Creating Legal Data for Public Health Monitoring and Evaluation: Delphi Standards for Policy Surveillance

David Presley, Thomas Reinstein, Damika Webb-Barr, and Scott Burris

Background

Surveillance in public health is the means by which people who are responsible for preventing or controlling threats to health get the timely, ongoing, and reliable information they need about the occurrence, antecedents, time course, geographic spread, consequences, and nature of these threats among the populations they serve.[1] "Policy surveillance" is the ongoing, systematic collection, analysis, and dissemination of information about laws and other policies of health importance.[2]

There is a long tradition of conducting "50 state surveys" to identify laws of public health significance.[3] A recent scan identified 135 websites offering 50 state health law surveys published since 2010, constituting a substantial level of legal research work.[4] In contrast, the use of scientific methods to create datasets of legal variables suitable for use in evaluation research has emerged in the last 20 years as a result of sustained research funding for legal evaluation in key areas, most notably alcohol and tobacco control. Unlike "traditional" legal research, policy surveillance entails use of systematic quantitative or qualitative coding to create scientific datasets and track policies over time.[5] Leading examples include the Alcohol Policy Information System (APIS),[6] CDC's State Tobacco Activities Tracking and Evaluation (STATE) System,[7] and NCI's Classification of Laws Associated with School Students (CLASS).[8] In the past three years, the Public Health Law Research (PHLR) Program of the Robert Wood Johnson Foundation has been creating and publishing legal datasets on LawAtlas, a prototype policy surveillance content management system. These systems exemplify the essential elements of policy surveillance: data is collected and coded in a transparent, replicable manner, and regularly updated; data is available without charge for research; and policy information is published on the web. Collectively, this body of work has demonstrated that law can be collected and coded for quantitative research with a high degree of accuracy and with sufficient nuance to capture important legal variation.

In a 2011 Report, the Committee on Public Health Strategies to Improve Health of the Institute of Medicine suggested wider use of policy surveillance as a public health tool. It noted, however, that such an enterprise would face both administrative and methodological challenges.[9] Although collecting and cod-

David Presley, J.D., *is a Policy Surveillance Manager with Public Health Law Research. Mr. Presley manages the creation of legal datasets using PHLR's Workbench system. He holds a Bachelor of Arts in History and Anthropology from the University of Tennessee in Knoxville, TN and a Juris Doctorate from Temple University James E. Beasley School of Law in Philadelphia, PA.* **Thomas Reinstein** *is a Ph.D. candidate in U.S. History at Temple University in Philadelphia, PA. He holds a Bachelor of Arts in History (magna cum laude) from Cornell University in Ithaca, NY.* **Damika Webb-Barr, J.D.,** *is a Senior Health Policy Analyst at AmeriHealth Caritas where she facilitates the development of health policy research on industry trends that could affect current and future business. She completed her J.D. at Thomas Jefferson School of Law in San Diego, CA and holds graduate certificate in Public Health from the University of North Carolina, Chapel Hill and B.S. in Biology from George Mason University in Fairfax, VA.* **Scott Burris, J.D.,** *is a Professor of Law at Temple University, where he directs the Center for Health Law, Policy and Practice, and the Robert Wood Johnson Foundation's Public Health Law Research program. His work focuses on how law influences public health, and what interventions can make laws and law enforcement practices healthier in their effects. He is the author of more than 100 books, book chapters, articles, and reports on issues including urban health, HIV/AIDS, research ethics, global health governance, and the health effects of criminal law and drug policy.*

ing legal data for scientific research is not a new practice, only recently have explicit methods for this work been articulated.[10] The several surveillance portals have been developed independently, and differ in such key dimensions as data format and availability. As part of a larger effort to develop tools and methods for public health law research, PHLR, working in collaboration with CDC's Office of State, Tribal, Local, and Territorial Support (OSTLTS), ChangeLab Solutions, the Network for Public Health Law, and the Public Health Law Center, surveyed a Delphi panel of experts to define basic standards and practices in the conduct of policy surveillance. Results are reported here in accordance with accepted methods for Delphi studies.[11]

In the past three years, the Public Health Law Research (PHLR) Program of the Robert Wood Johnson Foundation has been creating and publishing legal datasets on LawAtlas, a prototype policy surveillance content management system. These systems exemplify the essential elements of policy surveillance: data is collected and coded in a transparent, replicable manner, and regularly updated; data is available without charge for research; and policy information is published on the web. Collectively, this body of work has demonstrated that law can be collected and coded for quantitative research with a high degree of accuracy and with sufficient nuance to capture important legal variation.

