

**Panel on Disclosure Review Boards of Federal Agencies:  
Characteristics, Defining Qualities and Generalizability**

Presented at the  
Joint Statistical Meetings  
August 17, 2000 Indianapolis, Indiana

## Contents

[Introduction](#). Fritz Scheuren, The Urban Institute

[The Disclosure Review Board of the Bureau of Census](#). Easley Hoy

[The Disclosure Review Board of the National Center for Health Statistics](#). Alvan O. Zarate

[The Disclosure Review Board at the Bureau of Labor Statistics](#). George D. Stamas

[The Disclosure Review Board of the National Center for Education Statistics](#). Marilyn M.  
McMillen

[Discussion on Federal Review Boards](#). Gene D. Therriault

## **Introduction**

Fritz Scheuren, The Urban Institute

### **Background**

Federal agencies produce an enormous amount of electronic and other types of data in meeting their mandates to disseminate information to the public. Along with this obligation, is the equally compelling requirement to protect the privacy of those who have provided information to these agencies. Sophisticated means for disseminating data carry with them a risk that this same technology could be used to compromise the confidentiality assured those providers.

Methods for minimizing this risk have been available for some time, and in the past these have been applied by one or another agency staff member. If an agency has multiple or large number of data sets to review, it is necessary to organize a “team” of persons with the expertise needed to examine information before it is released to the public. In the federal government, these panels are usually referred to as Disclosure Review Boards (DRBs).

### **The Panel**

This panel is composed of members of the DRBs of four federal agencies (the Bureau of the Census, the National Center for Health Statistics, the Bureau of Labor Statistics and the National Center for Education Statistics). A fifth panel member is the executive secretary of the New York State Department of Health’s Data Protection Review Board.

Increasingly, the need has been expressed to learn more about these panels. Many persons would like to know what is involved in assembling the needed skills, reference materials, and documents. Not all are faced with a great number of data sets that must be released with the proper privacy protection, but it is precisely those who lack the experience of the agencies represented here, that feel most inhibited and unsure about releasing data in today’s environment.

### **The Panel’s Charge**

Each panel member was asked to cover the following seven key elements:

1. principal objectives;
2. kind of data reviewed;
3. intended use of data;
4. organization of the DRB;
5. decision making process;
6. types of materials reviewed; and, finally,
7. statistical disclosure methods used.

In addition to laying out the principal features of the DRB, each panelist was asked to address the question of the generalizability of these features to other settings. For example, to what extent could these functions be carried out by state or federal agencies that produce data files only occasionally? How about an Institutional Review Board, or a University (department), a hospital or health plan? How much adaptation would be necessary? This second goal may have been overly ambitious and is an ideal topic for a follow-up panel, but at least this panel will have

represented a solid initial step in making the experience of its members available to a wider audience.

One thing not really in scope for the session was the relationship of a DRB with other independent review mechanisms that also have a role to play. Two examples would be an Institutional Review Board (IRB), whose focus was primarily privacy, or a Computer Security Oversight (SCO) group, whose function was to provide a review of the physical arrangements made to be sure that confidentiality pledges given to respondents were met.

## **The Disclosure Review Board Of The Census Bureau**

Easley Hoy, U.S. Bureau of Census

### **Background**

The U.S. Bureau of Census (BoC) collects data under U.S. Code Title 13, Chapter 1, Section 9, which pledges confidentiality for data collected from its respondents. Such data collected under Title 13 are also exempt from Freedom of Information Act inquiries. The BoC is also charged to disseminate data for the determination of government policy and academic research. Such data includes demographic data (both tabular and microdata), economic data (tabular only), special tabulations, and more recently metadata. Demographic microdata have been provided since the 1960s. Given both the legal responsibility to protect confidentiality of its respondents and its legal responsibility to provide data to policy makers and their advisors, it was obvious that the BoC needed oversight of all publications for disclosure review and analysis prior to their release. For microdata, it established the Microdata Review Panel in 1981, and for macrodata, the review was done independently by the economic area and decennial area respectively. To insure disclosure oversight over all BoC data products and maintain consistent disclosure limitation policy both internally and over time, the BoC established the Disclosure Review Board (DRB) in 1995 which has responsibility for all data products.

### **Disclosure Review Board (DRB) Charter**

The purpose and functional responsibilities of the DRB are:

1. To establish and review the Census Bureau's disclosure-limitation policy regarding all publicly available data products.
2. To review and approve proposed disclosure-limitation procedures for the release of all publicly available data products.
3. Within the framework of the aforementioned disclosure policies, the DRB is responsible for communicating its disclosure-limitation policy to program managers, Census Bureau officials, data users, and prospective sponsors.
4. To initiate and coordinate research on the disclosure potential in microdata and tabular data and the effectiveness of disclosure-limitation techniques.
5. To revise the Census Bureau's disclosure-limitation policy as necessary.

