



# Journal of Regional Medical Campuses

## **Physician Practice Location—An Examination of Physician Workforce Data Sources, Their Spatial Concordance and Reliability**

Emily Onello M.D.; Patrick Bright M.A.; James Boulger PhD

DOI: <https://doi.org/10.24926/jrnc.v2i5.2141>

Journal of Regional Medical Campuses, Vol. 2, Issue 5 (2019)

[z.umn.edu/JRMC](http://z.umn.edu/JRMC)

All work in JRMC is licensed under CC BY-NC



## Physician Practice Location—An Examination of Physician Workforce Data Sources, Their Spatial Concordance and Reliability

Emily Onello M.D.; Patrick Bright M.A.; James Boulger PhD

### Abstract

**Purpose:** This study compares and contrasts locational identification (physician practice ZIP code) between several physician demographic databases and a research team-verified ZIP code to determine spatial concordance. Accuracy of physician location data is critical for both successful national physician workforce planning and for assessing the fulfillment of distinct regional medical school missions.

**Methods:** Three physician databases; the American Medical Association's Physician Masterfile, the National Provider and Plan Enumeration System (NPDES), and the Minnesota Board of Medical Practice's Licensure File were compared against each other as well as a set of actively verified practice location ZIP codes. A sampling frame of medical school alumni from 2003 to 2014 was selected. The sample included alumni from both regional and main campuses. From this alumni sample, a random sample of 400 individuals were selected for closer examination. Descriptive frequencies are presented for the concordance of ZIP codes for the sample of 400 alumni.

**Findings:** From an initial cohort of 2 605 University of Minnesota medical school alumni, a sample of 400 who also possessed a Minnesota medical license were randomly selected to examine concordance rates. The highest rate of concordance was the verified ZIP code and the Minnesota Licensure Board practice ZIP code at 68.8%. Only 42% of practice location ZIP codes matched across all databases.

**Conclusions:** The concordance rate across practice ZIP codes in the databases does not inspire confidence in ability to characterize true physician practice location. Some of the difficulties in defining, identifying, and maintaining information on physician practice location are discussed. Without accurate and precise practice location information, the development and implementation of comprehensive national health policies for the United States may face difficulties. The lack of reliable physician practice location data also presents a challenge to regional medical school campuses that depend upon accurate physician location data to evaluate progress toward mission-directed workforce goals.

### Introduction

In the United States health care system, physician location is important "to demonstrate to law- or policy-makers the geographic distribution of the health care workforce to assist them in making appropriate, evidence-based decisions...(and) to identify areas of potential need for certain medical specialties for purposes of creating effective workforce strategies to expand patient access to care".<sup>1</sup> Policy decisions that utilize spatial data have had major impacts on federal and state programs and policies that attempt to alleviate health workforce imbalances and to address population health inequities.<sup>2,3</sup>

Studies have examined physician location in several contexts, ranging from the effectiveness of medical education in reducing disparities and contributing to rural health,<sup>4,5</sup> to retention and workforce movement patterns,<sup>6,7</sup> and the role of spatial access to care and outcomes.<sup>8</sup> Having correct locational data has major implications in relation to all of the efforts listed. For regional medical campuses especially,

knowledge of physician location is critical towards the fulfilling the mission of the campus, such as training doctors to practice in rural or other targeted underserved communities.

In light of the importance of physician location, the validity of major physician databases is a concern. A major source of physician spatial data is the American Medical Association's (AMA) Physician Masterfile, which is the primary source of data on U.S. physicians. The AMA Masterfile staff collects and attempts to keep updated information on physicians from the time of entry to medical school through their practice career.<sup>9</sup> Many studies have utilized the Masterfile to study and analyze practice patterns and characteristics of the U.S. workforce.<sup>3-6,10-13</sup> The common, oft-cited expectation is that reports have been accurate and necessary in the development of state, regional, and national health workforce policy development. This may not be the case when the validity of the data utilized is examined. Knowing where physicians are practicing and whether databases that

**Emily Onello M.D.;** University of Minnesota Medical School, Duluth Campus

**Patrick Bright M.A.;** University of Minnesota Medical School, Duluth Campus

**James Boulger PhD;** University of Minnesota Medical School, Duluth Campus

Corresponding author: Emily Onello M.D.; University of Minnesota Medical School, Duluth Campus 1035 University Drive, Duluth, MN 55812-3031; econello@d.umn.edu



are utilized for this purpose are accurate and concur is essential for such policy development.

