



OPEN ACCESS

EDITED BY

Emanuele Monda,
University of Campania Luigi
Vanvitelli, Italy

REVIEWED BY

S. K. Ajim Ali,
Aligarh Muslim University, India
Rajan Dev Gupta,
Motilal Nehru National Institute of
Technology Allahabad, India
Vinay Kumar,
The Ohio State University,
United States

*CORRESPONDENCE

Jennifer H. Klein
jklein@childrensnational.org

SPECIALTY SECTION

This article was submitted to
Pediatric Cardiology,
a section of the journal
Frontiers in Pediatrics

RECEIVED 24 May 2022

ACCEPTED 20 July 2022

PUBLISHED 10 August 2022

CITATION

Klein JH, Gourishankar A and
Krishnan A (2022) Development of a
national dataset for geospatial analysis
of congenital heart disease.
Front. Pediatr. 10:952048.
doi: 10.3389/fped.2022.952048

COPYRIGHT

© 2022 Klein, Gourishankar and
Krishnan. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Development of a national dataset for geospatial analysis of congenital heart disease

Jennifer H. Klein^{1*}, Anand Gourishankar² and Anita Krishnan¹

¹Division of Cardiology, Children's National Hospital, Washington, DC, United States, ²Division of Hospital Medicine, Children's National Hospital, Washington, DC, United States

KEYWORDS

geospatial analysis, multicenter, geomapping, congenital heart disease, clustering

Introduction

The purpose of this opinion commentary is to describe a novel approach for studying the geospatial distribution of congenital heart disease. Congenital heart disease is estimated to occur in 1 in 100 children born in the United States each year (1, 2). The geographic location of congenital heart disease cases has emerged as an area of interest, especially as there is increased focus on health equity, delivery of care, and social determinants of health. However, developing a better understanding of the geographic characteristics and its impact on health (or disease, e.g., congenital heart disease) presents formidable challenges. We describe opportunities and challenges in developing a national dataset for multicenter geospatial analysis in congenital heart disease.

Context of geospatial analysis

Geospatial analysis, including the creation of geospatial maps—known colloquially as geomapping—is used across diverse industries, ranging from environmental conservation and national security to product delivery. It has a variety of public health applications including disease surveillance, emergency preparedness and response, community health, and environmental health (3). In healthcare, geospatial analysis aids in the study of disease etiology, influences, and outcomes, and can provide visualization tools to reinforce a conclusion, such as geographic disparities. For example, geospatial analysis tools can study relatedness of cases for spatial clustering, report drive time to tertiary care centers, analyze distance to a nearest point (such as a surgical center), or evaluate for association with an area's given attribute (percentage of high school graduates, or concentration of environmental air pollutants, for instance) (3, 4). Geographic Information System (GIS) with appropriate resources (e.g., software, hardware, and networking) and spatial analysis methods, including modeling (spatial, temporal, or both), provide a better understanding of the spatial organization of health care (5). Geography can thus provide a lens for understanding the cultural, socioeconomic, and built environment.

Application in congenital heart disease

Pediatric cardiology has embraced the study of geographic location as a key factor in caring for patients. The Baltimore-Washington Infant Study Group, a seminal population health study in cardiology, described variations in congenital heart malformations by race and socioeconomic status (6). Since that time, a compelling body of evidence highlighting significant sociodemographic disparities in many aspects of congenital heart disease care has evolved. Within single centers or regions, there is demonstrated geographic variation of the incidence of congenital heart disease (7, 8). Several other studies have evaluated sociodemographic factors' impact on surgical mortality (9, 10) and long-term outcomes from congenital heart disease (11–14). Geographic location and other sociodemographic factors are also implicated in missed prenatal detection of congenital heart disease (15–17).

Applying geospatial analysis as a method to study geographic and sociodemographic factors is a novel but powerful approach that has yet to be fully explored in pediatric cardiology. Researchers have published a single-center experience on geospatial analysis of specific cardiac diagnoses, such as Tetralogy of Fallot, hypoplastic left syndrome, and d-transposition of the great arteries (8, 18). However, to the best of our knowledge, no studies have carried out geospatial analysis of congenital heart disease on a multicenter level. The prior studies have been within a single defined region, examining cases in a single referral center, state, or catchment area. In our opinion, multicenter geospatial analysis allows for a broader investigation of risk factors for congenital heart disease, associations with the social disadvantage that may impact care, and predict where care is needed. A multicenter approach thus allows for new investigations that would otherwise be impossible.