### **Types of Data Products Reviewed**

The BoC releases publications (electronic and paper) which include tables, reports, and press releases. For publications, the DRB reviews the specifications for detailed tables for decennial and demographic data and determines the specifications/parameters for the automated cell suppression programs for the tables of economic data. The BoC also releases public use microdata files, and for these files the DRB reviews the specifications and record layouts for the decennial Public Use Microdata Sample (PUMS) files and the demographic current survey files (PUFs). In addition, the BoC upon formal request and funding, provides users special tabulations which must also satisfy DRB disclosure limitation guidelines. The BoC's Research Data Centers must also abide by the DRB's disclosure limitation guidelines.

### **DRB Administrative Process**

The DRB is composed of 9 voting members and several alternates. Three of the voting members are permanent members. The permanent members include the chairperson, who is a senior researcher in the Statistical Research Division (SRD); the secretariat, who is a representative from the Policy Office; and the disclosure-limitation principal researcher in SRD. The remaining six voting members are rotating members from the subject matter directorates. Two members are from the Demographic Directorate; one member is from the Decennial Directorate; and three members are from the Economic Directorate. One of the three members from the Economic Directorate also oversees the disclosure review function at the BoC's Center for Economic Studies and its Research Data Centers. The rotating members usually serve a three year term. The current alternates are from the SRD and the Policy Office.

When a data provider wants to submit a disclosure clearance request, he/she writes a formal memorandum to the chair of the DRB. The DRB has a regularly scheduled weekly meeting for one hour at the same place and time to review such clearance requests. The DRB averages about two requests per week, and the chair schedules each request as an agenda item. If there is not a quorum of members or there are no requests scheduled on the agenda, then the meeting is canceled. The formal requests are usually accompanied by a Bureau disclosure checklist, the questionnaire, a list of variables of interest, a record layout (microdata), table outlines, some cross tabulations, and some sample design identifiers (if applicable). The Bureau disclosure checklist asks the requestor basic questions about the content of the files to be released so that disclosure issues can be determined for discussion by the DRB. Likewise, the other information accompanying the formal request is used by the DRB to determine whether there are any other disclosure issues for discussion. Such materials are usually distributed to the DRB members several days prior to the meeting for review and discussion. Recently, a few requesters have initiated preliminary discussions with the DRB about possible disclosure issues during project planning and discussions with sponsors. This is an attempt to resolve significant disclosure issues prior to the actual collection and processing of the data and make the process more efficient.

Each agenda item (request) is presented to the DRB. If the request is a recurrence of a previously approved request with minimal changes, the chair presents the request to the DRB for discussion and approval. If the request is a recurrence of a previously approved request but has significant changes, or if the request is a new request, the chair may invite the requestor to attend the DRB meeting and present the request in person to the DRB. This attendance enables the requestor to make a personal plea for clearance and to answer any questions that the DRB may have regarding the request. After the presentation, there is discussion among the DRB members on assessing the risk of disclosure, and if necessary, suggesting possible solutions to gain approval. After the discussion, the members vote and the final decision of approval/denial of the request is made by consensus. If any member has strong objections against approving the request, then the DRB has further discussions to find a satisfactory solution so that the objection will be resolved. If the DRB is undecided, then it may request additional information for further discussion. When the final decision is made, the chair writes a formal memorandum to the requestor informing him/her of the decision. If the decision is a denial of the request, then there

is an appeal process to the Disclosure Executive Steering Group, which is composed of the Principal Associate Director for Programs, the Associate Director for Methodology and Standards, the Associate Director for Demographic Surveys, the Associate Director for Economic Censuses and Surveys, and the Associate Director for Decennial Censuses.

### **Disclosure Limitation Research Staff & Research**

The Disclosure Limitation Research Staff is composed of a principal researcher and three other researchers; one researcher is designated for each of the subject matter areas (decennial, demographic, and economic). The staff conducts internal research on disclosure limitation techniques such as those described in Office of Management and Budget's (OMB) [Statistical Working Paper # 22](#), dated May 1994. These techniques include determining minimum geographic thresholds, identifying possible skewed distributions, establishing criteria for top codes and bottom codes for continuous variables, determining criteria for recoding and collapsing categories, random swapping, identifying cells for suppression, etc. The research staff also reviews and monitors current external disclosure limitation research that is being conducted by academia and other researchers.

