Research Article

CLINICAL INVESTIGATION

Geographic distribution of clinical trials may lead to inequities in access

Background: We sought to describe the geographic distribution of clinical trial sites across the continental USA and to identify drivers of trial site location. **Methods/Results**: Locations of 174,503 clinical trial sites were collected from 2002–2007 from the US FDA's Bioresearch Monitoring Information System and geo-coded for spatial analysis. Predictors examined included urban population percentage (2000 US Census) and number of healthcare/social service and educational establishments (2002 Economic Census) per zip code. Extensive clustering of trial sites was detected. Urban composition and healthcare/social service facilities were strong predictors of the number of trial sites per zip code ($p < 0.0001$; $R^2 = 0.69$), but not their location (only 27% of clusters explained by these covariates). **Conclusion:** US clinical trial sites are highly clustered around urban areas with healthcare/social service facilities, which may partly explain why rural communities are underrepresented in clinical research.

Keywords: clinical trial access • clinical trial disparities • clinical trial distribution • geographic disparities in clinical trials • urban distribution of clinical trials

Elizabeth M Seidler[1], Aparna Keshaviah[2], Case Brown[3], Elizabeth Wood[4], Lisa Granick[4] & Alexandra B Kimball[*,5]

[1]Emory University School of Medicine, Atlanta, GA, USA
[2]Massachusetts General Hospital, Boston, MA, USA
[3]P-REX/Massachusetts Institute of Technology, Cambridge, MA, USA
[4]Fletcher Spaght, Inc., Boston, MA, USA
[5]Harvard Medical School, Boston, MA, USA
[*]Author for Correspondence:
Tel.: +1 617 726 5066
harvardskinstudies@partners.org

Background

Clinical trials should be easily and equally accessible to all populations for multiple reasons, including maintaining scientific integrity, ethical standards and, in some cases, access to novel treatments that may not otherwise be available. However, certain populations, such as racial and ethnic minorities, uninsured, socioeconomically disadvantaged, elderly and rural populations are often underrepresented in clinical research and can be difficult to recruit into clinical trials [1–3]. A host of economic, social, cultural and medical barriers to clinical trial accrual have been suggested, including lack of awareness of clinical trials, unequal access to the healthcare system, provider bias, mistrust of clinical research, poor past experiences with the healthcare system, poverty, substance abuse, homelessness, lack of insurance, and high comorbidity rates [1,4–6]. Some of these issues have been the focus of research and outreach efforts, including the 1993 NIH guidelines for inclusion in clinical research, which necessitates enrollment of women and racial/ethnic minorities [7]. Unfortunately, these guidelines do not extend to geographic minorities – that is, rural populations. While some efforts have been made to enroll underrepresented rural populations in clinical trials by collaborating with community-based, rural physicians [8–12], successful implementation of large-scale initiatives to improve rural enrollment in clinical trials is lacking.

In addition to the multiple barriers to participation listed above, another important and perhaps overlooked contributor may be physical distance to clinical trial sites. To test this theory, we sought to examine the geographic distribution of clinical trials across the continental USA, and to identify factors associated with the number of clinical trial sites in a given area. Our hypothesis was that, despite efforts to broaden access to clinical trials, clinical trial sites would be clustered around urban areas and areas with a greater presence of academic institutions and established healthcare and/or social services.

FUTURE SCIENCE part of fsg

Methods

Institutional Review Board approval with exempt status was obtained through Massachusetts General Hospital (MA, USA).

Datasets

The US FDA Bioresearch Monitoring Information System (BMIS) was used to collect geographic information on clinical trial sites. The system is publicly available and contains information submitted to the FDA identifying clinical investigators, contract research organizations, and institutional review boards involved in investigational new drug studies with human investigational drugs [20]. BMIS data are abstracted from FDA Forms 1571 and 1572 and other pertinent documents contained in investigational new drug submissions. The fields in the dataset include reviewer ID; the name, rank, and degree of the clinical investigator or institutional review board's chairperson; clinical trial site street address, city, state, zip code, and country; receipt date; and type of entry (clinical investigator, contract research organization or institutional review board).