Methods

A panel of experts was surveyed using web-based forms over two rounds using a Delphi process. A Delphi process consists of multiple rounds of Likert-scale questions and comment-based feedback to establish consensus among a group of experts while maintaining objectivity and giving equal weight to all participants' responses. The survey covered four areas of policy surveillance and legal data collection: (1) the defining elements of policy surveillance, (2) conceptualizing a legal dataset, (3) the legal research process, and (4) the coding process. The main technical propositions to be considered were set out as agreement items on a five-point Likert scale. The survey also included open-ended and multiple choice questions that were intended to test the depth and specificity of agreement ("context probes"). Every item was followed with an opportunity for comment.

The process was conducted in the winter and spring of 2013-2014. Fifteen experts were selected based on their experience conducting 50 state surveys, health surveillance, policy surveillance, and/or creating legal datasets. Consensus on the main technical propositions was defined as a mean agreement of 4. Three items that came within 0.1 of reaching consensus in rounds 1 and 2 were included as consensus standards based on comments and explanations from the panel and/or answers to context probes that demonstrated agreement on the core elements of the proposition.

The survey's first round consisted of 30 Likert-scale propositions, and 14 context probes, created by the research team based on the limited literature on policy surveillance.[12] Propositions that reached consensus in the first round were deemed accepted as technical standards and were not repeated in the second round unless comments or context probes raised important issues that warranted further consideration. The survey's second round consisted of 20 Likert-scale propositions and 5 context probes. These included both repeated and new items created in response to panel suggestions or evident disagreement or confusion. For all questions in round 2, the voting information and relevant comments from round 1 were provided to the panel.

Results

The Delphi panel reached consensus on 28 elements of policy surveillance and standards for its practice. See Table 1. A scientific definition of a legal dataset as a quantitative measurement of objective attributes was accepted as the foundation of policy surveillance. The mean was slightly below the designated consensus level because two respondents were concerned that qualitative measures (such as "strength" of a law) not be excluded from the definition. There was, therefore, strong consensus on the core proposition that a legal dataset is a quantitative representation of the attributes of a law. Likewise, the steps required to build a dataset, and its standard components, reached strong

Table 1

Standards of Policy Surveillance Derived from Likert-Scale Questions with Corresponding Scores

PROPOSITION	Round 1 Mean	Round 2 Mean
Section 1: Defining Elements of Policy Surveillance		
The full text of this prompt was derived from E. Anderson et al., "Measuring Statutory Law and Regulations for Empirical Research," in A. Wagenaar and S. Burris eds., *Public Health Law Research: Theory and Methods* (2013) *available at* <http://ssrn.com/abstract=2021191> (last visited February 4, 2015). They have been truncated for brevity in this table. The following captures the key steps in creating a legal dataset: Development and Scope Systematic Collection of the Law Coding Documentation of the Research in a Codebook and Protocol Dissemination Updating*	4.4	
A legal dataset is a collection of quantitative measurements that describe the apparent features of a specified body of law across jurisdictions and/or time.*†	3.9	3.9
A legal dataset must capture the effective date of the legal text.†‡	4.2	4.3
A codebook must accompany every completed legal dataset.	4.6	
A protocol must accompany every completed legal dataset.	4.6	
Ordinarily, datasets funded by public agencies and foundations should be available to users under a Creative Commons or similar license allowing free use for research and other public-interest purposes.	4.3	
Policy surveillance findings should be available to the public on the World Wide Web.	4.3	
Section 2: Conceptualizing a Legal Dataset		
The scope of a legal dataset should be defined through an iterative process of research, analysis and expert consultation.	4.5	
A domain expert should be consulted to help define the scope of the dataset.	4.4	
More than one expert may be needed to create a legal dataset.		4.2
A domain expert should have a sophisticated professional understanding of the law being collected.	4.1	
A domain expert who understands how the law to be measured is being implemented is desirable because he or she will be able to pick elements of the law most important for evaluation.*†‡	3.4	3.9
Section 3: The Legal Research Process		
Reliable legal research for policy surveillance requires the use of multiple search strategies including keyword searches in a legal database, searching in a legal text's table of contents, or using secondary sources to identify laws.*†‡	4.4	4.5
Replicable and transparent legal research requires recording all search terms and keywords, number of search results, specific databases searched, and exclusion and inclusion criteria.†‡	4.3	4.5
Replicable and transparent legal research requires legal text to be collected and retained in a readily accessible, organized record system.†‡	4.2	4.5
Reliable legal research for policy surveillance requires redundant research.	4.1	
An explicit quality control plan should be followed when conducting research. For example, 100% redundant research should be conducted until 95% of all redundant research is consistent with the original research. Once this is achieved, 20% of additional research is redundantly researched by another person unless the consistency drops below 95%.*		4.2
Research should be conducted with 100% redundancy and subject to timely review until the research strategy and each researcher are returning consistent results.†‡	4	3.9