### **Generalization**

In this information and Internet age, not only is there more information and greater access to such information for analysis, but also there is greater concern about the privacy and confidentiality of respondents' data. Therefore, more organizations, both private and government, need the disclosure limitation function, both to provide the needed analytical data and to protect the confidentiality of its source. This function has four aspects that include policy, administration, technical research, and subject matter expertise. This function can be performed by an individual or a group, depending on the volume and frequency of data requests. A generic checklist for raising disclosure issues has been developed and shared among the Federal statistical agencies. Those agencies have been encouraged to customize this generic checklist to meet their specific needs. The checklist has six sections and two appendices that asks questions and discuss various data disclosure risk situations and limitation methods. The six sections are 1) an introduction, 2) a cover sheet with basic information (e.g., survey name, project manager, organization, type of data, etc.) about the proposed data release, 3) microdata file and contents 4) demographic tabular data 5) establishment/organization tabular data and 6) selected references. The appendices provide a summary of statistical disclosure limitation methods application by type of release and some definitions of selected disclosure limitation methods. The checklist is a very useful instrument to educate and bring awareness to any organization about the disclosure and confidentiality issues pertaining to the dissemination of data. Also the generic disclosure limitation techniques are published in the OMB's [Statistical Policy Working Paper # 22](#), May 1994, and are available for any organization to use (see [www.fcsm.gov/](http://www.fcsm.gov/)). A generic automated cell suppression and auditing program has been developed for reviewing tables for disclosure limitation, and it is also available for any organization to use. Therefore, generic disclosure limitation procedures and tools are available to the general public to be customized for each specific use.

## **The Disclosure Review Board of the National Center for Health Statistics**

Alvan O. Zarate, National Center for Health Statistics

### **Background**

In recent years, the volume and frequency of electronic data products produced by NCHS has grown rapidly. At the same time the number and availability of external data bases that can potentially be matched to our public use files has grown even more. So too has the technological ability of potential abusers to develop and implement complex intrusion strategies. Adding to the complexity of confidentiality review of NCHS files is the fact that several surveys (including one conducted by another agency - AHRQ) now share the same sampling frame with the National Health Interview Survey (NHIS). Thus, review of public use files produced by the National Survey of Family Growth, the National Health and Nutrition Survey, NHIS and the Agency for Health Research and Quality's Medical Expenditure Panel Survey - all of which share respondents or sampling frames - must take into account possible interconnections with each other.

Considering the extreme sensitivity of much of the data collected, the increased public awareness of and concern for privacy, and NCHS' undiminished legal and moral obligation to fulfill its' guarantee of confidentiality, an organized, well coordinated and statistically sound procedure for establishing acceptable levels of disclosure risk is required.

This important task was, until 1998, the responsibility of the NCHS Confidentiality Officer who relied on the assistance of knowledgeable staff on an ad hoc basis. As the volume, complexity, and variety of data files requiring review increased, this arrangement proved unsatisfactory. While the Confidentiality Officer remains responsible for insuring that the Center's policies are observed in a statistically sound and legally permissible manner, the implementation of a formal mechanism for the review of files for release outside NCHS facilitates input on each data file from specialists in both sampling and survey statistics.

In addition, such a mechanism further facilitates timely and standardized review of statistical files in a manner that involves more staff in the process and encourages, thereby, a Center-wide capability and sensitivity to issues of confidentiality and disclosure limitation practices.

To this end, and guided by established procedures at other statistical agencies such as the U.S. Bureau of the Census and the National Center for Education Statistics, a NCHS Disclosure Review Board was established in 1997. It began operations in the following year.

### **Organization**

The Board is chaired by the NCHS Confidentiality Officer and, in addition is composed of the Chair of the NCHS Confidentiality Committee (who is also head of the Center's Research Data Center), a representative of the Office of Research and Methodology, and two other members selected at large. Alternates are named for the last three who serve for two years. A representative of the program producing the file to be reviewed is an automatic ad hoc member



and additional persons from within and outside of NCHS may be consulted as needed.

The DRB reviews micro data files for both public use, interagency sharing and other authorized release together with selected tabular materials following procedures established by the Confidentiality Officer.

Although DRB procedures are still under development, the following is a description of the manner in which micro data files are currently considered for release outside the Center:

### **Materials Reviewed**

After informing the Confidentiality Officer of plans to release a public use or other file, the requesting program is provided with an electronic version of the NCHS Disclosure Potential Checklist<sup>1</sup>. This document (patterned after one in use by the Census Bureau for many years and adapted for use at NCHS) contains a detailed description of potential problem areas for micro-data files together with suggestions for addressing those problems. The program, after having reviewed the proposed file in conjunction with the standards contained in the checklist, then fills out the checklist and submits it, together with file documentation, survey background and other essential documents to the DRB. In addition to materials specific to the file under consideration, the DRB has recourse to minutes of previous meetings at which related files were considered and decisions made.

### **The Review Process**

While the DRB initially met on an intermittent basis to consider proposed files, the schedule was changed to the first Tuesday of each month beginning in November of 1999. At these meetings, program staff provide a brief overview of the survey and the file. This is followed by discussion with DRB members of any problem areas together with possible approaches. The Board then meets in closed session to reach a final decision. Final decisions are not always made at the initial meeting; additional information requiring specific tabulations may be requested to clarify critical issues.

