# Data and Statistics Dissemination Policy
Updated on 5/12/2020

## Background

The *All of Us* Research Program holds a rich dataset that will eventually grow to hold health, lifestyle, genetic, and environmental information from one million or more diverse participants living in the United States. As such, privacy and security are of utmost importance to protect participants from re-identification and from the misuse of their information. While *All of Us* has technical, legal, and policy safeguards in place to protect participants, authorized data users conducting projects within the Researcher Workbench play a key role in ensuring that participant data is protected both within the Workbench and also when they disseminate the results of their research.

## Rule and Analysis

Authorized data users who access the registered and controlled tiers through the Researcher Workbench will be able to view individual-level data about participants, including demographic information; electronic health record information; and survey information. *All of Us* takes a number of steps to prevent the misuse of this data, including: 1) permitting analysis of data only within its secure, cloud-based platform; 2) publicly posting the research purpose of each project and information about data users; 3) requiring potential users to undergo identity verification and responsible conduct of research training; and 4) requiring all data users to sign the *All of Us* Data User Code of Conduct (DUCC).

Among other requirements, the terms of the DUCC prohibit data users from attempting to re-identify participants or their relatives; however, to more fully protect participants from re-identification, it is important that authorized data users be careful when distributing the results of their work, whether through publication or otherwise, to prevent others from using this information to re-identify *All of Us* participants.

For this reason, the DUCC also stipulates the authorized data users will:

- **NOT** take screenshots or attempt in any way to copy, download, or otherwise remove any participant-level data from the *All of Us* Researcher Workbench.
  - **NOT** publish or otherwise distribute any participant-level data from the *All of Us* Research Program database.
  - **NOT** publish or otherwise distribute any data or aggregate statistics corresponding to fewer than 20 participants unless expressly permitted under the terms of the *All of Us* Data and Statistics Dissemination Policy.

Thus, the purpose of this policy is to outline the circumstances under which authorized data users may publish or distribute data or aggregate statistics that correspond to fewer than 20 participants. The goal of the policy is to not only prevent data users from directly reporting that a value corresponds to a specific, small number of participants (n <20) but also to prevent the use of triangulation to deduce this information using other available information within a report or publication.

## Policy Statement

Under the *All of Us* Data and Statistics Dissemination Policy:

1. No participant count of 1 to 20 can be published or distributed directly (a count of 0 is permitted); and

2. No data or statistics can be reported that allow a participant count of 1 to 20 to be derived from other reported cells or information, including in text, tables, or figures. This includes the use of percentages or other mathematical formulas that in combination would allow an individual to deduce a participant count of less than 20.

This policy permits data users who wish to report data or aggregate statistics that correspond to fewer than 20 participants to obscure these values using scientifically accepted strategies, including collapsing data across cells, coarsening data, or cell suppression. More information about acceptable strategies and how to employ them are available via the User Support Hub in the Researcher Workbench.

If data users have a compelling reason to justify directly publishing or disseminating data or aggregate statistics that correspond to a participant count of less than 20, they may submit a request to the program for an exception. Exceptions will be rare.