Between 1 November, 2002 and 1 November, 2007, the BMIS contained data on 263,798 clinical trial site locations. Of these, 89,244 trial sites were excluded because they were located outside the USA (n = 88,634), were in Hawaii or Alaska (n = 497), had a missing zip code and thus could not be geo-coded (n = 70), or had a zip code of 000 (n = 43), which is used to de-identify 3-digit zip codes with fewer than 20,000 people (thus, 000 does not represent a single location). When possible, erroneous or missing zip codes were corrected using the street address.

The remaining 174,554 clinical trial sites were aggregated for spatial analysis by 3-digit zip code tabulation area (ZCTA). ZCTAs were introduced in the 2000 census in order to align census data tabulation to zip code areas. The aggregated clinical trial data were merged with a geo-coded dataset from GIS (ESRI Business Information Solutions, 2004) that contained total population in 2003 (within the timeframe data were gathered) for 877 identifiable 3-digit ZCTA regions. When the aggregated clinical trial data were merged with this geo-coded dataset, an additional 51 trial sites were excluded due to missing data on total population, leaving a total of 174,503 analyzed clinical trial site locations.

Aggregation by 3-digit zip code was chosen over 5-digit zip code primarily to examine disparities in clinical trial access on a large-scale basis. Such an approach is conservative with regard to analysis of spatial clustering, because very small rural areas are aggregated with larger neighboring areas within the same 3-digit ZCTA. Additionally, analysis by 5-digit ZCTA resulted in areas with many trial sites but very small populations (often 0) due to the fact that a number of hospitals and research centers comprise an entire 5-digit ZCTA. Thus, in clustering analyses that adjust for population, such areas would result in instability of estimates.

Geo-coded data (by 3-digit ZCTA) were gathered for four factors that we hypothesized would influence clinical trial site location. We used data from ESRI for total population in 2003, data from the 2000 US Census to calculate the urban population percentage (derived from total population, urban population, and rural population counts), and data from the 2002 Economic Census [21] to obtain estimates of the number of healthcare and/or social service facilities (sector 62), and the number of full-time educational establishments (sector 61; note: it was not possible to subset the detailed data to include only colleges, universities and professional schools; that is, code 6113). A total of 856 of the 877 identifiable 3-digit ZCTA regions had complete covariate data.

Since clinical trial data from BMIS did not contain descriptive information about the types of trials conducted, as a proxy, we examined data from clinical trials registered at ClinicalTrials.gov that had at least one site in the USA during the same time frame. ClinicalTrials.gov is a publicly available registry of clinical trials, maintained by the NIH, that contains information on who is sponsoring or funding the trial, the phase of the trial, whether it is interventional or observational, and whether it provides access to investigational drugs outside of clinical trials [22]. Only trials of drugs, biologics and devices are required to be registered, but additional trials may also be registered to provide up-to-date information to patients, family members, researchers and healthcare providers about clinical trials being conducted [23]. This proxy dataset contained 23,820 multisite trials conducted at a total of 144,856 US trial sites, but lacked detailed information on location that would allow for geo-coding. This dataset was therefore used solely to describe a sample of the types of trials being conducted in the time frame analyzed.

Statistical Methods

The spatial scan statistic [13] was employed to determine whether or not the locations of clinical trial sites across the continental USA exhibited complete spatial randomness. This method identified significant clusters of trial sites by continually moving a circle of varying radius around the map and comparing the observed number of clinical trial sites within the circle with the expected number of trial sites based

on the total population within the circle, and adjusting for multiple comparisons. Because the spatial scan statistic requires point data, each 3-digit ZCTA was represented by its centroid (spatial center). We first conducted a simple population-adjusted clustering analysis. Subsequently, we adjusted for urban population percentage, number of healthcare and/or social service facilities, and number of full-time educational establishments to determine if the clustering of clinical trial sites could be explained by these covariates. Significance was based on a Poisson model, and both high and low clusters were identified.

