

THE MAINTENANCE OF CONFIDENTIALITY OF HUMAN RESEARCH SUBJECT DATA

1. In surveys, the truthfulness of the data collected is the basis for the credibility of the survey. Any lapse in a survey which incurs a risk on the participants as a class or an individuals presents a breach of trust between scientists and the respondent. Such a breach of trust, in addition to raising the certainty of cancellation of the College backing of the project, raises the possibility of legal action against the grantee and the granting agency. Even more seriously, it is conceivable that a single lapse may be all that is needed to break the bond of trust between the scientists and respondents so that future surveys may become either meaningless and/or impossible to conduct by all researchers.
2. Investigators must follow the safeguards specified in Public Law 91-513 (Privacy Act) to ensure the confidentiality and security of all information obtained from human subjects. Questionnaires, inventories, interviews, schedules, and other data-gathering procedures must be carefully designed to ensure that only information relevant to the project will be obtained.
3. Forms before survey use should be kept secure, locked, and preferably serialized (or otherwise identifiable) by number in order to prevent compromise of the survey by release of the information thereby prejudicing the results, or the opportunity for counterfeiting the form.
4. Respondent forms and reference lists must be treated as confidential, and should be coded and kept in locked, secure files (i.e., a file or cabinet with a combination lock.) Access to forms should be limited to authorized persons, and a sign-out, sign-in procedure should be instituted to monitor access.
5. If the information is to be computer filed:
 - a. The shipment, delivery and transfer of all data, printouts and files between offices and institutions may require careful controls.
 - b. Distinct separation of the data from identifiable individuals must be maintained. In no case should the names of the participants, or the identifying code be computerized. If identifiable numerical identifiers are necessary for file editing, they should be deleted as soon as the editing is complete.
 - c. Upon coding, and computer filing, the original forms must be destroyed by shredding or witnessed burning.
 - d. Access to the computer sorted data should be on a limited basis, again by authorized and identified persons only.
 - e. Since complete data raises the possibility that some individuals might be identified as case studies if a sufficiently detailed question is asked of the data, particular care must be maintained to lump the information to prevent such identification. Responses from uniquely identifiable individuals, groups, or companies must never be filed in such a manner that the information can be identified by source.
 - f. Data must be embedded in computers in a long composite field; i.e., long strings of densely coded digits.
8. If analysis is conducted on a subsample of the population, special care must be taken to be sure that the small class size does not lead to unintentional disclosure.

THE MAINTENANCE OF CONFIDENTIALITY – 2

9. Reporting of the results of surveys, research, case histories, and facts resulting from studies places a special burden on the investigator whether the report is in the form of a research paper, news release, or a newspaper story. The following factors must be impressed on all conductors of the study.

a. There is no statute of limitations on the confidentiality of the subject. researchers must agree to refrain from identifying subjects thereafter.

b. Coincidental identification of a subject may (and probably will) occur with one any combination of the factors listed below:

- 1) The study involved a small sample size.
- 2) The total population from which the sample was drawn is small.
- 3) The general characteristics of the population aggregate are stated directly, indirectly, through the employment of the subject. Size of the town community character (e.g., industrial, agricultural center, suburban education community, etc.), and general location may be sufficient identify the town.
- 4) Any confluence of characteristics of family structure (size, sex distribution of children, ages, marital evolution), details of personal physical characteristics, or expressions of individuality which together would permit statistical identification.

10. The usual case history contains all of these factors and must not be released without being thoroughly disguised. One method of disguising such case histories is to reconstruct a case history by using "reversal" of characteristics (fat for thin, etc.). Such reversal should be analyzed against other similar cases to ensure that accidental reassignment does not occur. If this method is used, it is essential to be certain that the data not contain a "typical" case history based on the average characteristics. Again, it is essential to be certain the published study does not contain a perfect "average" individual.