Disparities in Clinical Trial Access Across US Urban Areas

Yevgeniy Feyman, BA; Frank Provenzano, PhD; Frank S. David, MD, PhD

Introduction

A lack of geographically convenient clinical trial sites hinders study participation, particularly for rural residents. But even in more densely populated regions, where most US residents live, trial locations may not be optimized to maximize access. In this cross-sectional study, we explored whether availability of trial seats in nonrural regions varies with population, racial and socioeconomic compositions, research activity, or amount of health care infrastructure.

Methods

We downloaded all records from ClinicalTrials.gov on August 29, 2016. Using automated methods and manual curation, we identified actively recruiting, interventional, industry-supported drug trials in phase 1/2 through phase 3 with at least 1 US site and an unambiguous number of total study locations. Records were downloaded on October 12, 2017, to update the sites and projected enrollment for each trial and exclude dormant trials. This yielded 41 770 nonunique US sites for 1926 trials (median, 18 sites/trial; interquartile range [IQR], 3-64 sites/trial).

Figure. Scatter Plot for 171 Combined Statistical Areas

![Scatter Plot](image_url)

Table. Subtractive Model of Clinical Trial Seats per CSA

<table>
<thead>
<tr>
<th></th>
<th>Bivariate</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of CSAs</td>
<td>171</td>
<td>171</td>
<td>171</td>
<td>171</td>
<td>171</td>
<td>171</td>
<td>171</td>
</tr>
<tr>
<td>Population/1000 (SE)</td>
<td>1 (-0.03)</td>
<td>0.88 (-0.393)</td>
<td>0.86 (-0.402)</td>
<td>0.895 (-0.399)</td>
<td>0.96 (-0.113)</td>
<td>0.902 (-0.395)</td>
<td></td>
</tr>
<tr>
<td>Median income/$1000 (SE)</td>
<td>157 (-22.2)</td>
<td>-4.6 (-10.3)</td>
<td>11.5 (-14.2)</td>
<td>-6.13 (-8.89)</td>
<td>-5.85 (-13.5)</td>
<td>-0.708 (-11.2)</td>
<td></td>
</tr>
<tr>
<td>White, % (SE)</td>
<td>-78.35 (-17.44)</td>
<td>-3.699 (-5.215)</td>
<td>-9.167 (-7.33)</td>
<td>-4.431 (-4.325)</td>
<td>-3.422 (-5.902)</td>
<td>-4.076 (-5.221)</td>
<td></td>
</tr>
<tr>
<td>Hospital beds, No. (SE)</td>
<td>0.435 (-0.016)</td>
<td>0.0352 (-0.202)</td>
<td>0.393 (-0.072)</td>
<td>0.0424 (-0.205)</td>
<td>0.0318 (-0.203)</td>
<td>0.0424 (-0.192)</td>
<td></td>
</tr>
<tr>
<td>NIH funding/$1000 (SE)</td>
<td>0.03 (-0.00196)</td>
<td>0.00208 (-0.00305)</td>
<td>0.00299 (-0.00335)</td>
<td>0.00188 (-0.00312)</td>
<td>0.00213 (-0.003)</td>
<td>0.00215 (-0.00281)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.873</td>
<td>0.855</td>
<td>0.872</td>
<td>0.872</td>
<td>0.872</td>
<td>0.872</td>
<td>0.872</td>
</tr>
</tbody>
</table>

Abbreviations: CSA, combined statistical area; NIH, National Institutes of Health.

Open Access. This is an open access article distributed under the terms of the CC-BY-NC-ND License.
We obtained zip codes for more than 98% of sites (41,279/41,770), imputing as needed by matching city, state, or institution name. We mapped zip codes to combined statistical areas (CSAs), each of which contains adjacent metropolitan, micropolitan, or both types of areas, and surrounding counties that share social and economic linkages.4,5 Because crosswalk files from zip codes to CSAs do not exist, we used a 2-step process involving core-based statistical areas as an intermediate. Our final analysis comprised 35,149 nonunique sites in CSAs from 1888 trials. To calculate mean seats per site, we divided each trial’s target enrollment by its total number of sites.

We obtained data on population (in 2016), median household income (in 2015), and racial demographics (in 2016) from the US Census; data on funding (in 2018) from the National Institutes of Health (NIH); and data on US inpatient hospital beds (in 2014) from the American Hospital Association’s annual survey. We mapped institutions to CSAs via zip codes as indicated above.