Continuous review of the accuracy of legal research is essential for an accurate legal dataset.	4.4	
It is essential for policy surveillance data to be kept current.	4.5	
The same research and coding procedures that were used to create the dataset should be followed for updating.	4.4	
Section 4: The Coding Process		
Coding legal data using software is superior in reliability to pencil and paper coding.*†‡	4	4.1
It is useful or desirable that the coding form (if electronic) allows legal text and coding questions to be displayed on the same screen.†‡	3.7	4
It is useful or desirable that the coding platform (if electronic) allows simultaneous coding by two or more researchers so that redundant coding or other simultaneous use can be achieved without manually merging different files.†‡	3.8	4.1
Reliable legal coding for policy surveillance requires redundant coding.	4.3	
An explicit quality control plan should be followed when coding legal data. For example, 100% redundant coding should be conducted until 95% of all redundant coding is consistent with the original coding. Once this is achieved, 20% of additional coding is redundantly coding by another person unless the consistency drops below 95%.*		4.1
Coding should be conducted with 100% redundancy and subject to timely review until scheme and each coder are returning consistent results.*†‡	4.1	4.1
Continuous review of the accuracy of coding is essential for an accurate legal dataset.	4.5	

* This prompt has been edited for clarity or brevity.
† This prompt or a permutation thereof appeared in both rounds. The texts of the prompts has been combined for clarity and brevity.
‡ This prompt or a permutation thereof appeared in both rounds. Relevant comments from the Delphi panel from the first round were reproduced with this prompt as it appeared in the second round.

consensus. The respondents agreed that defining the scope of a dataset, including the laws to be collected and the coding scheme, entailed an iterative process of research and analysis carried out with the assistance of experts in the topic area. They agreed on the need for explicit quality control procedures, including redundant research and coding; that coding with software was superior to pencil and paper methods; and that datasets should be regularly updated using the original protocol.

Probes aimed at defining more specific elements of the general standards had mixed results. There was unanimity or near unanimity on the required elements of a code book (e.g., variables, variable values) and protocol (e.g., search terms, inclusion criteria, and sources). In contrast, there was wide variance on how often datasets should be updated, with a range of as soon as a new law takes effect to one year. Similarly, while the need for a content expert with a sophisticated understanding of the topic area reached consensus, there was not consensus on specific required qualifications or experience: the nature of expertise required of a content expert could vary from dataset to dataset, and any individual's qualification to serve as an expert might not depend on whether that person was a J.D. or had field experience with the law in question.

Several reasons for this variance on specifics were apparent. Comments and responses to the probes indicated that the experts believed the general standards could be implemented in different ways depending on the project, and some respondents were explicitly averse to an overly-specific, "one-size-fits-all" methodological approach. Other panelists' comments explicitly or implicitly referred to resource concerns. Redundant research and coding, the norm in scientific research, is not common in traditional legal research and so would add to the cost of the work. Some respondents expressed concern that the use of multiple researchers and coders within a rigorous quality control framework, while ideal, might not always be feasible. Likewise, the panel agreed that datasets should be updated periodically, but feasibility considerations seemed to influence the variance on the appropriate interval, which ranged from whenever the law changed to annually.

