

TESTIMONY TO THE
DHHS CHRONIC FATIGUE SYNDROME ADVISORY COMMITTEE
Suzanne D. Vernon
Scientific Director
The CFIDS Association of America
October 28-29, 2008

Data Sharing

Data sharing is required in most academic research but is not ubiquitous. Most funding agencies, institutions, and publication venues have policies regarding data sharing because transparency and openness are considered by many to be an important part of the scientific method. A number of professional organizations are leading the way on data sharing, including the National Institutes of Health and the National Science Foundation (see <http://www.apa.org/science/data/homepage.html>).

Despite policies on data sharing and archiving, withholding of data still occurs. Authors may fail to archive data for public access, or they only archive a portion of the data set. Failure to archive data is not the only manner of withholding data; when a researcher requests additional information about study data reported at a scientific conference or publication, authors sometimes refuse to provide it. When authors withhold data like this, they run the risk of losing the trust of the science community.

Federal law

On August 9, 2007, President Bush signed the "America COMPETES Act" (or the "America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act") requiring civilian federal agencies to provide guidelines, policy and procedures, to facilitate and optimize the open exchange of data and research between agencies, the public and policymakers.

NIH and CDC have data sharing policies in place

Quoting from the final NIH Statement on Sharing Research Data posted on the NIH Office of Extramural Research Website on February 26, 2003: 'NIH reaffirms its support for the concept of data sharing. We believe that data sharing is essential for expedited translation of research results into knowledge, products, and procedures to improve human health. The NIH endorses the sharing of final research data to serve these and other important scientific goals. The NIH expects and supports the timely release and sharing of final research data from NIH-supported studies for use by other researchers.

'NIH recognizes that the investigators who collect the data have a legitimate interest in benefiting from their investment of time and effort. We have therefore revised our definition of "the timely release and sharing" to be no later than the acceptance for

publication of the main findings from the final data set. NIH continues to expect that the initial investigators may benefit from first and continuing use but not from prolonged exclusive use.'

The CDC/ATSDR also issued a policy on releasing and sharing data. The policy was originally dated 04/16/03 and updated 09/07/05 and can be found at the CDC's OD website (<http://cdc.gov/od/foia/policies/sharing.htm>). Here is an excerpt of that policy. 'CDC believes that public health and scientific advancement are best served when data are released to, or shared with, other public health agencies, academic researchers, and appropriate private researchers in an open, timely, and appropriate way. The interests of the public—which include timely releases of data for further analysis—transcends whatever claim scientists may believe they have to ownership of data acquired or generated using federal funds. Such data are, in fact, owned by the federal government and thus belong to the citizens of the United States. '

'However, although CDC recognizes the value of releasing data quickly and widely, CDC also recognizes the need to maintain high standards for data quality, the need for procedures that ensure that the privacy of individuals who provide personal information is not jeopardized, and the need to protect information relevant to national security, criminal investigations, or misconduct inquiries and investigations. The goal is to have a policy on data release and sharing that balances the desire to disseminate data as broadly as possible with the need to maintain high standards and protect sensitive information. '

Data sharing success stories

CDC approved the sharing of the chronic fatigue syndrome Wichita Clinical dataset with 25 intramural and extramural investigators in the CFS Computational Challenge and with CAMDA (Critical Assessment of Microarray Data Analysis) in 2006 and 2007. Over the past 3 years, sharing the Wichita Clinical dataset and partnering with expert extramural scientists resulted in more than 20 publications and there are more manuscripts delving into this dataset in the works. CDC's CFS Research Program has published about 15 papers based on the Wichita Clinical study in the past 3 years. In addition to yielding "more value for the money," this data-sharing effort generated new perspectives on CFS and confirmed previous observations about neuroendocrine and immune dysfunction in CFS.

Existing data

NIH has funded CFS research since 1988. The NIH CRISP database was searched from 1988 to present for all projects listing chronic fatigue syndrome in the abstract or thesaurus.