To determine if the magnitude of clinical trial sites per 3-digit ZCTA could be explained by the aforementioned covariates, we employed linear regression modeling. An ordinary least squares model revealed significant spatial correlation of the residuals based on the Breusch-Pagan test ($p < 0.0001$), and so we examined the fit of spatial regression models. Based on the lack of significance of the Lagrange multipliers ($p = 0.18$), we concluded that a spatial lag model [14] was a better fit to the data than a spatial error model ($p < 0.0001$). The weight matrix describing the nature of spatial dependence was chosen to ensure that all zip codes had at least one neighbor. In a sensitivity analysis that examined other weight matrices, the results were unchanged. Because the distribution of the number of clinical trial sites was highly skewed, the square root transformation was applied to improve symmetry.

Data cleaning and summary was conducted using SAS version 9.2. Clustering analyses were generated using SatScan version 9.0.1, and spatial regression was conducted in GeoDa version 9.8 [15]. All mapping was performed using arcGIS version 9.3. In all analyses, a two-sided alpha = 0.05 significance level was applied.

Results

A total of 763 (87%) out of the 877 identifiable 3-digit ZCTAs in the continental USA were home to at least one clinical trial site. The mean number of trial sites in a 3-digit ZCTA was 204, with a median of 43 and a range of 0 to 3,632 trial sites (Table 1). The mean and median total populations were 337,008 and 209,421, respectively, and ranged from 561 to 2,873,731. The mean and median urban population percentages were 66 and 63%, respectively, and ranged from 0 to 100%. The mean and median numbers of healthcare and/or social service facilities were 814 and 498, respectively, and ranged from 1 to 6520. The mean and median numbers of full-time educational establishments were 44 and 21, respectively, and ranged from 0 to 653 (Table 1). No further descriptive information on the clinical trial sites analyzed was available from the BMIS. However, the majority of clinical trial sites reported to ClinicalTrials.gov from 2002–2007 were industry sponsored (66%), Phase III (41%), interventional (94%) and drug (67%) trials.

Clustering analyses adjusting for total population detected 340 significant high or low clusters (Figure 1). High clusters (i.e., clusters with more trial sites than expected given the population) were largely located in urban areas, and Boston, MA (zip code 022) emerged as an extreme outlier, with a standardized trial rate of 185 trial sites per 10,000 people (2.8-times higher than the next highest rate).

After accounting for urban composition, number of healthcare and/or social service facilities, and number of full-time educational establishments, 248 significant high or low clusters remained; that is, these covariates together explained 27% of the clustering observed. The zip code 985 (which included Olympia, WA) had the largest cluster of clinical trial sites ($p < 0.0001$), with over nine-times more trial sites observed ($n = 1395$) than expected ($n = 149$) based on population size (418,858), population density (59.1% urban), healthcare and/or social service facilities (1101 facilities), and educational institutions (51 full-time establishments).

The spatial lag regression model used to examine drivers of the magnitude of clinical trial sites was able to account for the spatial autocorrelation in the data,

Table 1. Summary of covariates per 3-digit zip code tabulation area.

Variable Analyzed	Mean	SD	Minimum	Median	Maximum
Trial sites (n)	204	431	0	43	3632
Total population	337,008	357,162	561	209,421	2,873,731
Urban population (%)	66	26	0	63	100
Healthcare and/or social service facilities (n)	814	894	1	498	6520
Full-time educational establishments (n)	44	61	0	21	653

