - 84.71%	m = -0.51% per year r = -0.23, $p <$ 0.01	2 sample t-test (white vs. remaining races) $p < 0.001$
Asian	14.03%	11.02%-17.04%	m = 1.07% per year r = 0.279, $p < 0.01$	One-way ANOVA test (among remaining races) $p < 0.001$
African American	5.53%	4.50%-6.56%	m = -0.30% per year r = -0.34, $p <$ 0.001	

Race	Mean Participation Rate of Each Race	95% Confidence Interval	Pearson Correlation Coefficient (Change in Participation rate with Time)	p-value
Hispanic	3.97%	2.13%-5.81%	m = No Change with Time r = -0.01, p =0.97	
Other	3.99%	3.21%-4.77%	m = No Change with Time r = -0.04, p = 0.65	
Unknown	2.46%	0.85%-4.08%	m = No Change with Time r = 0.14, p = 0.55	

When looking at how participation rates changed with time, we found that the rate of white participants decreased at a rate of 0.51% per year (r = -0.23, $p < 0.01$) (Figure 3B). This decrease in participation was accompanied by an increase in the rate of Asian participants by a rate of 1.07% per year (r = 0.279, $p < 0.01$) (Figure 3C). However, the rate of African American participants decreased at a rate of -0.30% per year (r = -0.34, $p < 0.001$) (Figure 3D). The rates of participation among Hispanics (r = -0.01, p = 0.97), the other category (r = -0.04,p = 0.65), and the unknown category (r = 0.14, p = 0.55) did not change with time (Figure 3E, 3F, 3G), although we did not have enough studies report Hispanic demographics to draw meaningful conclusions about their representation.

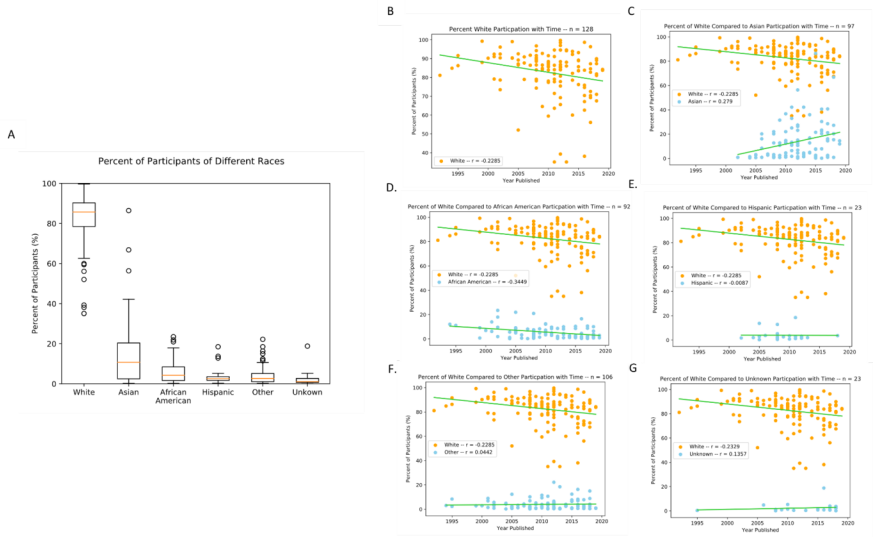

Figure 3. (A) Box Plots illustrating the percentage of participants of each race in all the studies analyzed that reported race. (B) A scatter plot with a Pearson correlation coefficient of r = -0.23 illustrating how the percent of white participants has decreased with time. (C) A scatter plot with a Pearson correlation coefficient showing that the percent of Asian participants has increased with time (r = 0.279) and percent of white participants has decreased with time (r = -0.23). (D) A scatter plot with a Pearson correlation coefficient showing that the percent of African American participants has decreased with time (r = -0.34) and percent of white participants has decreased with time (r = -0.23). (E) A scatter plot with a Pearson

correlation coefficient showing that the percent of Hispanic participants has not changed with time (r = -0.01) and percent of white participants has decreased with time (r = -0.23). (F) A scatter plot with a Pearson correlation coefficient showing that the percent of other participants has not changed with time (r = -0.04) and percent of white participants has decreased with time (r = -0.23). (G) A scatter plot with a Pearson correlation coefficient showing that the percent of unknown participants has not changed with time (r = 0.14) and percent of white participants has decreased with time (r = -0.23).