Advance notice of these procedures are provided to NCHS staff. The results of the Board's review are used by the Confidentiality Officer in reaching a final decision, and that decision is expeditiously shared with the program. In reaching decisions, every effort is made facilitate release of proposed files while conforming to NCHS' confidentiality law and practice.

Not all files are reviewed by the full DRB. At the discretion of the Confidentiality Officer, certain files (e.g. supplementary survey files following a full review of the main survey by the DRB) are reviewed by him in consultation with the program.

---

<sup>1</sup> A generalized form of the checklist can be found at <http://www.fcsm.gov/cdac/index.html>. This checklist is the product of the Federal Committee on Statistical Methodology's Confidentiality and Data Access Committee (CDAC).

In addition to extending the range of input available on confidentiality matters, the DRB has promoted an institutional capability for statistical disclosure limitation techniques as well as a broader awareness of NCHS confidentiality policies and practices.

### **Conclusions**

The above procedures have benefited NCHS efforts in a number of important ways by:

1. Providing an institutional capacity for the standardized review of micro data files and other materials. This is no longer the responsibility of only one person.
2. Producing an ongoing record of materials considered and decisions made. With this documentation, the DRB will need not consider each issue separately. Having the benefit of previous discussions and documentation the DRB can build upon a greater range of considerations. In addition the Center has created a record for purposes of accountability and future review.
3. In providing for member turnover, the DRB represents an important mechanism for the development and decentralization of disclosure limitation skills among NCHS staff.

### **Generalizability**

While the extent to which these procedures can be duplicated in other circumstances is clearly dependent upon both technical and financial resources, the principles underlying them should be evaluated for their applicability to other settings. Among the most important are:

- a) the need for certain kinds of expertise: mathematical and that related to the types of data under consideration for release. Proper disclosure review and analysis requires the appropriate background and experience.
- b) the principles of disclosure limitation are available and embodied in the checklist and other documents included in its reference materials.

Other arrangements are possible (ad hoc boards, modified Institutional Review Boards, contracts with qualified agents). This panel has discussed one approach used by four federal agencies and proposed in a number of others.

## **The Disclosure Review Board at the Bureau of Labor Statistics**

George D. Stamas,\* U. S. Bureau of Labor Statistics

As survey designers, we go through great efforts to collect data and produce accurate estimates efficiently, to serve the needs of policy makers and the public. At the same time, we often extend a pledge of confidentiality to our respondents. Many federal agencies that collect and disseminate data from the public have Disclosure Review Boards. With this panel, we have an opportunity to present the experiences from the Boards of three federal agencies. The current incarnation of the Disclosure Review Board at the Bureau of Labor Statistics (BLS) is relatively recent. Over the past year, the board has been discovered and is actively sought for advice and guidance by programs at BLS.

I am going to provide background on disclosure review, the charter and organization of the BLS Disclosure Review Board, the types of materials that the Board has been asked to review, some of the issues facing Disclosure Boards and whether this model can be applied to other agencies.

### **Background**

A standardized, defensible policy for limiting disclosure across programs within an agency and even across agencies would seem a desirable goal. It promotes a more efficient use of government resources. It makes decisions appear less capricious and void of favoritism. Most BLS Surveys are business establishment surveys. Complicating the disclosure review process for BLS, many of the Bureau's establishment surveys are collected in cooperation with state agencies in what the Bureau calls Federal/State Cooperative Programs. This includes Current Employment Statistics, Occupational Employment Statistics and Unemployment Insurance Covered Employment and Wages, the program that provides the sampling frame for most Bureau establishment surveys.

As computing power has grown, so it appears has interest in attaining micro data files or specially designed tabulations of survey estimates. Bureau management chartered the Micro Data Access Review Team to increase access to BLS micro data, while protecting confidentiality (BLS, 1999). They were given the mission of "easing the tension between access and confidentiality" and this would likewise be the goal of all Disclosure Review Boards. This team produced a report in June 1994, proposing methods to allow limited access to BLS micro data and many of their proposals have been adopted and implemented. The team was also designated as the Bureau's interim Disclosure Review Board. The team did not address the issue of disclosure in BLS tabular data. Though they varied by survey, BLS programs did have methods in place to avoid disclosure of information provided by survey participants. With advances in technology and more focus on tabular presentation, in the late Nineties the BLS Quality Council chartered the Disclosure Review Team to develop standard disclosure procedures for BLS programs. They reported in April 1999 and recommended the establishment of an active Disclosure Review Board (BLS, 1999). This recommendation was accepted with the

---

\* All opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.

establishment of a Disclosure Review Executive Committee and the Disclosure Review Board.

### **Organization**

The Disclosure Review Executive Committee is comprised of BLS Associate and Assistant Commissioners. Three program offices are represented: Prices and Living Conditions; Compensation and Working Conditions; and Employment and Unemployment Statistics. The Committee also has members from Administrative Services, Publications and Survey Methods Research. The Executive Committee presents issues or questions to the Disclosure Review Board for consideration and it receives recommendations from the Board and decides whether to accept them, reject them or send them back for further consideration, clarification or justification.