Several studies have raised the concern that the AMA Masterfile may not be accurate enough to assist in the development of health policy or to understand physician shortages in their full extent.<sup>5,14–18</sup> DesRoches et al. (2015) sampled 3 000 physicians from the National Provider and Plan Enumeration System (NPPES), and compared the listed addresses with the Masterfile, and with the data available from SK&A (by IQVIA, a private data-consulting firm). The AMA Masterfile had a low rate of matching practice address as listed by the NPPES (32% to 54% depending upon specialty). The NPPES and SK&A had higher address concordance rates (72% to 94% and 79% to 92% respectively, across specialties).

However, Henderson (2015) noted that DesRoches et al.'s methodologic use of the AMA's *Physician Preferred Address* prevented a meaningful comparison of the databases since the Physician's Preferred Address is subjectively provided by the physician and by definition could vary from practice ZIP, home residence ZIP, or other location. (The AMA allows physicians to indicate a preferred mailing address, which may not be a physician's actual practice address).<sup>19</sup> Henderson asserts that comparing this address to that in the NPPES and SK&A databases may have resulted in lower match rates. Freed et al. (2006), while assessing the distribution and number of pediatric cardiologists, found that only 58% were listed by both the AMA Masterfile and the American Board of Pediatrics. Of the 42% not in both databases, an additional 28% were only listed in the AMA Masterfile and 4% percent by only the Board. McLafferty et al. (2012), similarly to Henderson, noted that some unknown number of physicians report a home mailing address to the Masterfile rather than a work address. Reporting physician home address will result in misclassification of the physician in the workforce to an unknown degree.

However, other studies affirm the extensive work that is performed by the AMA's Division of Health Solutions Data Management. Some have verified generally strong concordance with other physician databases.<sup>7,20–22</sup>

Given the concerns raised on the validity of these data sources, a closer examination on the spatial aspects is warranted. This study aims to identify the degree of concordance between data sources for physician practice location, in terms of spatial location in the form of ZIP codes. ZIP code analyses are frequently used by state and federal agencies in delineating areas by their demographic, socio-economic, cultural, and environmental characteristics.<sup>23</sup> ZIP codes have often been used to create taxonomies such as the Rural-Urban Commuting Area Codes (RUCAs) for delineating rural and urban areas.<sup>24</sup>

State medical licensure boards offer another source of physician location ZIP code data that may be higher in accuracy and validity, though few studies have compared licensure location data to other datasets. A recent study by

Bell et al. (2018) examined a sample of South Carolina physician and nurse practitioner state license address data for correspondence to their place of employment, which was set as the location recorded in the National Committee for Quality Assurance Patient-Centered Medical Home (PCMH) provider file. Comparisons were also made to National Provider Identifier (NPI) file zip code data. At the zip code level, physician state license data was found to have an 85.5% concordance rate with the PCMH location and 88.2% concordance rate with the physician NPI file data. (Interestingly, similar zip code comparisons for Nurse Practitioners between state license zip code data, PCMH file, and NPI data demonstrated concordance rates of 35.8% and 76.5%, respectively).<sup>25</sup>

The study presented here uses a large cohort of medical school alumni that includes a regional medical campus. Actual physician practice location was compared to the practice address listed in 3 major datasets (AMA Masterfile, NPPES's National Provider Identification (NPI) number, and Minnesota Board of Medical Practice Licensure File) in order to investigate the validity of the data sources.

## Materials and Methods

Our initial study population was 2 605 medical school alumni from the University of Minnesota (UMN) Medical School graduating in 12 consecutive classes (2003–2014). The study intentionally included the span of 12 graduating classes using a cut-off date at least 3 years prior to the research study for 2 purposes: 1) to establish a large enough data set from which to randomly select the actual study sample of 400 alumni, and 2) to allow time for the majority of the alumni physicians to have completed residency and established their chosen medical practice location, thereby minimizing the chances of including physicians who are still in residency training. Graduates from the study timeframe's final year (2014) would have completed a 3-year residency (2014–17) and selected an initial practice location by the time of data collection in early 2018.