Methods and framework for multicenter geospatial analysis

The authors seek to establish a national dataset to study prenatal detection of congenital heart disease. Through collaboration with the Fetal Heart Society, a multicenter research collaborative, partner sites are recruited for data sharing. Collaboration across several departments within Children's National Hospital provides the foundation for this endeavor, requiring input from pediatric cardiologists, epidemiologists, legal experts, and statisticians. In addition, the Child Health Advocacy Institute at Children's National Hospital provides geospatial expertise and technical support. The guiding principles of this approach are as follows:

1. *Partners:* Not only does a national dataset require research partners across the country, but also a team of

geospatial analysts, epidemiologists, or statisticians familiar with geospatial analysis. This team approach is key to building a research network and performing rigorous geospatial analyses.

2. *Privacy:* Patient protection and privacy must be at the forefront of all data-sharing agreements. In addition to rigorously subscribing to Health Insurance Portability and Accountability Act and Institutional Review Board regulations, certain analyses may be precluded or might require redaction to protect patient privacy.
3. *Products:* Geographic information systems and geospatial analysis require advanced software or tools to perform the analyses and create maps. Several such tools exist. The authors have used ArcGIS® from Esri (Redlands, CA) and continue to investigate new or alternative tools with varying functionalities and strengths.

Discussion of challenges to multicenter geospatial analysis

Developing a national dataset of congenital heart disease poses noteworthy challenges. First, the selection and recruitment of partner centers require an investment of time, interest, and resources from each site. As partner sites are selected, it is also important to accurately represent the population desired to be sampled. Too few sites, or sites only within particular geographic regions or with specific demographic characteristics, may skew the analyses. Second, ensuring the security of protected health information may hinder the collection of data, as partner sites are appropriately cautious when sharing protected health information and may limit the types of analyses that can be performed. Geospatial analysis is most accurate when using the most specific level of detail available, meaning that street address is preferred to census tract, which is preferred to counties. However, as street address is considered protected health information, there may be limits to obtaining and analyzing addresses from multiple centers. Nevertheless, geographic manipulation (e.g., aggregation, masking, cross-hatching) and tabular methods (e.g., cell suppression, omission) can mitigate these privacy issues.

Third, planning suitable analyses presents challenges. First, spatial analysis is affected by lack of continuity in a given study area. For example, hot spot analysis is only feasible in a contiguous region. Hot spot analysis looks for non-random distribution of points in the form of statistically high or low clusters, such as was described in the case of environmental biodiversity (19). Extent of clustering is influenced by the area over which the analysis is performed (20). With partner centers contributing data from geographically and demographically unique settings, finding a common denominator by which settings can be compared poses a challenge. Furthermore,

TABLE 1 Challenges and solutions to multicenter geospatial analysis.

Challenge	Potential solutions
Recruitment of partner centers for data sharing	<ul style="list-style-type: none"> ➤ Use of existing networks/consortiums and prior research collaborators to recruit partner centers
Legal agreements for data sharing	<ul style="list-style-type: none"> ➤ Data-coordinating center for streamlined data sharing ➤ Consider options for sharing of de-identified patient data or non-protected health information geographic data ➤ Use secure cloud service or data management software when cloud-based GIS solutions are unavailable across multiple centers
Planning appropriate multicenter analyses	<ul style="list-style-type: none"> ➤ Tailor analyses to questions posed and data available ➤ Engage geospatial analysis experts ➤ Use geospatial techniques (such as spatial smoothing, collapsing, or combining data), as relevant to research question

using multicenter datasets to perform clustering analyses—a particularly useful form of geospatial analysis—can be ripe for error, highlighting problems with inconsistent boundaries (20, 21), such as with varying hospital catchment areas. In addition to the geographic challenges to data analysis, there are also statistical challenges, particularly as pediatric cardiology studies relatively rare diseases. The small number problem with numerator (cases) or denominator (population or live births) creates rate instability and imprecise estimates. In summary, identifying the most appropriate analyses is dependent on the questions posed, the data available, and the needs of the population served.

Potential solutions to each of the three challenges are outlined in Table 1.

Conclusion

In this opinion we present the initial experience in creating a dataset of infants with congenital heart disease for the purpose of geospatial analysis. The fundamental principles of partnership, privacy, and products will enable success. The knowledge obtained from a single-center geospatial study is invaluable, but the lessons from a multicenter approach could be transformative for understanding congenital heart disease nationwide. In summary, we recommend the use of geospatial analysis to further understand disparities in incidence and outcomes within pediatric cardiology.

