

## **APPENDIX F**

### **DMH Aggregate Data Release Standards**

#### **A. Purpose and Applicability**

As a Covered Entity and as a government agency, , that collects, stores, and reports data DMH has a responsibility to protect the privacy of individuals. At the same time, DMH is obligated to make a broad range of data publicly available in order to inform policy, effectively administer DMH services, and to comply with requests from other state agencies, commissions, boards, legislators, and other requestors. DMH complies with all applicable state and federal privacy laws. DMH utilizes data suppression standards in order to ensure that the confidentiality of data subjects is not inadvertently compromised in publicly available documents. DMH will apply these standards if the data, alone or in combination with other data, can be used to deduce an individual's identity. Examples of such data elements include, but are not limited to, geographic location, age, sex, diagnosis and procedures, admission or discharge dates, or date of death.

In order to protect the identity and privacy of data subjects and to avoid the risk of identification of individuals in small population groups, information reported in records must be reviewed and suppressed/redacted in accordance with the standards set forth below. This practice is commonly referred to as 'cell size suppression.'

These standards apply to data elements and any confidential information whether directly or indirectly displayed in records.<sup>1</sup> Confidential information includes any individually identifiable information<sup>2</sup> about a current, former, or deceased DMH applicant or client. For purposes of these standards confidential information also includes individually identifiable information about an individual contained in investigation, licensing, or other records maintained by DMH. Aggregate data<sup>3</sup> may also be considered confidential information unless it has been properly de-identified.<sup>4</sup>

#### **B. Standards**

The data suppression standards specify that no non-zero cells less than 11 and related complementary data fields may be displayed. Note that these standards are consistent

---

<sup>1</sup> Where applicable, "confidential information" includes individually identifiable information about a person who is not an applicant or client but who is included in an applicant's or client's household.

<sup>2</sup> Individually identifiable information includes any information readily associated with a particular individual that constitutes Personal Data, Protected Health Information, Patient Identifying Information, and/or Personal Information, as each of those terms are defined in the Privacy Handbook Glossary.

<sup>3</sup> See Privacy Handbook, Chapter 6, Section VIII. B. for definition of "aggregate data."

<sup>4</sup> See Privacy Handbook, Chapter 6, Section VIII. A. regarding de-identified health information.

with CMS current cell size suppression policy<sup>5</sup> and apply to individuals and related findings, observations, descriptions, and data elements that may result in the identification of an individual.

The reviewer should first determine whether small non-zero numbers less than 11 are included. If so, the reviewer needs to determine whether the data describes or reflects individuals (e.g. dates of admission, discharge, diagnosis, treatment, service or other information associated with unique individuals) as opposed to other metrics or indicators (e.g. mean patient days, average utilization, etc.) that are not representative of a unique individual.

Further, consider whether the document includes detailed descriptions of the data subjects and/or includes related criteria such as the complexity, incidence or prevalence of conditions (especially rare or unique conditions), and the granularity of data classification such as age bands, gender, geographic indicators, etc. Granularity and specificity of conditions or inclusion criteria increase the likelihood of re-identification and compromise confidentiality.

### **C. The Final Response to the Requestor:**

Consideration should be given to including the following language in the final response to the requestor:

“Please note that cells containing non-zero [*requested data, such as “discharge counts”*] under eleven (11) were redacted pursuant to the Health Insurance Portability and Accountability Act of 1996 (HIPAA) de-identification safe harbor, and in accordance with DMH aggregate data release standards, to avoid possible re-identification. (See 45 CFR §164.514(b).)”<sup>6</sup>

OR

“Non-zero numeric references less than 11 and related complimentary data fields have been masked (#) to protect confidentiality.”

---

<sup>5</sup> CMS current cell size suppression policy states that no cell (e.g. admissions, discharges, patients, services) containing a value of 1 to 10 may be displayed. A value of *zero* does not violate the minimum cell size policy. Also, percentages or other mathematical formulas may not be used or displayed if they result in the display or enable calculation of a cell of 1 to 10.

<sup>6</sup> Responses to public record requests must include this language.

## APPENDIX F

### Massachusetts ZIP Code Information

The HIPAA privacy rule allows de-identified data to contain a 3-digit ZIP Code field, as long as the population in that area is greater than 20,000. Massachusetts as of February, 2023 has 18 3-digit ZIP tracts with a population over the 20,000 mark.

Zip Code	Population
010XX	472,369
011XX	171,621
012XX	128,526
013XX	83,428
014XX	223,347
015XX	396,939
016XX	211,490
017XX	409,557
018XX	774,728
019XX	525,290
020XX	425,771
021XX	1,406,829
022XX	30,686
023XX	472,465
024XX	435,690
025XX	126,622
026XX	183,704
027XX	554,672

**For a listing of the Massachusetts towns by ZIP code see:**

<https://simplemaps.com/us-zips/MA/>

**For Census information by ZIP Code see:** the data dissemination tool feature of the U.S. Census Bureau Web page: <https://data.census.gov/>.