The list of University of Minnesota Medical School graduates was assembled from the published National Resident Matching Program (NRMP) Match List and the printed graduation programs of the medical school. The list was then compared by name to each of the data sources (AMA Masterfile, NPPES/NPI, MN Board of Medical Practice) to see if each graduate was identified in the data source. Three datasets were chosen for our analysis of physician spatial location concordance:

- 1) The AMA Masterfile. The Masterfile is publicly available and was purchased from Medical Marketing Services, an official broker for the AMA. Among many other data elements, the AMA Masterfile includes the physician's current office ZIP code. Data was pulled from the Masterfile on January 18, 2018; officeZIP was the zip code variable utilized ("OFFICE\_ZIP").

2) The National Plan and Provider Enumeration System's (NPPES). National Provider Identifier (NPI) File was downloaded from the Centers for Medicare & Medicaid Services (CMS) website in January 2018 and included the current practice ZIP code ("praczip").

3) Minnesota Board of Medical Practice's Licensure File. Data on physicians with a Minnesota (MN) medical license was obtained on March 14, 2018 via a data request to the Minnesota Board of Medical Practice. This data included office ZIP code information on file with the License Board ("zip").

The decision to utilize primary practice ZIP codes as our unit of analysis was made for several reasons. Practice ZIP code was found across all 3 datasets, ensuring a common variable for analysis. This is also an attempt to address the concerns of Henderson (2015), when comparing the databases by using office location rather than the physicians preferred address.

The data obtained from the AMA Masterfile did not include the street address for identification purposes; only the practice ZIP code. Practice ZIP codes were determined to be a reasonable approximation of location, noting that use of the actual street address could be too granular with the potential to overestimate the mismatch rates of physician location between datasets. For example, a physician may have a practice address at a hospital with one street address, but may also have an outpatient clinic practice address in a different clinic building on the same hospital campus with a different street address. However, both practice locations should share the same ZIP code. Lastly, ZIP codes were utilized as they are adequate for workforce planning.

Of the initial 2 605 alumni physician population, 1 384 physicians held a Minnesota license. This sub-group of alumni with MN medical licenses could be tracked using both the Minnesota Board of Medical Practice location data along with the AMA Masterfile and the NPI sources. From this group of 1 384 Minnesota license holders, a sample of 400 alumni was randomly selected ensuring 95% confidence of being within 5% of the true population count.<sup>26</sup> A representative sub-sample was utilized due to the logistical barriers of actively verifying the practice addresses of all 1 384 alumni holding a Minnesota license.

In early 2018, the research team worked to actively over a period of several weeks to verify the practice ZIP code for each graduate in the sample of 400. Practice addresses and ZIP codes were verified using hospital and clinic websites, professional network websites such as Doximity and LinkedIn, and contact via social media. In addition, direct personal knowledge of graduates was utilized to characterize location, including face-to-face encounters between alumni and research team members during site visits to practice locations, alumni networking events, statewide medical conferences. In addition to these face-to-face meetings, phone conversations were used to establish the verified practice location. This process was used to establish the physician's updated and verified practice location and

represented our best effort to obtain current physician location data.

Once this actively verified practice ZIP code was ascertained, it was compared to the ZIP codes of the AMA, NPI, and state medical board ZIP codes of practice. If the verified ZIP code matched, a designation of "Match" was given, otherwise the comparison was coded as "No Match". The following pairings were contrasted:

1. AMA Office ZIP code to NPI Practice ZIP code.
2. AMA Office ZIP code to Minnesota License Board ZIP code.
3. AMA Office ZIP code to verified ZIP code.
4. NPI Practice ZIP code to Minnesota License Board ZIP code.
5. NPI Practice ZIP code to verified ZIP code.
6. Minnesota License Board ZIP code to verified ZIP code.

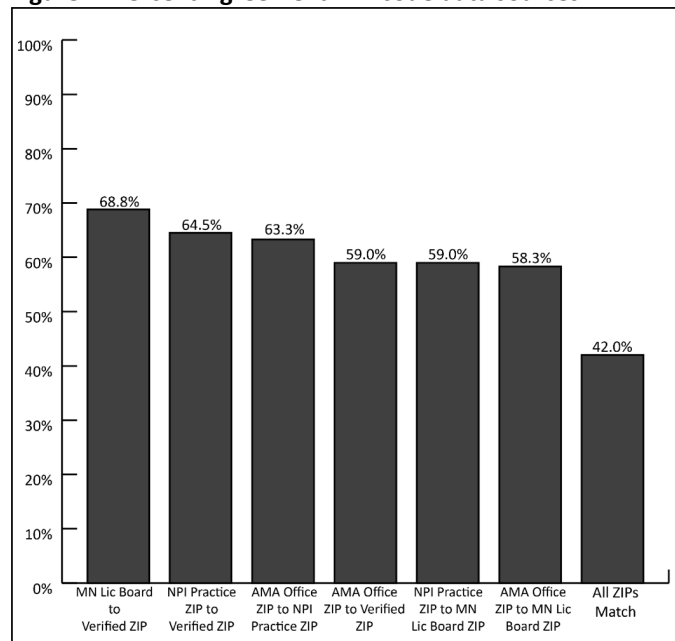
If the practice ZIP code was identical across all the datasets, locations were coded as "Complete Match".