The Disclosure Review Board has nine members: an economist and a mathematical statistician from each of the 3 program areas represented on the Executive Committee; a mathematical statistician from the Office of Survey Methods Research; a representative from Administrative Services to assist in finding legal interpretations; and a chair who by charter is the Director of the Mathematical Statistics Research Center. The tenure of the members is yet to be determined, although it is intended to be limited. Members will likely be replaced on a staggered basis.

From time to time the DRB designates subgroups of members to address specific issues. Any time that a subgroup is considering an issue specifically related to one of the program offices, the group will be comprised of mathematical statisticians and economists from the other programs and the Office of Survey Methods Research. Members from the affected program area will be available to provide program related technical expertise.

The DRB began with weekly meetings in an effort to get a quick start and now is on a schedule with meetings about every two weeks. The subgroups also schedule meetings and considerable communications goes on between meetings through e-mail. The DRB documents its meetings with minutes that are available to the Executive Committee. Votes are by simple majority, though the Board strives for consensus. Dissenting views are to be documented and made available to the Executive Committee. The Chair will announce a scheduled vote a week in advance. A vote can only be taken if a minimum of five members are present and this must include two of the statisticians and at least one economist from a program office. For convenience, the Chair will occasionally poll the Board by e-mail, but in such a case all members must participate in the decision. Memoranda from the Board Chair convey any decisions and any policy recommendations are put before the Executive Committee for approval.

The charter defines the DRB mission as follows. The DRB is to establish and update BLS disclosure limitation policy regarding all publicly available data products. It reviews and approves disclosure limitation procedures proposed by statistical program offices for the release of all publicly available data products. This requires the review of proposals and specifications to assure adherence to policy established by the DRB and approved by the Executive Committee. The DRB is responsible for communicating those policies to program managers. It also initiates research on disclosure potential in micro and tabular data and assesses the effectiveness of

disclosure limitation techniques. Finally, the DRB is to further develop and update, as needed, the policy for sharing confidential micro data for independent statistical research.

### **Procedures for Review**

At times the Executive Committee will ask the DRB to consider a specific policy issue. Most of the requests presented to the DRB are from program offices for guidance in responding to requests for unpublished data. These requests are directed to the DRB Chair by a program manager through an official memorandum, along with supporting documents. The DRB will review and discuss the materials. On occasions, the Board will ask a program manager or analyst to attend a Board meeting and present additional information or answer questions about the requested data set. The Board provides and receives feedback until they feel comfortable making a decision for or against the release of the data. The decision will at times put constraints on release of the data, treating it in some manner to limit disclosure risk. The review process can take as little as a week or as long as several months depending on the nature of the request. The DRB will evaluate data sets in terms population and sample composition and an n,k or p-percent rule. Some of the conditions or constraints put on release of data set considered are collapsing table cells, or number of groups in categorical data, top or bottom coding, stripping or recoding identifiers or items, and cell suppression. The Board has followed a practice of recommending that program offices provide as much of what is requested as possible, but no more.

### **Types of requests to date**

The DRB is to research and make policy recommendations, provide guidance to BLS programs in implementing the policy and as well as guidance in responding to specific requests for previously unreleased BLS data. Over the past year, the Board made the recommendation, concurred with by the Executive Committee, that BLS move to the p-percent standard for disclosure avoidance (OMB, 1994). This decision also conveyed that respondents were to be assured of an interval of protection about their response rather than protection from exact disclosure alone. The Board is reviewing software products that will assist programs in implementing this standard. The Board has also reviewed specifications from one of the Bureau's larger programs for a system that is design to apply the p-percent standard. The Board has not been involved in reviewing publication tables. It is the responsibility of the program offices to implement the disclosure avoidance methods adopted by the Bureau.

The Board has reviewed numerous requests for the release of micro and tabular data and one for survey design data. Micro data requests included one for data from a study conducted by the BLS Cognitive Research Lab. The researcher wanted to study seam effects, but had no specific interest in items purchased by the respondent. The DRB determined that if the item names were suppress along with any other respondent specific identifying information the data set did not pose a disclosure risk. Another micro data request asked for longitudinal data on price quotes from retail outlets. The DRB determined that if the program office reassigned outlet codes in a manner such that groups of items could not be determined as coming from the same outlet, then the data set would not pose a significant disclosure risk.

Requests for tabular data were very similar. Many of these were for price indexes on specific items and the DRB response was based on sample counts and the approximate size of the universe of possible outlet quotes for the items. In one that asked for data by state or Census region the DRB decided that state data might pose a disclosure risk because of the nature of the items requested. The board OK'd the release of regional data with few constraints. These are only a few examples of what the DRB has been asked to review and how they responded.