3.2 Rates of sex reporting and participation trends

From the 426 studies that were analyzed, 420 studies (98.8%) reported sex as a patient demographic (Figure 4A). Among those 420 studies, the mean rate of participation for males was 69.02% (95% CI 67.66% - 70.38%) and the mean for females was 30.98% (95% CI 29.62% - 32.34%), yielding a statistically significant difference in participation rates between males and females (p <0.001) (Table 2) (Figure 4B).

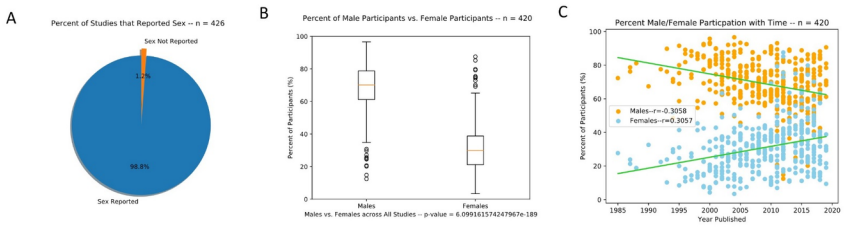

Figure 4. (A) Pie chart illustrating percentage of studies analyzed that reported sex as a patient characteristic. (B) Box Plots illustrating the percentage of participants of each sex in all the studies analyzed that reported sex (p-value < 0.001). (C) A scatter plot with a Pearson correlation coefficient showing that the percent of male participants has decreased with time (r = -0.31) and percent of female participants has increased with time (r = 0.31).

When looking at the change of participation rates over time, we found that the disparity between male and female participation has significantly improved. The rate of male participation has decreased by a rate of 0.65% a year since 1985 (r = -0.31, p <0.001) and the rate of female participation has increased by a rate of 0.65% a year since 1985 (r = 0.31, p < 0.001) (Figure 4C). Moreover, we found that the mean rate of participation from 2015-2019 for males was 65.21% and the mean rate of participation for females was 34.79%, which is markedly improved from the mean rate of participation from 1985-1995 when the mean participation rate was 73.60% for males and 26.40% for females (Table 2).

Table 2. Mean Rates of Participation of each Sex across all analyzed Lung Cancer Phase III Clinical Trials.

Sex	Mean Participation rate across all studies	95% Confidence Interval	p-value	Pearson Correlation Coefficient (Change in Participation rate with Time)	Mean from 1985-1995	Mean from 2015-2019
Male	69.02%	67.66% - 70.38%	2 sample t-test (Males vs. Females) p < 0.001	m = -0.65% per year r = -0.31, p < 0.001	73.60%	65.21%
Female	30.98%	29.62% - 32.34%		m = 0.65% per year r = 0.31, p < 0.001	26.40%	34.79%

4. Discussion

Our study is the first one to look at the reporting trends in race and sex in phase III lung cancer trials over last 35 years. The results of our analysis further substantiate the evidence that continues to highlight the disparities surrounding race in clinical trials.

Examining the trends in reporting of sex, we find more positive change over time. In the studies we reviewed, 98.8% reported participant sex. Analyzing these trends, we noted progressive increase in women representation since 1985. While the trends over time are promising, females are still underrepresented with a mean participation rate of 34.97% from 2015-2019. This is despite females accounting for 48.9% of lung cancer in the United States [13]. Similar patterns have also been seen in other, non-cancer clinical trials as well. For example, cardiovascular and lipid lowering agent clinical trials also showed underrepresentation of females at a participation rate of 33%-38% [14, 15].