## Results

Of the 400 alumni randomly sampled, the following was found: 36 (9%) had no AMA office ZIP code, 1 (0.2%) had no NPI practice ZIP code listed, and 16 (4%) we were unable to verify their practice ZIP code.

Comparisons between the data sets and the verified zip code demonstrated a range of ZIP code concordance rates. As Figure 1 shows, the highest concordance rate was between the verified ZIP code and the MN License Board ZIP at 68.8%. Only 42% of the sample had a complete ZIP code match across all 4 data sets.

**Figure 1. Percent Agreement: ZIP code data sources**



The 2 national level datasets, the AMA Masterfile and the NPPES, had a 63.3% ZIP code match rate. Notably, the Masterfile had only a 58.3% match to the Minnesota License Board and a 59% match to the verified ZIP code, while the

NPPES was marginally better with a 64.5% match to the verified ZIP code. Differences between all the pairs were significant ( $\chi^2= 71.65$ ,  $p<.001$ ).

These results clearly illustrate a difference in agreement between datasets when examining the practice ZIP code for concordance. The best agreement is that of the Minnesota State License Board ZIP with our verified ZIP code of 68.8%. Of concern, this still leaves a strikingly high amount of non-matched practice addresses (31.2%).

The finding that the Minnesota Board of Medical Practice's Licensure File provided the best match rate with the verified practice ZIP code was not surprising. The Minnesota board requires annual online renewal of medical licensure. As part of each physician's renewal application, updated practice information is requested and subsequently provided by the physician. Unlike the AMA Masterfile or the NPPES databases, the Minnesota State License Board information is certain to be reviewed annually at a minimum.

## Discussion

This study illustrates a discord between major physician databases that have been used to develop and determine national health positions and policies. These findings suggest an urgent need for a standardized nationwide data repository, valid and freely available, in order to identify exact physician practice location. To achieve this universal standard, more precise definitions of physician practice location will be required.

A major conceptual issue that remains unclear is the definition of a physician's "place". Is this where they spend the majority of their time in practice? What if a physician is primarily an urban physician but also practices one day a week in a rural location? Should this physician be counted as a rural physician despite not spending the majority of his/her time there? Some physicians, such as anesthesiologists for example, may practice in multiple locations within a given city or municipality (providing services at multiple hospitals or outpatient facilities), introducing ambiguity when attempting to assign a static definition of "place" to these practices. The rise of health systems has further muddled the concept of "place". These systems may utilize a central billing address for their physicians, some of whom may practice in multiple locations or in a satellite location distant from the central billing address. Do we count the physical location of the physician or the billing address as their "place"? This kind of ambiguity in defining a physician's place can have important implications for policy makers and public health strategists who interpret and plan around physician workforce data sets. Again, establishing and applying standardized definitions of physician practice location across datasets could reduce such ambiguities.

One potential alternative to improve data accuracy would be to acquire data from other sources such as professional medical specialty societies, such as the American Academy of

Family Physicians. Professional medical societies generally require annual membership renewal by the member physician. In addition, many of these societies require physicians to keep up maintenance of certification to retain their board certification in a given specialty. It is likely that the professional relationship propinquity may well improve the validity of the address communications. Further studies are required to ascertain potential means of improvement. This study does have several limitations. The first limitation is that the data sets were not acquired simultaneously from the 3 separate sources (AMA and NPPES in January 2018 and MN Board of Medical Practice in March 2018). But given the relative stability of physician practices, it is unlikely that the 2-month difference introduced a significant amount of mismatch between datasets.

The selection of zip code as the locator variable is not without complexity, as discussed earlier, as there are multiple potential classifications of zip code. For example, the AMA Masterfile includes 2 variables for zip code: MailZipCode, OfficeZIP. (The MailZipCode variable included 4 subcategories: physician's home ZIP code, physician's home-office ZIP code, physician's office ZIP code, or unknown). This study used OfficeZIP only, as it provided a more consistent and less confusing method of geographic location. Selection of the more specific zip code variable addresses the concern of McLafferty et al. (2012) and Henderson (2015) of potential ambiguity of actual practice location if this variable is utilized to classify practice location.

