

## National Institute on Alcohol Abuse and Alcoholism

### Limited Access Data Policy

#### I. Introduction

#### II. Dataset Requests

#### III. Procedures for Protection of Privacy for Limited Access Datasets

##### A. Institute Review and Approval of Limited Access Preparation

##### B. Guidelines for Limited Access Preparation

---

In general, the following study will be included under this policy:

Data based on the National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III) conducted by the Epidemiology and Biometry Branch (EBB) and sponsored by the National Institute on Alcohol Abuse and Alcoholism (NIAAA).

#### **I. Introduction**

The NESARC-III contains data from well-characterized population samples. As such, these data constitute an important scientific resource. It is the view of the NIAAA that their full value can only be realized if they are made available, under appropriate terms and conditions, in a timely manner to the largest possible number of qualified investigators.

Datasets distributed under this policy include “limited access data”, i.e., records without personal identifiers and other variables that might enable individual participants to be identified, such as birth date.

Limited access datasets are available for the NESARC-III. For the purpose of this policy, data is defined as phenotypic information only collected from study participants.

#### **II. Dataset Requests**

##### **Responsibilities of Investigators Seeking Access to Study Data**

To ensure that the confidentiality and security of study participants are protected, all investigators seeking access to data from the NESARC-III that are in the possession of

the Institute must execute and submit their requests through the standard Data Use Agreement (DUA). Investigators seeking access to study data must also submit an exemption letter or an approval letter from their Institutional Review Board (IRB). An expedited review from the IRB is also acceptable.

Investigators should recognize that they are bound by the conditions of the NESARC-III DUA. Failure to comply with it could result in denial of further access to NIAAA datasets. Moreover, violation of the confidentiality requirements in the DUA may lead to legal action against the recipients of the data by the U.S. Government

### **III. Procedures for Protection of Confidentiality for Limited Access Datasets**

#### **A. Institute Review and Approval of Limited Access Preparation**

The NIAAA requires that epidemiologic data be provided in a manner that protects the confidentiality and security of study participants. The Institute prepares appropriate documentation of the steps taken to protect their confidentiality in preparing a limited access dataset. A summary of all proposed modifications and deletions to be made to a dataset in preparing it for limited access was approved by the Chief, Epidemiology and Biometry Branch, NIAAA.

#### **B. Guidelines for Limited Access Preparation**

The following guidelines provided a framework for deletions (and modifications as specified) regarding preparation of the limited access dataset:

1. Names.
2. All geographic subdivisions except for region and urban/rural.
3. All elements of dates (except year) for dates directly related to an individual, including birth date; and all ages over 89, such ages and elements are aggregated into a single category of age 90 or older.
4. Telephone number.
5. Facsimile numbers.
6. Electronic mail addresses.
7. Social Security numbers.
8. Medical record numbers.
9. Health plan beneficiary numbers.
10. Account numbers.
11. Certificate/license numbers.
12. Vehicle identifiers and serial numbers, including license plate numbers.
13. Device identifiers and serial numbers.
14. Web universal resource locators (URLs).
15. Internet protocol (IP) address numbers.

16. Biometric identifiers, including fingerprints and voiceprints.
17. Full-face photographic images and any comparable images.
18. New identification numbers must replace original New identification numbers.
19. Interviewer identification numbers must be recoded or deleted.
20. All continuous measures of income (personal/family/household) related to the total universe should be topcoded to include at least  $\frac{1}{2}$  of 1 percent of all cases.
21. Race and ethnicity will be collapsed into five race-ethnic groups: white, not of Hispanic origin; black, not of Hispanic origin; Hispanic/Latino, Asian or Pacific Islander; and Native American or Alaskan Native.
22. Other variables may lead to identification of individuals. All such known variables were recoded, deleted, or masked.