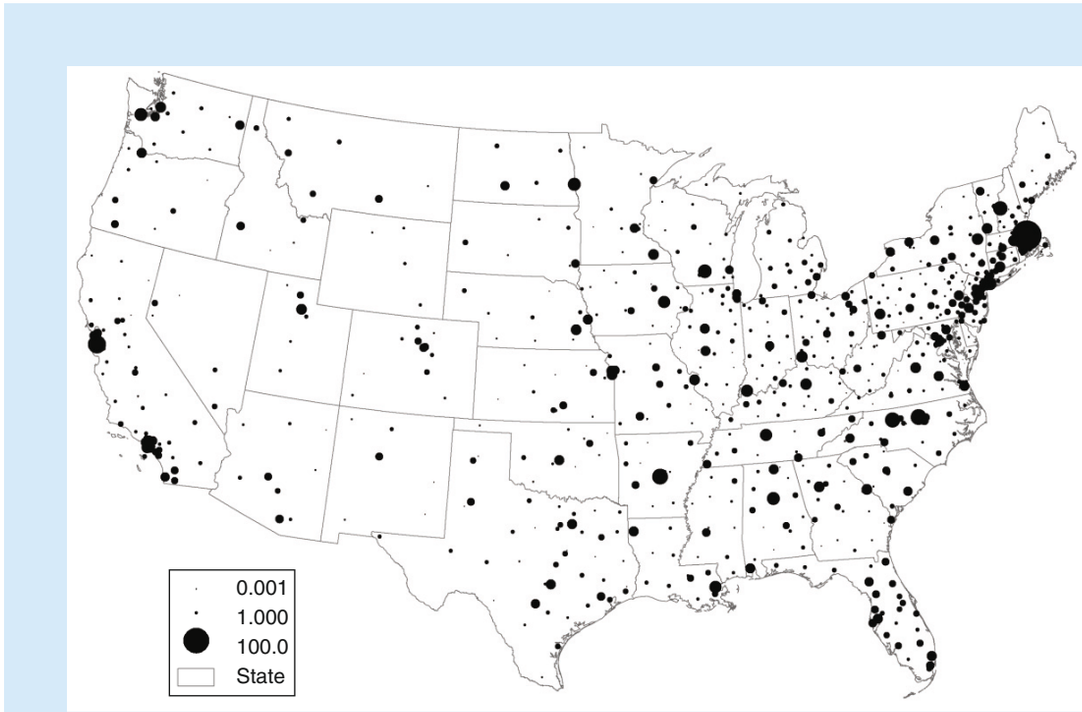

Figure 1. Number of clinical trial sites per 10,000 persons (2002–2007), by 3-digit zip code tabulation area. Circles are proportional to the number of clinical trial sites per 10,000 persons and are plotted at the centroid (spatial center) of each 3-digit zip code tabulation area. Data are based on 174,503 trial sites in the continental USA reported from 1 November, 2002 to 1 November, 2007 in the Bioresearch Monitoring Information System. Data taken from [12].

as evidenced by the fact that the residuals no longer exhibited significant spatial dependence ($p = 0.16$). Using this model, population density (urban population percentage), number of healthcare and/or social service facilities and number of educational establishments together explained the majority of the variability ($R^2 = 0.69$) in the number of clinical trial sites per 3-digit ZCTA (Figure 2 and Table 2). Despite the statistical significance of total population size ($p = 0.0021$), after adjusting for population density, population size had a negligible effect on the number of trial sites per 3-digit ZCTA (β estimate close to 0). Urban population percentage and number of healthcare and/or social service facilities were both significant predictors ($p < 0.0001$) of the number of clinical trial sites, but number of full-time educational establishments was not. On average, the number of clinical trial sites in a 3-digit ZCTA increased by 1 for every additional 114 healthcare and/or social service facilities, 83 educational establishments, and 165,289 people. The effect of population density was nonlinear, and Figure 2 shows that the number of clinical trial sites greatly increased in regions that were 70% or more urban.

Discussion

Our analyses revealed that the distribution of clinical trial sites per person was highly clustered around urban centers across the continental USA. As hypothesized, within a region, population density, availability of healthcare and/or social service facilities, and educational establishments together were strong predictors of the number of clinical trial sites per region, accounting for 69% of the variability seen.

