Dear Dr. Micky Tripathi,

We applaud your efforts to improve clinical data collection through the United States Core Data for Interoperability version 3 (USCDI v3). While primarily intended to improve collection of clinical data through electronic medical records (EMRs), USCDI standards improve public health surveillance and allow CDC and its public health partners to better understand the health of communities. The current version of USCDI includes standard values for Sex (Assigned at Birth), Sexual Orientation, and Gender Identity. USCDI v3 will allow for better classification for each of these values and provides an opportunity to align values with clinical, research, and surveillance standards.

We write to provide comments on the values for each demographic field, as well as to share and promote standard questions that may be provided to clinical entities and EMRs to further align clinical data collection across the country with public health surveillance and research data collection. We also would call your attention to a recently released report from the National Academies of Sciences, Engineering, and Medicine (NASEM), which reflects standardized questions and answers to measure sex, sexual orientation, and gender identity and improve data quality[[1]](#footnote-1), as well as a proposal by the gender harmony project of the Health Level Seven International (HL7) vocabulary working group [[2]](#footnote-2).

With regards to Sex (Assigned at Birth): Current values include “Female (F),” “Male (M),” and “Unknown (UNK).” These values align with the NASEM report and appropriately capture necessary information for public health surveillance of Sex (Assigned at Birth). Similar to the NASEM report, we would recommend that this demographic variable be linked with Gender Identity and asked in a two-step, paired fashion (e.g., Q1: What sex were you assigned at birth, on your original birth certificate? Q2: What is your current gender?). By combining Sex (Assigned at Birth) with Gender Identity, you can improve public health surveillance for many diseases and outcomes of interest (including, but not limited to HIV, STIs, viral hepatitis, suicidal ideation, interpersonal violence, firearm-related violence, and substance use/overdose), allowing for better calculation of rates and impact in cisgender and transgender communities.

With regards to Gender Identity: Current values include “Male,” “Female,” “Female-to-Male (FTM)/Transgender Male/Trans Man,” “Male-to-Female (MTF)/Transgender Female/Trans Woman,” “Genderqueer, neither exclusively male nor female,” “Additional gender category or other, please specify (OTH),” and “Choose not to disclose (ASKU).” These values provide more categories than are recommended by the NASEM report, although all values of USCDI v3 can appropriately match to a NASEM-recommended value while allowing for discrete values to enhance clinical data collection. We would recommend inclusion of “Non-Binary” as a value to reflect standard public health surveillance values, and would suggest inclusion of this value with Genderqueer (i.e., “Genderqueer or Non-Binary, neither exclusively male nor female”). Additionally, we agree with the NASEM report recommendation to include “Two-Spirit” as a value for Gender Identity for those who identify as American Indian/Alaska Native, recognizing the need to be inclusive of all communities.

At CDC, we have discussed the HL7 gender harmony project’s recommendations, which propose a more limited set of values for Gender Identity: “female,” “male,” “non-binary,” and “unknown.” We note that these values are only able to capture accurate data on Gender Identity when paired with Sex (Assigned at Birth) in a two-part question, as described above. While there are benefits to a more limited set of values for Gender Identity (including the potential for less data ambiguity and less stigma associated with the questions), there are also limitations. Many transgender individuals exclusively identify as “transgender male” or “transgender female,” while others may exclusively identify as “male” or “female.” A limited set of values may result in more transgender individuals selecting “unknown” and limit the accuracy of the data for public health surveillance (as well as for clinical care). We stress that more important than distinguishing between the use of two (male, female) or four (male, female, transgender male, transgender female) values for Gender Identity is the use of a two-part question paired with Sex (Assigned at Birth).

With regards to Sexual Orientation: Current values include “Lesbian, gay or homosexual,” “Straight or heterosexual,” “Bisexual,” “Something else, please describe (OTH),” “Don’t know (UNK),” and “Choose not to disclose (ASKU).” While these values correspond with the NASEM report-recommended values, we would advocate for the inclusion of two additional values to align with public health surveillance, including “Asexual” and “Questioning.” While these values may be captured in the current value of “Something else, please describe (OTH),” standardization of values will ensure better data collection and surveillance, especially of youth and adolescents. We agree with the NASEM report-recommended question for Sexual Orientation (i.e., “Which of the following best represents how you think of yourself?”).

In addition to the above recommendations, we provide 2 additional comments on USCDI v3 related to Sex (Assigned at Birth), Gender Identity, and Sexual Orientation:

* Sex (Assigned at Birth) is considered a fixed data point, reflected on an individual’s original birth certificate. Gender Identity and Sexual Orientation are data points that may change with time, depending on an individual’s understanding of their own gender identity or sexual orientation, or an individual’s comfort with disclosing these data in a clinical or administrative setting. We would recommend that USCDI guidance provide recommendations on how frequently to ask these questions and update data values for an individual.
* USCDI v3 does not include demographic values for non-binary sex identification (also known as intersex identity). We recommend that ONC consider inclusion of non-binary sex in future iterations of USCDI and agree with the NASEM report that intersex identity and intersex traits be identified with a stand-alone question (e.g., “Have you ever been diagnosed by a medical doctor or other health professional with an intersex condition or a difference of sex development (DSD) or were you born with (or developed naturally in puberty) genitals, reproductive organs, or chromosomal patterns that do not fit standard definitions of male or female?”).
* The HL7 Gender Harmony project has recommended that in addition to Sex (Assigned at Birth) and Gender Identity, an additional demographic detail be captured in a field titled “Sex for Clinical Use.” This is defined as “a sex classification element based on one or more clinical observations, such as an organ survey, hormone levels, and chromosomal analysis.” While not a current standard value of USCDI v3, we suggest that ONC consider adding this data element, possibly under the name “Sex for Clinical Use Note” included in the “Clinical Notes” data class.

We appreciate your willingness to consider our recommendations to improve USCDI v3 and strengthen clinical data collection and public health surveillance. We look forward to ongoing collaboration as we work to make health data more inclusive for all Americans, including those who identify as sexual and gender minorities.

1. National Academies of Sciences, Engineering, and Medicine. 2022. Measuring Sex, Gender Identity, and Sexual Orientation. Washington, DC: The National Academies Press. https://doi.org/10.17226/26424. [↑](#footnote-ref-1)
2. R.C. McClure et al. Gender harmony: improved standards to support affirmative care of gender-marginalized people through inclusive gender and sex representation. Journal of the American Medical Informatics Association, Volume 29, Issue 2, February 2022, Pages 354–363, https://doi.org/10.1093/jamia/ocab196 [↑](#footnote-ref-2)