The study relies on an actively verified ZIP code that was established by exhaustive inquiry and cross-checking. While there is no clear metric to provide absolute confirmation of the veracity of our attempt at verifying physician location, these methods used to locate the physicians should result in a more valid determination of physician practice location than intermittent self-report by physicians or their employers to national organizations. Additionally, regional medical campuses may have more detailed personal knowledge of their graduates than larger central campuses due to smaller class sizes characteristic of regional medical schools. Another study limitation is that there were a few graduates who could not be located. Some of the followed graduate cohort did not have an office ZIP code on file in the Masterfile, or did not have a ZIP code listed in the NPPES database. Reliance upon manual searching for missing graduate practice data and a lack of an NPPES NPI number for identification did introduce a margin of error about current practice information. Several instances of mismatches occurred as a result of the physician's maiden names which was not at times reflected in the NPPES database. Furthermore, it is possible that a few of the graduates who could not be located may not have "officially" graduated, even if they had been listed in the medical school's printed graduation ceremony program. Similarly, some of the graduates may not have ever actually entered medical practice in the United States, despite earning the degree of



medical doctor. The total number of alumni who could not be located was very small in relation to the overall cohort and it is unlikely that this small number exerted any significant influence on the resulting analysis.

While this study revealed that the Minnesota Board of Medical Practice data demonstrated the highest match rate with the verified physician location, the accuracy of state licensing board information may vary in other states. Some state boards do not update annually, but do so less frequently, which could result in higher mismatch rates.

## Conclusion

Surprisingly low concordance rates were found between data sources for physician practice location when comparing ZIP code data across the AMA Masterfile, the NPPES NPI database, and Minnesota State Board of Medical Practice against a verified standard. State licensing board data demonstrated the highest rate of agreement with verified data.

Our study confirms earlier findings that suggest it is more difficult to identify current physician practice location than previously thought. Findings such as this have major implications for policy makers who utilize national level datasets for workforce decisions. Furthermore, incorrectly attributing a physician's practice location to an incorrect ZIP code will distort the picture of the workforce landscape. This also has the potential to make it more difficult to accurately characterize the effect of regional medical campus graduates on workforce needs in rural and underserved areas of the U.S.

It is hoped that attention to this finding of spatial discord (in terms of reliability and validity), will result in the recognition of the importance of strong efforts to provide better information to policy and decision makers for not only the good of the general population but to correctly ascertain the positive effects regional medical schools have across the landscape.