From a practical standpoint of conducting clinical research, it is understandable to see the observed increase in clinical trial sites in areas with greater numbers of healthcare and/or social service facilities, which are often necessary for the procedures involved in interventional trials. Additionally, densely populated urban areas may be attractive to sponsors due to high visibility of advertisements, fewer travel barriers for participants, and access to larger and more racially and ethnically diverse populations, all of which may facilitate rapid enrollment. However, practicality should not be the main determinant of clinical trial site location. Expanding clinical research to rural areas may improve the heterogeneity of the sample population and generalizability of the research

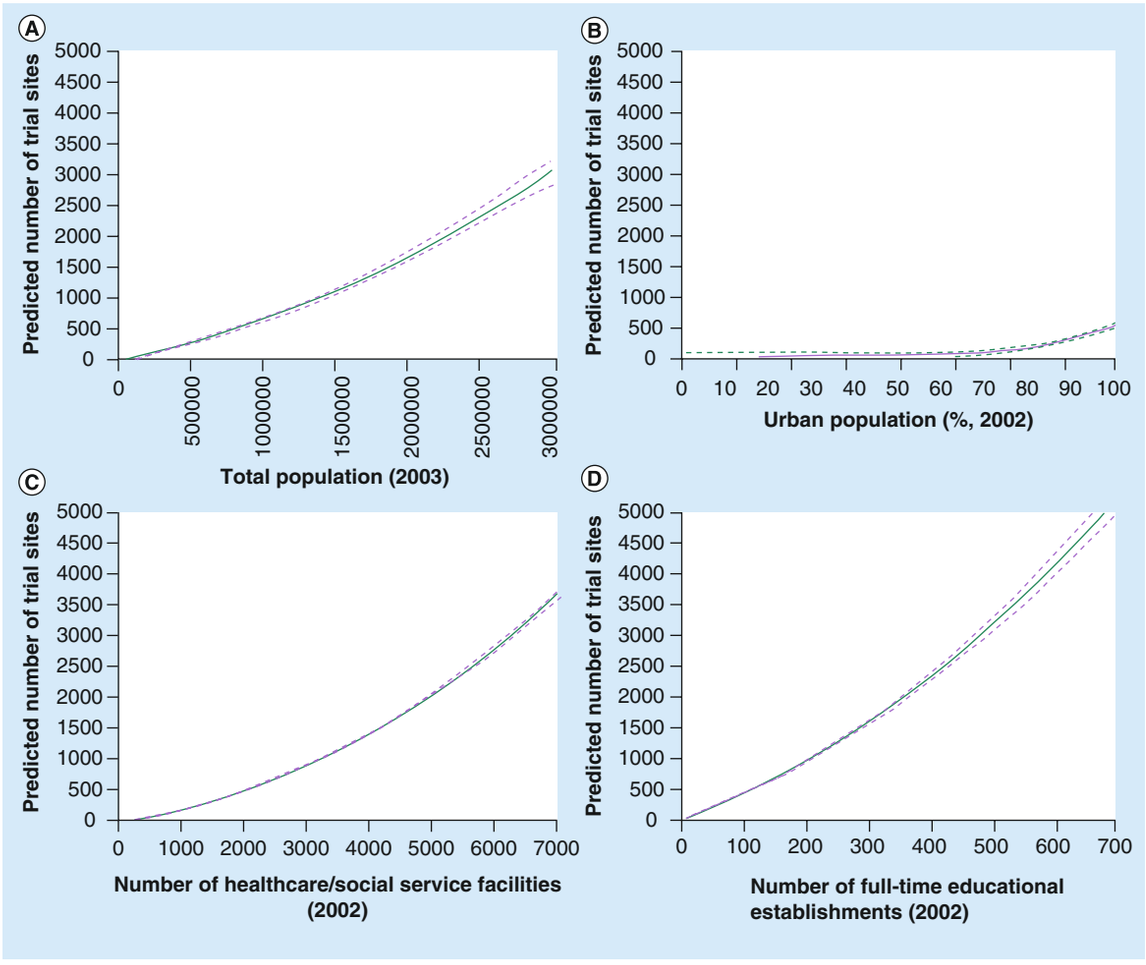

Figure 2. Predicted number of trial sites by covariates in multivariate spatial lag regression model. The figure shows the predicted number of clinical trial sites per 3-digit zip code tabulation area (solid line), as well as the 95% CI (dashed lines), based on a multivariate spatial lag regression model that included **(A)** total population, **(B)** urban population, **(C)** number of healthcare/social service facilities and **(D)** number of full-time educational establishments.

findings. Moreover, as research increasingly focuses on the effects of biologic and genetic variants on health outcomes and treatment efficacy, reaching a diversity of populations may be essential.