Investigators	Pubmed Publications by	CFS-Specific Publications by
---------------	------------------------	------------------------------

funded for CFS and/or applicable research	these Investigators	NIH-funded investigators
75	4250	900

CDC has conducted CFS research since the early 1990s. The CDC CFS Research Program is in the Division of Viral and Rickettsial Diseases, Chronic Viral Diseases Branch with Dr. William C. Reeves serving as branch chief since 1989. Pubmed was searched for all CFS relevant publications with Dr. WC Reeves as a co-author resulting in an estimated 50 publications with the first one published in 1992. The following table lists studies that have been conducted or funded by the CDC's CFS Research Program from 1997 to 2007 (studies were identified using the Federal Register and OMB and is likely not a comprehensive list).

Study Name	Start date	Publications
Wichita Longitudinal	1997	25
National Survey	2001	1
Wichita Clinical	2002	25
Interferon Alpha Induced Sickness	2000	11
Dose Response of CRH	2002	0
Impact Study of Recruitment	2006	0
Georgia Survey	2004	3
Georgia Clinical Study (Emory GCRC)	2005	0

The Solution

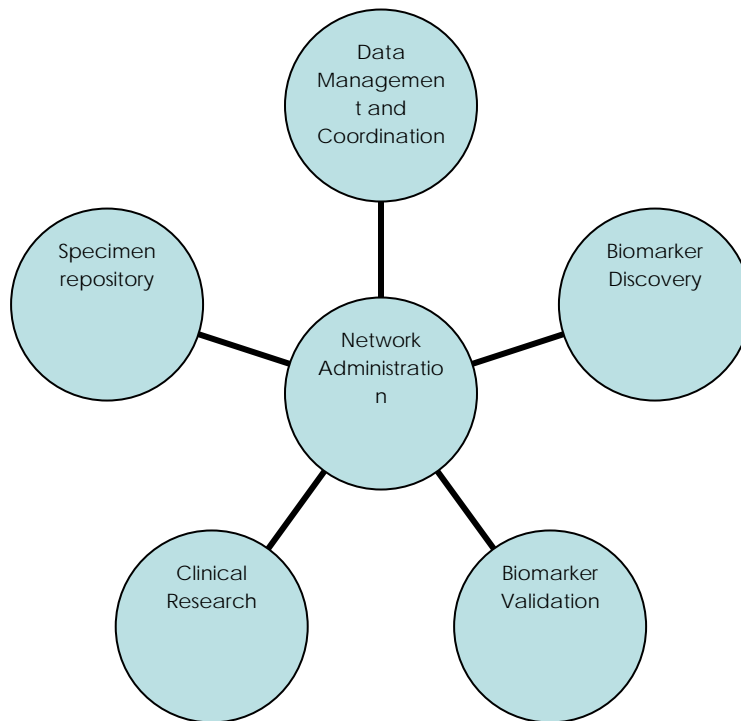
An enormous amount of data has been collected on CFS over the past 20 years. This is partially reflected in the more than 4,000 CFS specific articles that can be found in PubMed. However, we have not gotten the best value for the Federal dollars that have

been invested into CFS research as reflected by the lack of publications, measureable outcomes and products.

There is an immediate solution. We can gather, make publicly available, and use readily available information technology approaches to organize and better understand the data that has been collected to date. This will allow us to establish a valid and evidence-based CFS knowledgebase that can be used to direct further CFS research.

Further, there should be a required inventory of all specimens that have been collected by federal agencies and federally funded investigators. These specimens should be deposited into a biorepository and managed accordingly.

The 20 years of information collated in the form of a CFS knowledgebase CFS would direct the overall objectives of a collaborative CFS research network. The ultimate goal of this research network would be effective strategies for treatment and control by describing risk factors, identifying biomarkers, and elucidating CFS pathophysiology. Many examples of successful research networks exist and many have the following structure:



Such a network should be administered by an entity with proven accomplishments, accountability to stakeholders and a clear understanding of the CFS research priorities. The NIH would provide both intramural and extramural funding for the clinical and biomarker discovery research. CDC would provide extramural support for biomarker

discovery investigations and intramural support for biomarker validation as well as specimen repository. Data management and coordination would occur by a public or private entity with familiar with CFS and proven record of data management. Private and academic partnerships would be central to the overall success of such a research network.

This solution is best value for our tax payer dollars dedicated to furthering CFS research and could be a model for the study of other complex chronic conditions that impose such a burden on the individual, the community and the nation.