Community-Based Data Share: Achieving Exponential Return on Original Research

<table>
<thead>
<tr>
<th>Journal:</th>
<th>2010 AMIA Clinical Research Informatics Summit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>AMIA-121-C2010.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Abstract Presentation</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>26-Jan-2010</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Green, Cynthia; Duke University Medical Center, Biostatistics and Bioinformatics  
                             Nahm, Meredith; Duke Translational Medicine Institute, Center for Health Informatics  
                             Shostak, Jack; Duke Clinical Research Institute |
| Classification for Review Purposes: | 3. Data Access, Integration, and Analysis |
Community-Based Data Share: Achieving Exponential Return on Original Research

Cynthia L. Green, PhD\textsuperscript{1}, Meredith Nahm, MS\textsuperscript{2}, and Jack Shostak, MBA\textsuperscript{3}

\textsuperscript{1}Duke University Medical Center, Durham, NC; \textsuperscript{2}Duke Translational Medicine Institute, Durham, NC; \textsuperscript{3}Duke Clinical Research Institute, Durham, NC

Abstract

We have developed a public-use data share web site for the NIDA CTN to share data from their completed substance abuse trials. The NIH expects and supports the timely release and sharing of final research data from NIH-supported studies. Data sharing multiplies the scientific contribution of the original research by encouraging secondary analyses. Web site logistics included the site “look and feel”, access and storage, documentation, as well as data anonymization, structure and format. A framework for necessary data share components and lessons learned are presented to aid other researchers in data share construction.

Introduction

The National Institute on Drug Abuse (NIDA) established the Clinical Trials Network (CTN) to enhance the study and delivery of novel therapies and treatments for drug addiction to patients in community-based settings in an arena that had primarily been comprised of specialized research in restricted patient populations. The CTN mission is “to improve the quality of drug abuse treatment throughout the country using science as the vehicle.”

Background

The CTN established a public-use data share web site to collect and store study data and documentation for each completed CTN clinical trial. These data can be used for secondary efficacy and safety analyses or to design future trials. These types of analyses improve the quality of drug abuse treatment by increasing the scientific contribution of the original research.

Methods

Implementing a data share portal involves various considerations, all of which must be considered in an iterative process of creating a standard data structure and portal to effectively share the data. The CTN data share considerations are outlined in Figure 1. We utilized the CDISC Submission Data Tabulation Model (SDTM) to provide data from multiple clinical trials using common data formats and structure. We also removed any patient, site or other sensitive information from each completed trial’s final database to make the data suitable for public use.

Results

The CTN data share web site launched in May 2006 and is the gateway that researchers can use to obtain substance abuse data and documentation from completed CTN clinical trials (www.ctndatashare.org). We have found that effective data-sharing includes, rendering the data into usable formats, de-identifying and anonymizing data according to HIPAA requirements, providing adequate documentation and communicating that the data are available. Whether or not data standards are considered in the design of the original data-collection instruments, a master data model should support later consolidation. This content model should be designed based on the known data collected for the therapeutic area so that the model can accommodate future trials while maintaining backward compatibility.

Discussion

Data sharing multiplies the scientific contribution of the original projects. Universal standards for data shares can further magnify this impact by providing uniformity across therapeutic areas. However, consideration should also be given to researchers who will access the data to determine the best possible data structure and format.

\textit{Funded by NIDA RFP-N01-DA-05-2207}