## References

1. American Medical Association. Health Workforce Mapper. <https://www.ama-assn.org/about/research/health-workforce-mapper>. Published 2018. Accessed March 1, 2019.
2. American Academy of Family Physicians. Geographic Distribution of Primary Care Physicians Affects Health Care, Says Policy Brief. <https://www.aafp.org/news/government-medicine/20130620geodistpolicy.html>. Published 2013. Accessed August 13, 2018.
3. Rosenthal MB, Zaslavsky A, Newhouse JP. The geographic distribution of physicians revisited. *Health Serv Res.* 2005;40(6 Pt 1):1931-1952. doi:10.1111/j.1475-6773.2005.00440.x
4. Chen FM, Fordyce MA, Andes S, Hart LG. U . S . Rural Physician Workforce : Analysis of Medical School Graduates from 1988-1997. 2008;(October).
5. Ko M, Heslin KC, Edelstein RA, Grumbach K. The role of medical education in reducing health care disparities: The first ten years of the UCLA/Drew Medical Education Program. *J Gen Intern Med.* 2007;22(5):625-631. doi:10.1007/s11606-007-0154-z
6. Cullen TJ, Hart LG, Whitcomb ME, Rosenblatt RA. The National Health Service Corps: rural physician service and retention. *J Am Board Fam Pract.* 1997;10(4):272-279. <http://www.ncbi.nlm.nih.gov/pubmed/9228622>. Accessed August 15, 2018.
7. Ricketts TC, Randolph R. Urban-rural flows of physicians. *J Rural Heal.* 2007;23(4):277-285. doi:10.1111/j.1748-0361.2007.00104.x
8. Wan N, Zhan FB, Zou B, Wilson JG. Spatial Access to Health Care Services and Disparities in Colorectal Cancer Stage at Diagnosis in Texas. *Prof Geogr.* 2013;65(3):527-541. doi:10.1080/00330124.2012.700502
9. American Medical Association. AMA Physician Masterfile. <https://www.ama-assn.org/life-career/ama-physician-masterfile>. Published 2018. Accessed August 15, 2018.
10. Kletke PR, Marder WD. The Supply of Renal Physicians: An Analysis of Data From the American Medical Association Physician Masterfile. *Am J Kidney Dis.* 1991;18(3):384-391. doi:10.1016/S0272-6386(12)80100-4
11. Waters TM, Lefevre F V, Budetti PP. Medical school attended as a predictor of medical malpractice claims. *Qual Saf Health Care.* 2003;12(5):330-336. doi:10.1136/qhc.12.5.330
12. Schwartz MR. The physician pipeline to rural and underserved areas in Pennsylvania. *J Rural Heal.* 2008;24(4):384-389. doi:10.1111/j.1748-0361.2008.00185.x
13. Lango MN, Handorf E, Arjmand E. The geographic distribution of the otolaryngology workforce in the United States. *Laryngoscope.* 2017;127(1):95-101. doi:10.1002/lary.26188
14. Konrad TR, Slifkin RT, Stevens C, Miller J. Using the American Medical Association physician masterfile to measure physician supply in small towns. *J Rural Heal.* 2000;16(2):162-167. doi:10.1111/j.1748-0361.2000.tb00450.x
15. Freed GL, Nahra TA, Wheeler JRC. Counting physicians: Inconsistencies in a commonly used source for workforce analysis. *Acad Med.* 2006;81(9):847-852. doi:10.1097/00001888-200609000-00017

16. McLafferty S, Freeman VL, Barrett RE, Luo L, Shockley A. Spatial error in geocoding physician location data from the AMA Physician Masterfile: Implications for spatial accessibility analysis. *Spat Spatiotemporal Epidemiol.* 2012;3(1):31-38. doi:10.1016/j.sste.2012.02.004
17. DesRoches CM, Barrett KA, Harvey BE, et al. The Results Are Only as Good as the Sample: Assessing Three National Physician Sampling Frames. *J Gen Intern Med.* 2015;30(S3):595-601. doi:10.1007/s11606-015-3380-9
18. Xierali IM. Physician Multisite Practicing: Impact on Access to Care. *J Am Board Fam Med.* 2018;31(2):260-269. doi:10.3122/jabfm.2018.02.170287
19. Henderson M. Assessing the Accuracy of Three National Physician Sampling Frames. *J Gen Intern Med.* 2015;30(10):1402-1402. doi:10.1007/s11606-015-3483-3
20. Cherkin D, Lawrence D. An evaluation of the American Medical Association's physician masterfile as a data source--one state's experience. *Med Care.* 1977;15(9):767-779. <http://www.ncbi.nlm.nih.gov/pubmed/578284>. Accessed August 15, 2018.
21. Werner RM, Polsky D. Comparing the supply of pediatric subspecialists and child neurologists. *J Pediatr.* 2005;146(1):20-25. doi:10.1016/j.jpeds.2004.08.061
22. Vanasse A, Ricketts TC, Courteau J, Orzanco MG, Randolph R, Asghari S. Long term regional migration patterns of physicians over the course of their active practice careers. *Rural Remote Health.* 2007;7(4):812. doi:812 [pii]
23. Onega T, Weiss JE, Alford-Teaster J, Goodrich M, Eliassen MS, Kim SJ. Concordance of Rural-Urban Self-identity and ZIP Code-Derived Rural-Urban Commuting Area (RUCA) Designation. *J Rural Heal.* 2019;0:jrh.12364. doi:10.1111/jrh.12364
24. Hart LG, Larson EH, Lishner DM. Rural definitions for health policy and research. *Am J Public Health.* 2005;95(7):1149-1155. doi:10.2105/AJPH.2004.042432
25. Bell N, Lòpez-DeFede A, Wilkerson RC, Mayfield-Smith K. Precision of provider licensure data for mapping member accessibility to Medicaid managed care provider networks. *BMC Health Serv Res.* 2018;18(1):974. doi:10.1186/s12913-018-3776-4
26. McGrew JC, Monroe CB. *An Introduction to Statistical Problem Solving in Geography: Second Edition.* Second. Long Grove: Waveland Press; 2009.