Conclusion

This Delphi study reached high agreement on a basic set of standards for conducting policy surveillance and constructing legal datasets. The consensus standards defined in this survey formed the basis of a draft technical guide for policy surveillance created by the staff of the Public Health Law Research Project.[13] A consensus on how to best put these standards into operation is still emerging, however, and would be a fruitful topic of further discussion in the field. Panelists indicated that resources are a continuing concern. Policy surveillance on a scientific model is valuable and efficient, but it is not free. Putting the Institute of Medicine's recommendation into practice may require new resources, or the redirection of legal research resources now being used in less efficient ways.

Acknowledgements and Funding

Work on this project has been funded the Robert Wood Johnson Foundation, and by a subcontract between PHLR and ChangeLab Solutions, under Collaborating Agreement Number CDC-RFA-OT13-1302, funded by the Centers for Disease Control and Prevention. Opinions expressed here are those of the authors, and not necessarily those of the funding agencies.

References

1. J. W. Buehler, "Surveillance," in K. J. Rothman, S. Greenland, and T. L. Lash, eds., *Modern Epidemiology*, 3rd ed. (Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008): at 459-480.
2. J. F. Chriqui, J. C. O'Connor, and F. J. Chaloupka, "What Gets Measured, Gets Changed: Evaluating Law and Policy for Maximum Impact," *Journal of Law, Medicine & Ethics* 39, no. 1, Supp. 1 (2011): 21-26, at 21.
3. W. Fowler, "Principal Provisions of Smallpox Vaccination Laws and Regulations in the United States," *Public Health Reports* 56, no. 5 (1941): 167-189; C. H. Greve, "Provisions of State Laws Governing Local Health Departments," *Public Health Reports* 61, no. 1 (1953): 31-42; L. O. Gosti, "Controlling the Resurgent Tuberculosis Epidemi: A 50-State Survey of TB Statutes and Proposals for Reform," *JAMA* 269, no. 2 (1993): 255-261.
4. D. Presley and S. Burris, "A Scan of Existing 50 State Survey and Policy Surveillance Resources" (2014), *available at* <http://ssrn.com/abstract=2523011> (last visited February 4, 2015).
5. See Chriqui, *supra* note 2.
6. National Institute on Alcohol Abuse and Alcoholism, "Alcohol Policy Information System (APIS)," *available at* <http://apps.nccd.cdc.gov/statesystem/Default/Default.aspx> (last visited February 4, 2015).
7. Centers for Disease Control and Prevention, "State Tobacco Activities Tracking and Evaluation System," *available at* <http://apps.nccd.cdc.gov/statesystem/Default/Default.aspx> (last visited February 4, 2015).
8. National Cancer Institute, "Classification of Laws Associated with School Students (CLASS)," *available at* <http://class.cancer.gov> (last visited February 4, 2015).
9. The Committee on Public Health Strategies to Improve Health, Institute of Medicine, *For the Public's Health: Revitalizing Law and Policy to Meet New Challenges* (Washington, D.C.: The National Academies Press, 2011).
10. E. Anderson, C. Tremper, S. Thomas, and A.C. Wagenaar, "Measuring Statutory Law and Regulations for Empirical Research," in A. Wagenaar, and S. Burris, eds., *Public Health Law Research: Theory and Methods* (San Francisco: John Wiley & Sons, 2013): at 237-260; C. Tremper, S. Thomas, and A. C. Wagenaar, "Measuring Law for Evaluation Research," *Evaluation Review* 34, no. 3 (2010): 242-266.
11. I. R. Diamond, R. C. Grant, B. M. Feldman, P. B. Pencharz, S. C. Ling, and A. M. Moore et al., "Defining Consensus: A Systematic Review Recommends Methodologic Criteria for Reporting of Delphi Studies," *Journal of Clinical Epidemiology* 67, no. 4 (2014): 401-409.
12. See Anderson et al., *supra* note 10; C. LaFond, T. L. Toomey, C. Rothstein, W. Manning, and A. C. Wagenaar, "Policy Evaluation Research: Measuring the Independent Variables," *Evaluation Review* 24, no. 1 (2000): 92-101.
13. Public Health Law Research, "A Technical Guide for Policy Surveillance," (2014) *available at* <http://ssrn.com/abstract=2469895> (last visited February 4, 2015).