The balance of the year has been spent with two subcommittees developing guidelines for the release of micro data and tabular data.

Several issues continue to surface as the DRB goes about its work. One of these issues is applying disclosure limitation practices to sample data. Sampling is thought to provide protection against disclosure as long as specific identifying information such as names and addresses is not released. This is less the case for establishments surveys, where the probability of selection increases to certainty depending on measures of size. Some programs may not have resources to screen data for disclosure across all items collected so they will try to protect the respondents based on measures from a key item that they collect, like employment, wages or revenues. The properties of this approach are not clearly understood. Another question is how to identify records with characteristics so unique that releasing information on those characteristics amounts to revealing the identity of the respondent. And then there is a whole set of issues related to matching or otherwise comparing files. Programs should compare any request against previous releases from the same data set. Not only are the logistics of this difficult, it can create a scenario where data requesters determine the level of detail eventually released by a program, along the possible dimensions of the data set, on a first come basis. Another related issue is longitudinal data. Finally, while not considered a big risk there is the potential for linking information in a BLS data set to other data sets available to the users.

### **Generalization**

While every agency faces unique issues in pursuing disclosure avoidance, the establishment of the DRB at BLS appears to rely on the fact that many of the approaches can be applied across agencies. As mentioned at the outset of this presentation, this DRB at BLS is relatively new. We look at how other agencies have implemented boards and benefit from what they have learned. Likewise, there can be benefit from sharing disclosure avoidance methodology, stories of software development efforts and the software itself. BLS, along with the other agencies represented on this panel, participates in the [Committee for Data Access and Confidentiality](#) (CDAC) as one forum to exchange these ideas. We also participate at professional conferences such as this one to share information on this topic.

### **References**

Executive Office of the President, Office of Management and Budget (OMB) (1994), [Statistical Policy Working Paper 22](#), Report on Statistical Disclosure Limitation Methodology.  
U.S. Bureau of Labor Statistics (BLS), internal document, Final Report-Disclosure Review Team, April 1999.

## **The Disclosure Review Board of the National Center for Education Statistics**

Marilyn M. McMillen, National Center for Education Statistics

The National Center for Education Statistics (NCES) is Congressionally mandated to "collect, analyze, and disseminate statistics" related to education in the United States. In so doing, NCES is required by law to develop and enforce standards designed to protect the confidentiality of individually identifiable respondents.

### **LAWS**

There are three laws that pertain to confidential data at NCES:

- § The Privacy Act of 1974;
- § The Computer Security Act of 1987; and
- § The National Education Statistics Act of 1994.

The first two laws pertain to all Federal agencies. Specifically, the Privacy Act of 1974 protects the privacy of personal data maintained by the Federal government. It imposes specific requirements on Federal agencies to safeguard the confidentiality and integrity of personal data, and limits the uses of these data. The Federal Information Processing Standard Publication (FIPSPUB) 41, Computer Security Guidelines for Implementing the Privacy Act of 1974, provides guidance to ensure that government-provided individually identifiable information is adequately protected in accordance with Federal statutes and regulations. Unlawful disclosure is a misdemeanor and is subject to a fine up to \$5,000.

The second law, the Computer Security Act of 1987, relates to sensitive information which is defined as any unclassified information which could adversely affect the:

- § National interest;
- § Conduct of Federal programs; or
- § Individual privacy under the Privacy Act of 1974.

This law requires each Federal agency to identify all Federal computer systems that contain sensitive information and implement security plans to protect these systems against loss, misuse, disclosure, or modification. Unlawful disclosure is a misdemeanor and is subject to a fine up to \$5,000.

The third law, the National Education Statistics Act of 1994, pertains only to NCES. This Act replaces a 1988 law that contained the same confidentiality provisions. It authorizes NCES to collect and disseminate information about education in the United States. The Act incorporates and expands upon the Privacy Act of 1974 by requiring strict procedures to protect the privacy of individual survey respondents. The Act requires NCES to develop and enforce standards to protect the confidentiality of persons in the collecting, reporting, and publication of data. The Act also protects the confidentiality of individual schools in the National Assessment of Educational Progress (NAEP) data.

More specifically, under this law, no person may:

- § Use any individually identifiable information for nonstatistical purposes;
- § Make any publication whereby the data furnished by any particular person can be identified; or
- § Permit anyone other than the individuals authorized by the NCES Commissioner to examine the individual reports.

A confidentiality violation of this law is a class E felony, punishable by up to five years in prison, and/or a fine up to \$250,000.

## **DISCLOSURE REVIEW BOARD**

The 1988 Hawkins-Stafford Amendment required NCES to establish procedures for disclosure risk analyses of proposed public use micro-data files that contain any individually identifiable data. In October of 1989 the first NCES Disclosure Review Board (DRB) was convened.