While the aforementioned covariates predicted the number of trial sites within a region, together they explained only 27% of the spatial distribution of clinical trial sites, indicating the presence of additional factors not analyzed that might also drive the spatial distribution of clinical trial sites. Such factors may include infrastructural features such as the location of clinical trial sponsors (e.g., pharmaceutical companies or government establishments), as well as population demographics, including socioeconomic status, insurance status, education level, and racial/ethnic composition of the population [1–3, 5]. Given that rural populations are less racially and ethnically diverse than urban populations (82 vs 66% non-Hispanic white residents, respectively) [16], it is possible, if ironic, that the well-intentioned NIH guidelines to increase enrollment of racial/ethnic minorities may actually be further disadvantaging more racially homogeneous rural populations, essentially trading one underserved population for another. Paradoxically, another explanation for the lack of representation of rural populations in clinical trials research may be that rural populations more often report fair to poor health and have chronic health issues, and in clinical trials with strict inclusion/exclusion criteria designed to isolate the disorder being studied, subjects with high co-morbidity are likely to be excluded [24]. Expanding clinical trials research to include subjects with more co-morbid conditions could potentially benefit rural populations.

Our analyses highlight the relative lack of clinical trial sites in rural areas as a potential explanation for

Table 2. Results from multivariate spatial lag regression model.

Covariate (overall $R^2 = 0.69$)	Estimate (β)	Standard error	z-value	p-value
Total population (2003)	-6.05×10^{-6}	1.97×10^{-6}	-3.08	0.0021†
Urban population percentage (quadratic term)	0.0015	0.00034	4.24	<0.0001†
Urban population percentage (linear term)	-0.069	0.044	-1.56	0.12
Number of healthcare and/or social service facilities	0.0088	0.0011	7.67	<0.0001†
Number of full-time educational establishments	0.012	0.010	1. 23	0.22

The estimate (β) represents the predicted increase in number of clinical trial sites in a given 3-digit ZCTA associated with a 1-unit increase in the covariate.
†Significant at the alpha = 0.05 level.

why rural populations are underrepresented in clinical trial research. However, it is important to note that geographic proximity is just one barrier to gaining access to clinical trials. Even when clinical trial sites exist nearby, gaining access to them can be difficult. For example, despite increased outreach efforts, racial and ethnic minority populations in urban areas remain underrepresented in clinical trials research [25]. Additional challenges to implementing and accruing to clinical trials include raising awareness of the goals of clinical research, creating workable infrastructure, training and maintaining research personnel, securing funding, and establishing a culture of research in these areas [1,17,18]. Policy-level changes (akin to the NIH guidelines for inclusion of women and racial/ethnic minorities) may also be needed to incentivize pharmaceutical companies and research organizations to recruit from geographically disadvantaged regions. Having such incentives in place would make recruiting rural subjects into clinical trials more practical.

The use of spatial analytic methods to describe and visualize the geographic distribution of clinical trials is essential to capture spatial dependence (i.e., the propensity for attributes in neighboring areas to be similar to one another). The spatial lag regression model used was superior to an ordinary least squares model in that it was able to capture spatial autocorrelation. Nevertheless, spatial analyses have some limitations. One such limitation, often unavoidable in spatial analyses, is that our results may be sensitive to 3-digit ZCTA scale at which the data were aggregated, a bias known as the modifiable areal unit problem scale effect [19]. Had the data been aggregated by 5-digit ZCTA, for example (which was not done for the reasons described in the Methods section), our results may have differed. Additionally, the study was limited by the extent of publicly available data, in that datasets with enough information on trial site location to allow for geo-coding did not provide much descriptive information about the trials. We were therefore unable to control for either trial demographics (phase, type of trial, sample size, or inclusion and exclusion criteria) or population demographics (age, gender, race, ethnicity, socioeconomic status, insurance status or co-morbidity). As a result, it was not possible to adjust for the truly eligible population for a clinical trial, as opposed to the total population (e.g., controlling only for the number of people 18 years or older for adult clinical trials). However, given that population size was a negligible predictor in the spatial lag model after controlling for the urban composition, the authors do not believe that the findings would be substantially altered. Another potential limitation is the possibility that trials reported in the BMIS between 1 November, 2002 and 1 November, 2007 were not representative of all clinical trials being conducted during that time frame.

