Challenges in Identifying Refugees in National Health Data Sets

Wagahta Semere, MD, Katherine Yun, MD, Cyrus Ahalt, MPP, Brie Williams, MD, and Emily A. Wang, MD

Objectives. To evaluate publicly available data sets to determine their utility for studying refugee health.

Methods. We searched for keywords describing refugees in data sets within the Society of General Internal Medicine Dataset Compendium and the Inter-University Consortium for Political and Social Research database. We included in our analysis US-based data sets with publicly available documentation and a self-defined, health-related focus that allowed for an examination of patient-level factors.

Results. Of the 68 data sets that met the study criteria, 37 (54%) registered keyword matches related to refugees, but only 2 uniquely identified refugees.

Conclusions. Few health data sets identify refugee status among participants, presenting barriers to understanding refugees' health and health care needs.

Public Health Implications. Information about refugee status in national health surveys should include expanded demographic questions and focus on mental health and chronic disease. (*Am J Public Health.* 2016;106:1231–1232. doi:10.2105/AJPH.2016.303201)

By the end of 2014, 14.4 million individuals had been forcibly displaced and were living as refugees around the world, the highest number reported in nearly 20 years.[1] Refugees often flee from violence and human rights violations in unstable countries to places of exile, where nearly two thirds will spend at least 5 years awaiting a permanent home.[2,3] Living in uncertainty, without economic or social stability, refugees are vulnerable to poor health outcomes.

Since 1975, the United States has heavily invested in refugee resettlement, admitting more than 3 million refugees from around the world. Yet, relatively little is known about refugees' health status and health care needs. Existing evidence on refugee health primarily involves community-level research with small nonrandom sample sizes.[4,5] Our goal was to evaluate a representative sample of publicly available health data sets to determine whether they could be used to study the health of resettled refugees living in the United States.

METHODS

We used 2 comprehensive online sources of publicly available health research data, the Society of General Internal Medicine (SGIM) Dataset Compendium and the Inter-University Consortium for Political and Social Research (ICPSR), to identify data sets. The SGIM compendium comprises links to health service research data, and the ICPSR database includes social and behavioral health data sets.[6,7] We included in our analysis US-based data sets with publicly available documentation and a self-defined health-related focus that allowed for an examination of patient-level factors.

For each data set meeting our inclusion criteria, we analyzed publicly available documentation including study questionnaires, codebooks, publications, and online data search tools. In our search, we used a broad selection of demographic keywords associated with refugees selected by 3 researchers (including 2 of the authors) experienced in refugee health research: country, born, birth, foreign, origin, citizen, visa, immigrant, immigration, migrant, refugee, asylee, asylum, race, ethnicity, ancestry, ancestor, United States, and America. We then cataloged data sets that specifically included refugee health data. Finally, we characterized the study focus (health care access/use/costs, social determinants of health, health status/risk behaviors/outcomes) and sample construction (longitudinal, cross sectional) of each data set.

RESULTS

Of the 161 SGIM and ICPSR data sets reviewed, 68 met our inclusion criteria. Thirty (44%) of the data sets did not register any keyword matches in terms of information related to refugees. Thirty-seven data sets (54%) registered some keyword matches but did not uniquely identify refugees (Figure A and Figure B, available as supplements to the online version of this article at http://www.ajph.org). Fewer than half (n = 32) of the data sets specified whether study participants were foreign-born. Although several studies asked whether respondents were US citizens

ABOUT THE AUTHORS

Wagahta Semere and Emily A. Wang are with the Section of General Internal Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT. Katherine Yun is with the Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia. Cyrus Ahalt and Brie Williams are with the Department of Medicine, University of California, San Francisco.

Correspondence should be sent to Wagahta Semere, MD, Section of General Internal Medicine, Yale University School of Medicine, 333 Cedar St, SHM IE-61, New Haven, CT 06510 (e-mail: wagahta.semere@yale.edu). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

This article was accepted March 14, 2016.

doi: 10.2105/AJPH.2016.303201

TABLE 1—US Data Sets That Can Be Used to Study Current or Former Refugees

Data Set/Study	Sample	Focus	Keyword Match Items
National Epidemiologic Survey on Alcohol and Related Conditions	Longitudinal	Health status, risk behaviors, outcomes	Were you EVER a refugee—that is, did you flee from your home to a foreign country or place to escape danger or persecution? How old were you the FIRST time this happened? How long were you a refugee? What is your country of origin or heritage? Were you born in the United States? How many years have you lived in the United States?
New Immigrant Survey	Longitudinal	Health care access, use, costs Social determinants of health Health status, risk behaviors, outcomes	In what country were you born? What is your country of citizenship? In what month and year did you arrive in the country in which you now live? (follow-up questions detail migration history) Temporary visa categories: refugees, asylees, parolees

(n = 21), only a minority subcategorized noncitizenship status (n = 8).