The DRB is charged with responsibility for reviewing disclosure risk analysis documentation for any data files prepared for public release that include individually identifiable data, and for reviewing and evaluating proposed changes to data files intended to make them suitable for public release. In evaluating the risk of disclosure, the Board considers the resources required to disclose individually identifiable information, the age of the data, the reliability and accessibility of external files, and the detail and specificity of the data. Following the completion of a review, the DRB prepares and submits a Decision Memorandum to the NCES Commissioner for the proposed public use data file.

### **Organization and Process**

The Board is chaired by a senior mathematical statistician from the Chief Statistician's staff, and is comprised of the Chief Statistician, senior technical reviewers from each Division (four), and a representative from the Census Bureau. The NCES representatives serve by virtue of their positions, and thus do not rotate. There is no dedicated staff for the DRB, each DRB member has multiple responsibilities; thus, the Board meets on an as needed basis, with much of the communication accomplished electronically. As soon as a report is submitted it is distributed to DRB members with a due date for comments.

If any member requests a meeting for discussion of the report, a meeting is scheduled. DRB decisions are reached by voting, however every effort is made to reach consensus before a recommendation is submitted to the Commissioner for review.

### **Type of Data Products Reviewed**

Reviews requiring the full Board are limited to micro-data files. Program staffs submit disclosure risk analysis reports and supporting documentation. Each report identifies external data sets that could be used for matching, summarizes the disclosure risk analysis procedures and/or software used and the variables that were analyzed. Each report also summarizes the results of the analysis and describes any recoding, data suppression, swapping, or statistical perturbations that occurred as a result of the analysis. Each report must be accompanied by supporting tabulations from the proposed public use data file.



As described above, the DRB members review the submission prior to making a recommendation to the Commissioner. After 10 years of DRB reviews, there have been no documented disclosures resulting from NCES public use data files.

Although reviews of tabulations are not conducted by the full Board, NCES statistical standards specify that:

" . . . each publication cell must have at least three (unweighted) observations in it and subsequent tabulations (e.g. crosstabulations) must not provide additional information which could disclose individual identities." (NCES Statistical Standards, 1992).

As a part of the standard NCES technical review process, each tabulation is reviewed by at least two members of the DRB who are responsible for ensuring that all NCES Statistical Standards for reporting are upheld. There are also no documented disclosures resulting from NCES publications.

### **RESTRICTED USE DATA**

Not all data files pass the Disclosure Review Board, and in most cases those that do have sensitive data items that have been deleted, top-coded, aggregated or altered in some way. While these alterations do not affect the analyses of most data users, there are times when a specific analysis requires some of the information that has been altered or deleted. Eventually, with the widespread availability of the Internet, external data users may be able to conduct analyses of restricted data without getting "close enough" to the micro-data file to risk disclosure. In the interim, NCES has adopted two approaches designed to increase access to restricted use data-- the Data Analysis Systems (DAS) and the Restricted Use Data Licensing System.

#### **Data Analysis Systems on the Web**

This system combines a data base system with a spreadsheet program to allow data users to request tabulations from restricted-use data files. The Data Analysis Systems are available as Windows software applications at the NCES Web site. With the DAS, users specify data tables by creating table parameter files that are uploaded to the DAS Web site. The request is processed and tables are placed in a pickup directory for the user to download. The tables include weighted data (usually expressed as percentages of students) and the corresponding standard errors that have been calculated taking into account the complex sampling procedures used in the data collection. Users may also request correlation matrices through the same process.

To avoid the risk of disclosure, the data produced by a DAS are categorical, no unweighted counts are provided, and estimates are only produced for cells with at least 30 respondents.

Data Analysis Systems are currently available for most of the postsecondary sample surveys. And while they are convenient for data users, they are limited to the set of data items programmed into the DAS. As a rule of thumb, a DAS is created to accompany each postsecondary data analysis report to facilitate further analysis of the items in the report. Thus, a new DAS is developed for each analysis report, and users are not able to use this tool to access additional restricted data.

### **Restricted Use Data Licensing System**

In 1989, the Associate Commissioner for the Statistical Methodology Division at NCES explored the feasibility of establishing a data licensing system at NCES with the Chief Statistician at OMB. A decision was made to proceed on a trial basis. NCES staff developed a protocol for the licensing system and worked with a lawyer from the Department's Office of the General Counsel to develop the legal documents required for the system.

The licensing system established a mechanism for NCES to provide external researchers access to individually identifiable NCES data covered under Federal statutes and regulations. By executing the license agreement and supporting materials, the external researcher assumes full responsibility for the confidentiality requirements established in the law that governs NCES confidential data. That is, the researcher agrees to assume the responsibilities and penalties for violations that apply to NCES staff. The license is granted under the authority of the Commissioner to permit individuals to examine respondent records. Under the license agreement, authorized users agree to unannounced inspection visits. Details of the licensing procedure are described in the [NCES Restricted-Use Data Procedures Manual](#), 1995.