We were surprised to find that no phase III lung cancer study, in our database, between 1984-1992 reported on race. In reviewing the literature for evidence, we encountered significant variability in race reporting across different studies. As an example, Brahan et al. found significant increases in race reporting from 29.6% to 63.5% in pediatric studies conducted in 1991-1993 and 2000-2002 [16]. In a 2011 study, Geller et al. looked at 86 clinical trials across nine high-impact journals and found that one-fifth of the studies failed to report the distribution of participant race [17]. Unlike the findings from these studies, our findings showed that a significantly larger proportion of lung cancer phase III clinical trials did not report the distribution of participant race. There can be several reasons for such degree of variance. Pediatric clinical studies, especially pediatric oncology studies, have more centralized review and sponsoring processes, and are more likely to be funded by the NIH. Logically, such studies are more likely to comply with the NIH policies including the provisions of the NIH Revitalization Act and hence the significant change in the race and sex reporting as noted by Brahan et al. This increase in race reporting, while encouraging, falls short of the ideal as a third of pediatric oncology studies still did not report race as a variable. Another reason for lack of reporting on race may be that majority of the studies are multinational in nature and may include geographical areas with less cultural heterogeneity. This is likely an important limitation of our study. However, the rapid globalization and human migration over the last two decades calls for a more acute evaluation. The results of our study in the context of recent demographic changes highlight the significance of addressing this trend as health inequality and disparities are becoming more of a global issue.

Our findings of minority underrepresentation in lung cancer phase III clinical trials have also been observed in other studies as well. A 2004 study looking at race-based disparities in cancer clinical trials also found White participants to represent 85.6% of all participants and another study found White participants to represent 76.3% of all participants in cancer drug clinical trials [4, 18]. One might question the generalizability of this finding in the context of the demographic make-up of the population and the epidemiology of the disease under study. For lung cancer, this is especially important as the incidence in African Americans, the largest minority population in the United States, is 59.5 per 100,000 compared to 61.6 per 100,000 in Whites [13]. Loree et al in 2019 evaluated the reporting of race and other demographic variables in the landmark clinical trials leading to the approval of antineoplastic therapies and found a similar finding of decline in African American participation between 2008 and 2018 [18].

The principles of justice, beneficence and the medicine's professional code of conduct require diversity and inclusion as mission critical to its social contract with the society. These higher moral values call out for all efforts to ensure adequate access of all segments of society to the benefits of biomedical research. While it may be debatable that non-reporting is not synonymous with not-being included, it is strong enough evidence to support the conclusion of non-inclusion. Although positive changes have been made in the reporting of sex, there is clearly room for much improvement when it comes to the reporting of race. The responsibility to move the needle forward falls on shoulders of all stakeholders. The clinicians and patient care teams should review and discuss the importance of race and sex reporting with the patients in the broader context of its value to the society and in addressing the disparities. The clinical trials' sponsors should work with the study teams to ensure recording sex and race amongst other demographic information in case report

forms. The medical publishers should mandate the reporting of race and sex as a criterion for acceptance of publication.

Our study does have *limitations* . Given that this study was a retrospective analysis, it carries the same weaknesses associated with such studies. As noted in the methods section, we could not reliably decipher which studies were solely U.S. based versus from other regions as majority of the manuscripts had author affiliations from various institutions across the globe.

5. Conclusions

In conclusion, we found that disparities in reporting and participation of minority racial groups continue to exist. We believe that requiring reporting of patient demographic data in clinical trials, in compliance with the NIH Revitalization Act, will improve transparency and the rate of reporting. Furthermore, it will critically inform health policy makers and other stakeholders to address the root causes of underrepresentation of minorities by addressing social barriers and hesitancy to participate in clinical trials [19].

References

1. Addressing Health Inequalities in the European Union. 2020.

2. Geller SE, Koch AR, Roesch P, Filut A, Hallgren E, Carnes M. The More Things Change, the More They Stay the Same: A Study to Evaluate Compliance With Inclusion and Assessment of Women and Minorities in Randomized Controlled Trials. Acad Med. 2018;93(4):630-5. doi: 10.1097/ACM.0000000000002027.