Conclusion

Utilization of spatial analytic methods revealed that the population-adjusted distribution of clinical trials across the continental United States is highly clustered. Populations living in urban areas with more healthcare and/or social service facilities have a greater number of clinical trial sites within reach. However, geographic proximity is just one barrier to gaining access to clinical trials. Researchers must consider social, economic and cultural factors when examining barriers to accrual. As the need to diversify clinical trial participation grows, new efforts will be needed to reach underserved populations.

Financial & competing interests disclosure

The work grew out of a project originally conducted by Fletcher Spaght Inc. *pro bono* on behalf of the Massachusetts Technology Collaborative, Life Sciences Consortium. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Background
- Geographically remote, rural populations are often underrepresented in clinical trials.

Methods
- Publicly available data were used to gather zip codes from 174,503 clinical trial sites from 1 November, 2002 to 1 November, 2007. Data were geo-coded for population-adjusted spatial analyses.
- Urban population percentage, number of healthcare and/or social service facilities, and number of full-time educational establishments in a given 3-digit zip code were examined as predictors of the location and number of clinical trial sites.

Results
- Extensive clustering of clinical trial sites around urban areas with healthcare and/or social service facilities was detected.
- The analyzed covariates together explained only 27% of the geographic clustering of trial sites, but fully 69% of the variability in the number of clinical trial sites per region.

Discussion
- Geographic proximity is just one barrier to gaining access to clinical trial research.
- Additional factors that may drive the geographic clustering of clinical trial sites include infrastructural features (e.g., location of pharmaceutical or government sponsors) and population demographics.
- Factors that may affect access to trials regardless of proximity include socioeconomic status, education level, racial/ethnic composition, medical co-morbidity, insurance status, awareness of the trials, research personnel and funding.
- Expanding clinical research to rural areas may increase the heterogeneity of the population sample and generalizability of the research findings.

Conclusion and future perspective
- The distribution of clinical trial sites across the continental USA is highly clustered.
- Efforts to establish clinical trials sites in rural areas may improve access to clinical trials and increase the generalizability of clinical research.

References

Papers of special note have been highlighted as:

- **of interest**

1 Ford JG, Howerton MW, Lai GY *et al.* Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review. *Cancer* 112(2), 228–242 (2008).

- **A systematic review of 65 studies focusing on recruitment of underrepresented populations found numerous barriers to participation including awareness, opportunity and acceptance/refusal barriers and sociodemographic factors.**

2 Gifford AL, Cunningham WE, Heslin KC *et al.* HIV Cost and services utilization study consortium. participation in research and access to experimental treatments by HIV-infected patients. *N. Engl. J. Med.* 346(18), 1373–1382 (2002).

3 Menezes P, Eron JJ Jr, Leone PA,Recruitment of HIV/AIDS treatment-naïve patients to clinical trials in the highly active antiretroviral therapy era: influence of gender, sexual orientation and race. *HIV Med.* 12(3), 183–191 (2011).

4 King TE Jr. Racial disparities in clinical trials. *N. Engl. J. Med.* 346(18), 1400–1402 (2002).

5 Colon-Otero G, Smallridge RC, Solberg LA Jr *et al.* Disparities in participation in cancer clinical trials in the United States: a symptom of a healthcare system in crisis. *Cancer* 112(3), 447–454 (2008).

6 Cargill VA, Stone VE. HIV/AIDS: a minority health issue. *Med. Clin. North Am.* 89(4), 895–912 (2005).

7 National Institutes of Health. NIH guidelines on the inclusion of women and minorities as subjects in clinical research. *Fed. Regist.* 59, 11146–11151 (1994).

