HSR Health Services Research

DOI: 10.1111/1475-6773.12351
EDITORIAL

Editorial

County Health Rankings and the Cult of the Imperfect

Sir Robert Watson-Watt, who developed early warning radar in Britain during World War II to counter the rapid growth of the Luftwaffe, proposed a "cult of the imperfect," which he stated as "Give them the third best to go on with; the second best comes too late, the best never comes" (Brown 1999). Just as imperfect radar detected incoming German aircraft to protect British citizens, imperfect public health surveillance systems are widely used to measure and monitor the health of populations to mobilize action toward community health.

Defined as the ongoing systematic collection, analysis, interpretation, and dissemination of health data for use in planning, implementing, and evaluating public health practice, public health surveillance serves as the centerpiece of all community health improvement efforts (Remington and Flood 2014). In practice, public health surveillance is a continuous process involving four diverse components: (1) data collection, (2) analysis, (3) interpretation, and (4) dissemination. Each of these steps in the process requires different skills and systems. Information technology and systems engineering are needed to design data collection systems. Epidemiologic and biostatistical methods are needed for data analysis, especially for small areas where data may be sparse or missing. But public health surveillance also requires skills in communications and program planning, to assure that the data are translated into useful information that supports community health improvement efforts (Remington and Nelson 2010).

The long-standing tradition of public health surveillance has been simply to "disseminate" results to those working in the public health system. Increasingly, public health surveillance systems have focused on developing specific communication plans to define the purpose of communicating information, understand audiences, develop messages, select appropriate channels,

market the information, and evaluate the process and outcomes. The ultimate purpose of these systems is to improve the health of populations by informing or persuading individuals or policy makers (Parvanta et al. 2002).

The *County Health Rankings* have become one of the most widely recognized public health surveillance systems in the nation (Remington, Catlin, and Gennuso 2015). Annually since 2010, the University of Wisconsin Population Health Institute and the Robert Wood Johnson Foundation have produced the *Rankings*, which serve as a "population health checkup" for the nation's more than 3,000 counties. We base the *Rankings* on a conceptual model of population health that includes both health outcomes (mortality and morbidity) and health factors (health behaviors, clinical care, social and economic factors, and the physical environment). Data for over 30 measures are available at the county level for over 3,000 counties in the United States, and they are assembled and combined to create composite measures that are then ordered and counties are ranked from best to worst health within each state.

There is ample evidence that the *County Health Rankings* is achieving its goal to mobilize action toward community health by stimulating interest among the media and policy makers. A Google search of "County Health Rankings" returns about 340,000 hits with stories from every state and countless communities. Rankings serve as a hook for the media and play on our competitive instincts. It is much easier to say "the best" or "the worst" than it is to discuss age-adjusted rates or compare relative or absolute differences. This experience is not unique to health rankings and has been noted in rankings of educational institutions. Hazelkorn (2009) suggests that the practice of "naming and shaming" introduces a competitive element that positively influences institutional behavior and increases the quality of higher education.

However, the *County Health Rankings* are far from a perfect surveillance system. Recently, some have examined the methods used in health rankings and have expressed concerns about the reliability of the results (Erwin et al. 2011; Kanarek, Tsai, and Stanley 2011; Remington and Booske 2011; Hendryx, Ahern, and Zullig 2013; Arndt et al. 2013). In this issue of *Health Services Research*, Courtemanche, Tchernis, and Soneji (2015) examine the performance of the *Rankings* in two states—Wisconsin and Texas. Using sophisticated and appropriate methods, the authors conclude that the reliability of the *Rankings* varies between states; the results would be strengthened if

*Address correspondence to Patrick L. Remington, M.D., M.P.H., Department of Population Health Sciences, School of Medicine and Public Health, University of Wisconsin-Madison, 750 Highland Ave, Rm 4263; e-mail: plreming@wisc.edu.

we incorporated alternate weighting methods, and the *Rankings* should account for uncertainty.

The reliability of the measures used in the *County Health Rankings* varies greatly from county to county. Courtemanche et al. observed this in Texas, where 89 of 254 counties have fewer than 10,000 residents and 46 have fewer than 4,000 residents—including Loving County, which with only 46 residents is the least populated county in the nation. To account for unreliability nationwide, we exclude the least populated counties from the *Rankings* (124 counties in 2011, about 4 percent of all counties). In 2011, Texas accounted for 25 percent of the unranked counties in the nation (31/124).