Only 2 studies, the National Epidemiologic Survey on Alcohol and Related Conditions and the New Immigrant Survey, uniquely identified current or former US refugees (Table 1).

DISCUSSION

Our analysis of the nation's largest compendia of publicly available health data showed that only 2 studies (3%) could be used to study the health of refugees. We recommend expanding demographic questions in longitudinal data sets that focus on mental health and chronic disease, topics increasingly relevant to refugee health.[4,5,8–10] This can be done while maintaining sensitivity in questions pertaining to immigration status. For example, in Denmark, national data on country of birth and year of arrival can be linked to health service registry data, providing a proxy for identifying refugees in health data sets and opportunities for expanded research on refugee health.[11]

In the United States, the National Epidemiologic Survey on Alcohol and Related Conditions includes questions about self-reported refugee status, country of origin, and duration of US residency that may be appropriate for surveys focusing on mental health or access to care (Table 1). The New Immigrant Survey includes questions about country of first asylum, which may be particularly important in surveys focusing on young adults or chronic disease. In our experience, these questions may be less stigmatizing than other immigration-related questions (e.g., those intended to identify undocumented individuals).

Although we reviewed only a sample of all available health data sets, we included the most widely accessed national studies of health. Our study demonstrates that several frequently used health data sets may include refugees in their samples but fail to identify refugee status in their coding methodologies. As the number of refugees worldwide continues to grow, policies and programs that best meet the health needs of these populations will require adequate data. Collecting information about refugee status in national health data sets is a critical first step in illuminating the health care needs of refugees living in the United States.

PUBLIC HEALTH IMPLICATIONS

Information about refugee status in US national health surveys should include expanded demographic questions and should focus on mental health and chronic disease. AJPH

CONTRIBUTORS

W. Semere, C. Ahalt, B. Williams, and E. A. Wang contributed to the study concept and design. W. Semere, K. Yun, and E. A. Wang contributed to analysis and interpretation of the data. All of the authors contributed to preparation and critical reviews of the article.

ACKNOWLEDGMENTS

Katherine Yun is supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant 1K23HD082312).

HUMAN PARTICIPANT PROTECTION

No protocol approval was needed for this study because no human participants were involved.

REFERENCES

1. United Nations High Commissioner for Refugees. Global trends: forced displacement in 2014. Available at: http://www.unhcr.org/556725e69.html. Accessed April 2, 2016.

2. Refugee Act of 1980 (Pub L No. 96–212, 1980).

3. US Department of State. Protracted refugee situations. Available at: http://www.state.gov/documents/organization/157622.pdf. Accessed April 2, 2016.

4. Steel Z, Chey T, Silove D, Marnane C, Bryant RA, van Ommeren M. Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement: a systematic review and meta-analysis. *JAMA*. 2009; 302(5):537–549.

5. Fazel M, Wheeler J, Danesh J. Prevalence of serious mental disorder in 7000 refugees resettled in Western countries: a systematic review. *Lancet*. 2005;365(9467): 1309–1314.

6. Smith AK, Ayanian JZ, Covinsky KE, et al. Conducting high-value secondary dataset analysis—an introductory guide and resources. *J Gen Intern Med*. 2011;26(8): 920–929.

7. Inter-University Consortium for Political and Social Research. ICPSR home page. Available at: https://www.icpsr.umich.edu/icpsrweb/content/membership/about.html). Accessed April 2, 2016.

8. Yanni EA, Naoum M, Odeh N, Han P, Coleman M, Burke H. The health profile and chronic diseases comorbidities of US-bound Iraqi refugees screened by the International Organization for Migration in Jordan: 2007–2009. *J Immigr Minor Health*. 2013;15(1):1–9.

9. Yun K, Fuentes-Afflick E, Desai MM. Prevalence of chronic disease and insurance coverage among refugees in the United States. *J Immigr Minor Health*. 2012;14(6): 933–940.

10. Mowafi H, Spiegel P. The Iraqi refugee crisis: familiar problems and new challenges. *JAMA*. 2008;299(14): 1713–1715.

11. Norredam M, Kastrup M, Larsen-Helweg K. Register-based studies on migration, ethnicity, and health. *Scand J Public Health*. 2011;39(suppl 7): 201–205.