

Guidelines for Navigating Human Subjects Review and Preparing Data Sets for Sharing with the ISDS Technical Conventions Committee

Atar Baer*¹, Marcus Rennick², Aaron Kite-Powell⁴ and David Atrubin³

¹Public Health - Seattle & King County, Seattle, WA, USA; ²Marion County Health Department, Indianapolis, IN, USA; ³Florida Department of Health, Tampa, FL, USA; ⁴MIT Lincoln Laboratory, Lexington, MA, USA

Objective

The purpose of this panel is to facilitate the dissemination of surveillance-related use cases by public health practitioners with accompanying benchmark datasets to method developers. The panel will present practitioners' experiences with preparing patient-level emergency department data sets to accompany a use case submitted to the ISDS Technical Conventions Committee.

Introduction

Collaborative relationships between academicians and public health practitioners are necessary to ensure that methodologies created in the research setting translate into practice. One barrier to forging these collaborations is restrictions on the sharing and availability of public health surveillance data; therefore, most academics with expertise in method development cannot access "real world" surveillance data with which to evaluate their approaches.

The ISDS Technical Conventions Committee was established in 2013 to facilitate and expedite the development, evaluation, and implementation of technical methods for public health surveillance. The purpose of the committee is to bridge a long-standing gap between technical challenges in public health practice and solution developers needing both understanding of these challenges and representative data.

Methods

Sharing benchmark datasets can pose a significant barrier for the public health practitioner wishing to collaborate with other entities or organizations on use cases. While simulated data sets can be developed to support these use cases, these generated data sets may not be sufficiently realistic to support some types of analyses.

Even for syndromic surveillance data where the data are "limited" per HIPAA and no direct patient identifiers are included, it may be possible to infer the identity of individual patients for rare events, or by cross-referencing these data sets with other sources. The decision whether it is allowable to share these data with other entities is dependent on numerous factors, including whether the proposed activity constitutes routine public health surveillance or research, and whether the data have been sufficiently de-identified prior to release. Another barrier is that syndromic surveillance data are often acquired under data use agreements that restrict or prohibit sharing patient-level data with third parties. Even if data are not protected health information, they may be considered proprietary by the contributing entity and need to be treated as confidential.

Keeping with the theme of the conference of translating research and surveillance into action, we will discuss strategies that others can follow to help streamline the process of preparing a data set to share with the ISDS Technical Conventions Committee. We will describe the experiences of two local health departments that were collaborating to submit a data quality use case to the ISDS Technical Conventions Committee. We will review these practitioners' experiences and lessons learned while navigating the human subjects review process,

and highlight a tool that is available to help de-identify data without needing programming expertise. We will consider the technical steps required to successfully de-identify a data set and the possible implications of data manipulation on analysis. Finally, we will review and discuss legal and ethical considerations of sharing data beyond public health, even when the data have been sufficiently de-identified.

Results

1. Methods for cultivating a data set that is appropriate for sharing with the ISDS Technical Conventions Committee and ensuring that the data set is sufficiently de-identified. One talk will focus on programming and statistical strategies to de-identify data (Marcus Rennick), and another talk will review a tool that was developed to de-identify data without needing programming expertise (Aaron Kite-Powell).

2. Experiences with and strategies for navigating the IRB human subjects review process (Atar Baer).

3. Legal and ethical issues related to sharing syndromic surveillance data with third parties (David Atrubin).

A discussion will follow, inviting reactions to the talks and giving audience members the opportunity to comment on their experiences sharing data with the research community.

Keywords

emergency department; surveillance; data sharing; legal and ethical issues; data de-identification

*Atar Baer

E-mail: atar.baer@kingcounty.gov