Statistical techniques can be used to improve the reliability of data from less-populated counties. When we compared results using models with and without demographic fixed effects, we found that counties whose ranks had wide confidence intervals had smaller populations or ranked in the middle of all counties for health outcomes (Athens et al. 2013). Incorporating covariates produced narrower confidence intervals, but rank estimates remained imprecise for many counties. We recommended that local health officials, especially in less-populated and midperforming communities, consider these limitations when interpreting the results of the *Rankings*.

Efforts are underway to improve the reliability of the survey-based measures that we use in the *County Health Rankings*, such as the Behavioral Risk Factor Surveillance System. Dwyer-Lindgren et al. (2015) at the Institute for Health Metrics and Evaluation at the University of Washington has recently developed small area models incorporating spatial and temporal smoothing to improve the reliability of estimates from the Behavioral Risk Factor Surveillance System. The Centers for Disease Control and Prevention has implemented similar models for obesity measures (http://www.cdc.gov/diabetes/pdfs/data/calculating-methods-references-county-level-estimates-ranks.pdf) and is extending this approach to include other measures from the Behavioral Risk Factor Surveillance System.

Although these statistical methods will improve reliability of the measures used in the *Rankings*, there are limitations. Spatial smoothing assumes that geographic proximity leads to more similar health factors and outcomes —an assumption that is true in certain areas but does not apply when neighboring communities have stark difference in factors that influence health outcomes. Perhaps more important, local stakeholders may find these methods complicated and difficult to understand.

Other practical approaches are possible to improve the reliability of the *Rankings*, such as changing the size of populations ranked or collecting more

and better data. In their seminal work "Eight Americas: Investigating Mortality Disparities across Races, Counties, and Race-Counties in the United States," Murray et al. (2006) arranged the 3,141 counties in the United States into 2,072 units by merging smaller counties with adjacent counties to form units with total population of at least 10,000 residents. The advent of electronic health records provides the potential to collect health information from entire health systems or communities, thus avoiding the limitations inherent in survey methods (Remington and Wadland 2015). The Institute of Medicine recommends developing and testing systems to collect patient-level information, to share de-identified data across systems, and to make them available at the local, state, and national levels to monitor and improve health outcomes (Institute of Medicine [IOM] 2012).

Courtemanche et al. also found that the population health model used in the *County Health Rankings* performed well in Wisconsin but not in Texas. Using a process to determine factor weights that incorporated population, spatial covariance, and missing data, they found that these data-derived factor weights in Texas differed from standard *Rankings* model weights. This is not surprising, since health factors and outcomes vary across states (e.g., Texas may have greater variation in low birth weight rates than other states). Separate models could be generated for each state, based on the observed data in that state. Interestingly, Cofiño et al. used the *County Health Rankings* population health model in Asturias, Spain, but reduced the weight given to "health care" (from 20 to 10 percent) since all residents have access to health care (Cofiño et al. 2014).

Finally, Courtemanche recommends that the *County Health Rankings* should account for variability by providing error estimates for the ranks. Error estimates are provided for all measures on the *County Health Rankings* website (see www.countyhealthrankings.org). We are collaborating with others to adapt methods that have been used to estimate confidence intervals for cancer mortality rankings (Zhang et al. 2014), including discussions about Type I error (i.e., stating that a difference in rank exists when it does not). Is the 95 percent confidence interval used in research the most appropriate standard when considering differences in the health of populations? Perhaps policy makers would prefer an 80 or 60 percent confidence interval when comparing the health rankings of counties in their state.

In conclusion, the goals of the *County Health Rankings* are to build awareness of the multiple factors that influence health and provide a reliable, practical, and sustainable source of local data for communities. Rather than providing the final answers, the *Rankings* serve as a "call to action" to collect

better local data and engage community partners in collective action. To this end, we strive to balance efforts to improve the quality of the data with the need to provide the most recently available data for over 3,000 counties in the United States, in a clear and understandable way. Perfection does not have to be the enemy of the good. Instead, the work of Courtemanche and others challenges the "cult of the imperfect" and offers practical suggestions that will help make the *County Health Rankings* the best public health surveillance system possible.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: Support for this work was provided by a grant from the Robert Wood Johnson Foundation (ID 69835).

Disclosures: None.

Disclaimers: None.

Patrick L. Remington

REFERENCES

Arndt, S., L. Acion, K. Caspers, and P. Blood. 2013. "How Reliable Are County and Regional Health Rankings?" *Prevention Science* 14: 497–502.