3. Duma N, Vera Aguilera J, Paludo J, Haddox CL, Gonzalez Velez M, Wang Y, et al. Representation of Minorities and Women in Oncology Clinical Trials: Review of the Past 14 Years. J Oncol Pract. 2018;14(1):e1-e10. doi: 10.1200/JOP.2017.025288.

4. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA. 2004;291(22):2720-6. doi: 10.1001/jama.291.22.2720.

5. Zavala VA, Bracci PM, Carethers JM, Carvajal-Carmona L, Coggins NB, Cruz-Correa MR, et al. Cancer health disparities in racial/ethnic minorities in the United States. Br J Cancer. 2021;124(2):315-32. doi: 10.1038/s41416-020-01038-6.

6. de Groot P, Munden RF. Lung cancer epidemiology, risk factors, and prevention. Radiol Clin North Am. 2012;50(5):863-76. doi: 10.1016/j.rcl.2012.06.006.

7. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol. 2016;893:1-19. doi: 10.1007/978-3-319-24223-1_1.

8. Subramanian J, Madadi AR, Dandona M, Williams K, Morgensztern D, Govindan R. Review of ongoing clinical trials in non-small cell lung cancer: a status report for 2009 from the ClinicalTrials.gov website. J Thorac Oncol. 2010;5(8):1116-9. doi: 10.1097/JTO.0b013e3181e76159.

9. Jung KJ, Jeon C, Jee SH. The effect of smoking on lung cancer: ethnic differences and the smoking paradox. Epidemiol Health. 2016;38:e2016060. doi: 10.4178/epih.e2016060.

10. Meza R, Meernik C, Jeon J, Cote ML. Lung cancer incidence trends by gender, race and histology in the United States, 1973-2010. PLoS One. 2015;10(3):e0121323. doi: 10.1371/journal.pone.0121323.

11. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261-72. doi: 10.1038/s41592-019-0686-2.

12. Hunter JD. Matplotlib: A 2D graphics environment. Computing in science & engineering. 2007;9(03):90-5.

13. Cancer Statistics Center. http://cancerstatisticscenter.cancer.org Accessed.

14. Khan SU, Khan MZ, Raghu Subramanian C, Riaz H, Khan MU, Lone AN, et al. Participation of Women and Older Participants in Randomized Clinical Trials of Lipid-Lowering Therapies: A Systematic Review. JAMA Netw Open. 2020;3(5):e205202. doi: 10.1001/jamanetworkopen.2020.5202.

15. Jin X, Chandramouli C, Allocco B, Gong E, Lam CSP, Yan LL. Women's Participation in Cardiovascular Clinical Trials From 2010 to 2017. Circulation. 2020;141(7):540-8. doi: 10.1161/CIRCULATIONAHA.119.043594.

16. Brahan D, Bauchner H. Changes in reporting of race/ethnicity, socioeconomic status, gender, and age over 10 years. Pediatrics. 2005;115(2):e163-6. doi: 10.1542/peds.2004-1437.

17. Geller SE, Koch A, Pellettieri B, Carnes M. Inclusion, analysis, and reporting of sex and race/ethnicity in clinical trials: have we made progress? J Womens Health (Larchmt). 2011;20(3):315-20. doi: 10.1089/jwh.2010.2469.

18. Loree JM, Anand S, Dasari A, Unger JM, Gothwal A, Ellis LM, et al. Disparity of Race Reporting and Representation in Clinical Trials Leading to Cancer Drug Approvals From 2008 to 2018. JAMA Oncol. 2019;5(10):e191870. doi: 10.1001/jamaoncol.2019.1870.

19. Jayakrishnan T, Aulakh S, Baksh M, Nguyen K, Ailawadhi M, Samreen A, et al. Landmark Cancer Clinical Trials and Real-World Patient Populations: Examining Race and Age Reporting. Cancers (Basel). 2021;13(22). doi: 10.3390/cancers13225770.