Protecting Highly Sensitive and Identifiable Data for Research on Vulnerable Populations
Natalie M. Matthews, MSW; Brooke Spellman; Teresa Doksum
Abt Associates, Inc.

Submission Type: Programmatic
Topic Area: Ethics & Risk
Poster Number: 27

Statement of the Problem and Intended Outcome: National evaluations of government housing programs often use identifiers to link client survey data to the sensitive administrative data of other organizations. Some of these organizations transmit to us more data than authorized, or use unsecure methods in the transfer. These types of transmission errors often violate the consent and relevant laws regarding authorization of access to this data and/or security requirements. The Evaluation of the Rapid Re-housing Demonstration (RRHD) is a study of homeless families who received US Department of Housing and Urban Development housing assistance. As part of the evaluation, the research team collected administrative data from homeless programs, including special needs (mental health, substance abuse, physical disabilities, HIV/AIDS status). This administrative data was collected from 23 evaluation sites and covered both identifiable data on consenting participants and de-identified data on non-consenting participants. We sought to minimize the risk of the loss of confidentiality and security incidents by implementing several strategies related to the data agreement and the design for data matching and sharing to collect the data.

Description of the Program: The research team and IRB worked very closely with our Information Security and Contract Operations team to develop a comprehensive Data Use Agreement that was executed with all sites for the purposes of collecting this highly sensitive administrative data. The researchers and IRB carefully designed the data to be exchanged to minimize the identifiers and data shared and reduce the risk of harm in the event of a breach. The team also established a secure file sharing site for purposes of data submission, and developed and delivered remote training for all researchers on the project and the individuals from the sites who would be transferring the administrative data. Finally, the team monitored all data submissions and implemented a protocol for identifying potential security incidents. The team worked with each site to ensure that all responses to potential security incidents were properly addressed.

Evaluation measures: Staff awareness and attitudes, numbers of staff trained, and number and nature of data security incidents.

Suggestions for Future Program Usage and Implementation: We believe that the practices of the Evaluation of RRHD can be applied to future research projects that involve the exchange and matching of highly sensitive and identifiable data between two institutions. The data use agreement, training, and response to potential incidents are all strategies that can be replicated.