References

1. Hoffman JLE, Kaplan S. The incidence of congenital heart disease. *J Am Coll Cardiol.* (2002) 39:1890–900. doi: 10.1016/S0735-1097(02)01886-7
2. Donofrio MT, Moon-Grady AJ, Hornberger LK, Copel JA, Sklansky MS, Abuhamad A, et al. Diagnosis and treatment of fetal cardiac disease: A scientific statement from the American Heart Association. *Circulation.* (2014) 129:2183–242. doi: 10.1161/01.cir.0000437597.44550.5d

Author contributions

JK conceived of and drafted this opinion piece. AG and AK provided content expertise and manuscript revisions. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the American Heart Association Grant #821139/JK/2021.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

5. McLafferty SL. GIS and health care. *Annu Rev Public Health.* (2003) 24:25–42. doi: 10.1146/annurev.publhealth.24.012902.141012
6. Correa-Villaseñor A, McCarter R, Downing J, Ferencz C, Baltimore-Washington Infant Study Group. White-Black differences in cardiovascular malformations in infancy and socioeconomic factors. *Am J Epidemiol.* (1991) 134:393–402. doi: 10.1093/oxfordjournals.aje.a116101
7. Cronk CE, Gangnon R, Cossette S, McElroy JA, Pelech AN. Modeling geographic risk of complex congenital heart defects in eastern Wisconsin. *Birth Defects Res A Clin Mol Teratol.* (2011) 91:631–41. doi: 10.1002/bdra.20828
8. Nelson JS, Stebbins RC, Strassle PD, Meyer RE. Geographic distribution of live births with tetralogy of fallot in North Carolina 2003 to 2012. *Birth Defects Res A Clin Mol Teratol.* (2016) 106:881–7. doi: 10.1002/bdra.23566
9. Dean PN, McHugh KE, Conaway MR, Hillman DG, Gutgesell HP. Effects of race, ethnicity, and gender on surgical mortality in Hypoplastic Left Heart Syndrome. *Pediatr Cardiol.* (2013) 34:1829–36. doi: 10.1007/s00246-013-0723-3
10. Bucholz EM, Sleeper LA, Newburger JW. Neighborhood socioeconomic status and outcomes following the Norwood procedure: an analysis of the Pediatric Heart Network Single Ventricle Reconstruction Trial public data set. *J Am Heart Assoc.* (2018) 7:7065. doi: 10.1161/JAHA.117.007065
11. Peyvandi S, Baer RJ, Moon-Grady AJ, Oltman SP, Chambers CD, Norton ME, et al. Socioeconomic mediators of racial and ethnic disparities in congenital heart disease outcomes: a population-based study in California. *J Am Heart Assoc.* (2018) 7:1–12. doi: 10.1161/JAHA.118.010342
12. Favilla E, Faerber JA, Hampton LE, Tam V, DeCost G, Ravishankar C, et al. Early evaluation and the effect of socioeconomic factors on neurodevelopment in infants with Tetralogy of Fallot. *Pediatr Cardiol.* (2021) 42:643–53. doi: 10.1007/s00246-020-02525-6
13. Cassidy A, Drotar D, Ittenbach R, Hottinger S, Wray J, Wernovsky G, et al. The impact of socio-economic status on health related quality of life for children and adolescents with heart disease. *Health Qual Life Outcomes.* (2013) 11:1–8. doi: 10.1186/1477-7525-11-99
14. Bucholz EM, Sleeper LA, Goldberg CS, Pasquali SK, Anderson BR, Gaynor JW, et al. Socioeconomic status and long-term outcomes in single ventricle heart disease. *Pediatrics.* (2020) 146:e20201240. doi: 10.1542/peds.2020-1240
15. Krishnan A, Jacobs MB, Morris SA, Peyvandi S, Bhat AH, Chelliah A, et al. Impact of socioeconomic status, race and ethnicity, and geography on prenatal detection of Hypoplastic Left Heart Syndrome and Transposition of the Great Arteries. *Circulation.* (2021) 143:2049–60. doi: 10.1161/CIRCULATIONAHA.120.053062
16. Ailes EC, Gilboa SM, Riehle-Colarusso T, Johnson CY, Hobbs CA, Correa A, et al. Prenatal diagnosis of nonsyndromic congenital heart defects. *Prenat Diagn.* (2014) 34:214–22. doi: 10.1002/pd.4282
17. Quartermain MD, Pasquali SK, Hill KD, Goldberg DJ, Huhta JC, Jacobs JP, et al. Variation in prenatal diagnosis of congenital heart disease in infants. *Pediatrics.* (2015) 136:e378–85. doi: 10.1542/peds.2014-3783
18. Klein J, Dwyer G, Donofrio MT, Krishnan A. Geographic distribution of congenital heart disease: a single surgical center experience. *J Pediatr.* (2022) 240:117–21. doi: 10.1016/j.jpeds.2021.08.080
19. Possingham HP, Wilson KA. Turning up the heat on hotspots. *Nature.* (2005) 18:919–20. doi: 10.1038/436919a
20. Beale L, Abellan JJ, Hodgson S, Jarup L. Methodologic issues and approaches to spatial epidemiology. *Environ Health Perspect.* (2008) 116:1105–10. doi: 10.1289/ehp.10816
21. Fotheringham AS, Rogerson PA. GIS and spatial analytical problems. *Int J Geogr Inf Syst.* (1993) 7:3–19. doi: 10.1080/02693799308901936