We performed linear regression and tabulations in Stata MP version 13 (StataCorp). Underlying data are available at Open Science Framework under a CC-BY-4.0 license.

The study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. This study was based on publicly available data that did not include patient contact or medical record review, and therefore did not require institutional review board approval or informed consent.

Results

Across 171 CSAs in the United States, trial seats per 1000 residents ranged from 0 to 3.83 (median, 0.64; IQR, 0.25-1.01). In a bivariable model, population size explained 87% of the variation in seats per CSA (Figure). In a multivariable model, median income, percentage of the population identified as white, number of hospital beds, and NIH funding explained scant remaining variability and were not statistically significant (Table). We obtained similar results when we specifically focused on cancer studies (n = 1012), early-stage studies (phase 1/2 and phase 2; n = 1206), and late-stage studies (phase 2/3 and phase 3; n = 682) (data not shown).

Discussion

Unequal distribution of trial sites relative to population may impede patients’ participation in clinical studies and contribute to the well-documented low representation of minorities in certain studies.6 We found that population size explained 87% of the variability in the number of seats in industry-supported studies across nonrural regions throughout the United States. Accounting for racial demographic categories, income, NIH funding, and hospital beds did not substantively improve the bivariable model’s explanatory power. Thus, we conclude that the number of available seats in industry-sponsored trials across CSAs is largely explained by population size.

Our work has several limitations. Indication-specific differences in seat access and the distances patients are willing to travel to participate in trials were not examined. Although our methods minimized selection bias, our focus on industry-sponsored studies in CSAs limits the generalizability of our findings. In addition, the small number of CSAs (geographic units) limited our ability to explore other factors. Lastly, we could not explore time trends in trial access in this cross-sectional analysis.

Notwithstanding these limitations, we demonstrated that population size is the main determinant of the availability of industry-sponsored clinical trial seats in nonrural areas. Factors related to trial recruitment and retention, not site selection per se, are most likely responsible for remaining racial and socioeconomic inequities in trial access and participation in these regions.
ARTICLE INFORMATION
Accepted for Publication: January 7, 2020.
Open Access: This is an open access article distributed under the terms of the CC-BY-NC-ND License. © 2020 Feyman Y et al. JAMA Network Open.
Corresponding Author: Frank S. David, MD, PhD, Pharmagellan LLC, 499 Adams St, PO Box 94, Milton, MA 02186 (frank@pharmagellan.com).
Author Affiliations: Pharmagellan LLC, Milton, Massachusetts (Feyman, Provenzano, David); Boston University School of Public Health, Boston, Massachusetts (Feyman); Department of Neurology, Columbia University, New York, New York (Provenzano).
Author Contributions: Dr David had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Mr Feyman and Dr Provenzano contributed equally to the article.

Concept and design: All authors.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: All authors.
Critical revision of the manuscript for important intellectual content: Feyman, David.
Statistical analysis: Feyman.
Administrative, technical, or material support: Provenzano, David.
Supervision: David.

Conflict of Interest Disclosures: Mr Feyman and Dr Provenzano reported receiving personal fees from
Pharmagellan during the conduct of the study. Dr Provenzano reported holding patents and having pending patents; serving as a consultant to Imij Technologies; and having equity in Imij Technologies. Dr David reported receiving grants from the Biotechnology Industry Organization; having a patent pending; serving as a full-time biopharmaceutical industry consultant through the advisory firm Pharmagellan; earning most of his income by providing strategic advisory services to various companies in the biotech, pharmaceutical, and device industries; being a part-time employee of Brigham and Women’s Hospital in the Brigham Research Institute; and that his wife is employed by and expects to receive stock compensation from Sanofi Genzyme.

Additional Contributions: We thank Julie Lin, MD, MPH (Sanofi Genzyme), for helpful comments, and Matthew Murphy (Accenture) for assistance with background research. Dr Lin did not receive compensation for her contributions. Mr Murphy was a part-time employee of Pharmagellan at the time of his contributions.

Additional Information: US Census data were obtained from https://factfinder.census.gov and National Institutes of Health data were obtained from https://report.nih.gov. Zip code to CSA crosswalks were derived from data from the US Department of Housing and Urban Development (https://www.huduser.gov/portal/datasets/usps_crosswalk.html) and the US Census Bureau (https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/delineation-files.html).

REFERENCES