Athens, J. K., B. B. Catlin, P. L. Remington, and R. E. Gangnon. 2013. "Using Empirical Bayes Methods to Rank Counties on Population Health Measures." *Preventing Chronic Disease* 10: E129.

Brown, L. 1999. *Technical and Military Imperatives: A Radar History of World War 2*. Bristol, UK: IOP Publishing Ltd.

Cofiño, R., M. Prieto, O. Suárez, and K. Malecki. 2014. "The Art of Drawing Numbers and Stories in the Air: Epidemiology, Information, Emotion and Action." *Journal of Epidemiology and Community Health* 68: 1109–11.

Courtemanche, C., R. Tchernis, and S. Soneji. 2015. "Modeling Area-Level Health Rankings." *Health Services Research* (this issue).

Dwyer-Lindgren, L., A. D. Flaxman, M. Ng, G. M. Hansen, C. J. L. Murray, and A. H. Mokdad. 2015. "Drinking Patterns in US Counties from 2002 to 2012." *American Journal of Public Health*. Published online ahead of print April 23, 2015. e1–8.

Erwin, P. C., C. R. Myers, G. M. Myers, and L. M. Daugherty. 2011. "State Responses to America's Health Rankings: The Search for Meaning, Utility, and Value." *Journal of Public Health Management and Practice* 17 (5): 406–12.

Hazelkorn, E. 2009. *Impact of Global Rankings on Higher Education Research and the Production of Knowledge*. Occasional Paper, No. 18. Unesco Forum on Higher Education, Research and Knowledge.

Hendryx, M., M. M. Ahern, and K. J. Zullig. 2013. "Improving the Environmental Quality Component of the County Health Rankings Model." *American Journal of Public Health* 103 (4): 727–32.

Institute of Medicine [IOM]. 2012. "Chapter 5: Surveillance and Assessment." In *Living Well with Chronic Illness: A Call for Public Health Action*, pp. 187–227. Washington, DC: National Academies Press.

Kanarek, N., H. L. Tsai, and J. Stanley. 2011. "Health Ranking of the Largest US Counties Using the Community Health Status Indicators Peer Strata and Database." *Journal of Public Health Management and Practice* 17 (5): 401–5.

Murray, C. J. L., S. C. Kulkarni, C. Michaud, N. Tomijima, M. T. Bulzacchelli, T. J. Iandiorio, and M. Ezzati. 2006. "Eight Americas: Investigating Mortality Disparities across Races, Counties, and Race-Counties in the United States." *PLoS Medicine* 3 (12): e545.

Parvanta, C., E. Maibach, E. Arkin, D. E. Nelson, and J. Woodward. 2002. "Public Health Communication: A Planning Framework." In *Communicating Public Health Information Effectively: A Guide for Practitioners*, edited by D. E. Nelson, R. C. Brownson, P. L. Remington, and C. Parvanta, pp. 11–31. Washington, DC: American Public Health Association.

Remington, P. L., and B. C. Booske. 2011. "Measuring the Health of Communities–How and Why?" *Journal of Public Health Management and Practice* 17 (5): 397–400.

Remington, P. L., B. B. Catlin, and K. P. Gennuso. 2015. "The County Health Rankings: Rationale and Methods." *Population Health Metrics* 13 (11): 12.

Remington, P., and T. Flood. 2014. "Public Health Surveillance." Oxford Bibliographies [accessed on April 26, 2015]. Available at http://www.oxfordbibliographies.com/view/document/obo-9780199756797/obo-9780199756797-0149.xml

Remington, P. L., and D. Nelson. 2010. "Communicating Surveillance Information for Action." In *The Principles and Practice of Public Health Surveillance*, 3rd Edition, edited by L. M. Lee, S. M. Teutsch, M. E. St Louis, and S. B. Thacker, pp. 146–65. New York: Oxford University Press.

Remington, P. L., and W. C. Wadland. 2015. "Connecting the Dots: Bridging Patient and Population Health Data Systems." *American Journal of Preventive Medicine* 48 (2): 213–4.

Zhang, S., J. Luo, L. Zhu, D. G. Stinchcomb, D. Campbell, G. Carter, S. Gilkeson, and E. J. Feuer. 2014. "Confidence Intervals for Ranks of Age-Adjusted Rates across States or Counties." *Statistics in Medicine* 33 (11): 1853–66.