### **NEW INITIATIVES**

The NCES DRB is currently undertaking two additional activities. First, the DRB is sponsoring the development of a software package that builds on the experiences of numerous analysts who have conducted disclosure analyses on NCES data. The goal is to produce a software package that will result in consistency and uniformity of disclosure risk analysis procedures across NCES data collections. Second, as a part of a larger project to revise and update the NCES 1992 Statistical Standards, the NCES members of the DRB are revising the NCES standard for maintaining confidentiality.

### **REFERENCES**

- U.S. Government, Computer Security Guidelines for Implementing the Privacy Act of 1974, Federal Information Processing Standard Publication (FIPSPUB 41). Washington DC: U.S. Government Printing Office.
- U.S. Government, The Privacy Act of 1974, 5, U.S. Code Section 522a, As Amended, Washington DC: U.S. Government Printing Office.
- U.S. Government, The Computer Security Act of 1987, Public Law 100-235, 100<sup>th</sup> Congress: Washington DC: U.S. Government Printing Office.
- U.S. Government, Hawkins-Stafford Elementary and Secondary School Improvements Amendments of 1988, Public Law 100--297, 100<sup>th</sup> Congress, Washington DC: U.S. Government Printing Office.
- U.S. Government, [NCES Statistical Standards](#), 1992, National Center for Education Statistics, Washington DC: U.S. Government Printing Office.
- U.S. Government, [The National Education Statistics Act of 1994](#), 20, U.S. Code 9001, Section 401, Washington DC: U.S. Government Printing Office.
- U.S. Government, [Restricted-Use Data Procedures Manual](#), 1995, National Center for Education Statistics, Washington, DC.

## **Discussion on Federal Review Boards**

Gene D. Therriault, New York State Department of Health

I would like to extend my congratulations to the three speakers for providing us with excellent summaries of the disclosure review boards operating in their agencies. I am struck by the many similarities in their approach to protecting the privacy and confidentiality of their data subjects. It appears that we are at a relatively infant stage in the development of these review boards and can learn from each other as we seek to improve and refine our mechanisms and procedures. Clearly the challenge that all of us face is to balance the legitimate need by the research community and the potential benefits of their work with the risks associated with the release of sensitive and potentially identifiable data.

My view of this subject has certainly been formed, in part, by my background as a user of this type of information in research. I understand the benefits of using data, analyzing it to draw conclusions and applying results to form policy and develop programs. But I also appreciate the need to conduct responsible research and the need to ensure that privacy concerns are addressed. I have also been influenced by my position as Executive Secretary of the Data Protection Review Board at the New York State Department of Health. This board reviews applications for access to hospital discharge data that contains potentially identifiable data. While our work is not completely analogous to the work of these federal boards, it is clear that the challenges we face in striking the appropriate balances between benefit and risk are very similar.

There are several common themes that I have identified. All three federal agencies have been able to establish an agency wide process for consistent review and evaluation of developed data products. Adherence by various programs within an agency to a single review process is a goal that all of us need to strive to attain. In many organizations, including my own, multiple review processes, some grounded in legislation, others in regulations and others in procedures, exist in a single agency resulting in various approaches to release policies and various views of the need to address privacy and confidentiality issues.

I also note that all three review boards are comprised entirely of federal employees. Understanding that there may be requirements that dictate such a membership composition, I would nevertheless raise the issue of developing mechanisms to include, either formally or in an advisory position, users of the data in the review process. Broadening the input to include non federal users may result in a more complete review and a better decision making process. In fact, the Data Protection Review Board of the New York State Department of Health has a membership composition that is almost totally comprised of non governmental representatives.

Another common theme of all three presenters was the need to be consistent in their review and consistent in their decisions. Previous applications that are similar in nature need to be referred to in order to insure that release of data is grounded in a common understanding of the responsibilities of the review board which do not vary from application to application or vary markedly over time.

Dr. Hoy from the Bureau of the Census emphasized two important points that are worth repeating. We who are charged with reviewing sensitive material and determining whether or not the release of the data could jeopardize privacy or confidentiality need a self awareness of the potential for inadvertent disclosure and we need a certain level of subject matter expertise. Without some background and appreciation of these two topics the task becomes impossible to perform.

Dr. Zarate from the National Center for Health Statistics mentioned the impact that HIPAA will have in the public health research community. These new and emerging rules demand a consistent review process for the release of sensitive health related information. These federal agencies can serve as models for those organizations that fall under the requirements of HIPAA.

Finally, Dr. Stamas from the Bureau of Labor Statistics summarized the common thread that ran through all the presentations when he described the role of review boards. What is the job description for these review boards? We need to ease the tension between data access and confidentiality. This states our challenge in words everyone can understand and appreciate. It may be a simple statement, but it is not a simple task.

*(compiled April 5, 2002)*