8 Paskett ED, Cooper MR, Stark N *et al.* Clinical trial enrollment of rural patients with cancer. *Cancer Pract.* 10(1), 28–35 (2002).

- **A Reaching Communities for Cancer Care (REACH) interventional study aimed at increasing enrollment of rural patients into cancer clinical trials through multicomponent, educational, community-based efforts was not successful.**

9 Petereit DG, Rogers D, Govern F *et al.* Increasing access to clinical cancer trials and emerging technologies for minority populations: the Native American Project. *J. Clin. Oncol.* 22(22), 4452–4455 (2004).

10 Maurer LH, Davis T, Hammond S *et al.* Clinical trials in a rural population: professional education aspects. *J. Cancer Educ.* 16(2), 89–92 (2001).

- **Rural hospitals in New Hampshire and Vermont were targeted to develop clinical trial programs and measure accrual rates and changes in patterns of care, which were similar to national rates and trends, demonstrating that expansion of clinical trial programs to include patients from rural communities is feasible.**

11 Aumiller BB, Parrott R, Lengerich EJ *et al.* Responses to a theoretically adapted clinical trial education session: faith-based sites versus rural work site dissemination. *J. Cancer Educ.* 28(4), 698–708 (2013).

- **A total of 90 church and rural community members in eight focus groups in central Pennsylvania were found to be receptive to education about cancer clinical trials, which decreased anxiety, increased knowledge and intent to participate in clinical trials, and provided suggestions for adapting future educational sessions to be even more culturally relevant.**

12 Corbie-Smith G, Odeneye E, Banks B *et al.* Development of a multilevel intervention to increase HIV clinical trial participation among rural minorities. *Health Educ. Behav.* 40(3), 274–285 (2013).

- **Intervention mapping was used to systematically develop a multilevel intervention aimed at increasing rural, racial and ethnic minority participation in HIV clinical trials.**

13 Kulldorff M. A spatial scan statistic. *Commun. Stat Theory Methods* 26, 1481–1496 (1997).

14 Anselin L. *Spatial Econometrics: Methods and Models.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 22–23 (1988).

15 Anselin L, Syabri I, Kho Y. GeoDa, an introduction to spatial data analysis. *Geographical Analysis* 38, 5–22 (2006).

16 Johnson K. *Demographic Trends in Rural and Small Town America. Reports on Rural America*. Carsey Institute, NH, USA, 1–35 (2006).

17 Minasian LM, Carpenter WR, Weiner BJ *et al.* Translating research into evidence-based practice: the National Cancer Institute Community Clinical Oncology Program. *Cancer* 116(19), 4440–4449 (2010).

- **Reviews success of the National Cancer Institute Community Clinical Oncology Program over more than 25 years and provides practical guidance to provider-based research networks.**

18 McCaskill-Stevens W, McKinney MM, Whitman CG *et al.* Increasing minority participation in cancer clinical trials: the minority-based community clinical oncology program experience. *J. Clin. Oncol.* 23(22), 5247–5254 (2005).

19 Openshaw S. *The Modifiable Areal Unit Problem CATMOG 38*. Geo Books, Norwich, UK (1984).

20 Bioresearch monitoring information system file submissions from 1 November, 2002 and 1 November, 2007, US FDA. www.fda.gov/cder/foi/special/bmis/index.htm

21 ZIP Code statistics from the economic census (2002). www.census.gov/econ/census02/guide/g02zip.htm

22 ClinicalTrials.gov file submissions from 1 November, 2002 and 1 November, 2007, US NIH. www.clinicaltrials.gov

23 FDAAA 801 requirements. ClinicalTrials.gov, US NIH. http://clinicaltrials.gov/ct2/manage-recs/fdaaa

24 Health care disparities in rural areas: selected findings from the 2004 National Healthcare Disparities Report. Agency for Healthcare Research and Quality. http://archive.ahrq.gov/research/ruraldisp/ruraldispar.pdf

25 Clinical trials shed light on minority health. food and drug administration: consumer health information. www.fda.gov/downloads/ForConsumers/ConsumerUpdates/UCM349488.pdf