



March 30, 2023

Rochelle P. Walensky, MD, MPH
Director
U.S. Centers for Disease Control and Prevention (CDC)
1600 Clifton Road
Atlanta, Georgia 30329-4027

Submitted electronically to:

REcode@cdc.gov

PHINVS@cdc.gov

pz1@cdc.gov

Re: CDC Race and Ethnicity 2022 Code System (CDCREC)

Dear CDC Director Walensky:

Health Level Seven (HL7) International welcomes the opportunity to offer feedback on the CDC Race and Ethnicity 2022 Code System (CDCREC) in accordance with a call for comments or requests related to concept additions, synonyms, or other Code System changes that will be considered by CDC in the maintenance process.

HL7 is the global authority on healthcare interoperability and a critical leader and driver in the standards arena. Our organization has more than 1,600 members from over 50 countries, including 500+ corporate members representing healthcare consumers, providers, government stakeholders, payers, pharmaceutical companies, vendors and consulting firms. HL7 standards, implementation guides (IGs) and related tools provide both a fundamental and innovative backbone to achieving our national goals of an interoperable and equitable health system.

The HL7 product family is robust end-to-end and is well supported by the healthcare industry, as reflected by our Accelerator community, long-standing Work Group structure, and expanding technical capabilities to support the HL7 development and implementation divisions. HL7 product and services management groups and related work groups manage, author and maintain HL7 standards such as HL7 Version 2 Messaging Standards (V2), Fast Healthcare for Interoperability Resources (FHIR) and the Clinical Document Architecture (CDA). Critical to all the standards is the use of terminology.

HL7's detailed comments on the CDC Race and Ethnicity 2022 Code System (CDCREC) below reflects the perspectives of our leadership, Co-Chairs of the HL7's Terminology Services Management Group, Terminology Authority, Terminology Infrastructure Work Group, Structured Documents Work Group, Orders and Observation Work Group and US Realm Steering Committee.

Overall, HL7 appreciates the effort to reflect in the CDCREC 2022 a detailed, up-to-date resource for collecting, coding, exchanging, analyzing and reporting race and ethnicity data in accordance with federal standards for classifying those data. However, HL7 recommends CDC create a new version of CDCREC (rather than create a new code system) and retain the concept identifiers from CDCREC 2020 that will be equivalent in CDCREC 2022. And, critically we urge CDC to engage with HL7 terminology subject matter experts to further discuss concerns before any updates are officially published.

Other key issues highlighted in HL7's comments are:

- Vocabulary management practices (HL7 and CDCREC);
- Data integrity;
- Critical implementation considerations and minimizing burden;
- Key concerns if CDCREC 2022 advances as outlined; and
- Other issues

Should you have any questions about our attached comments, please contact Charles Jaffe, MD, PhD, Chief Executive Officer of Health Level Seven International at cjaffe@hl7.org or 734-677-7777. We look forward to continuing this discussion with CDC and other relevant stakeholders.

Sincerely,



Charles Jaffe, MD, PhD
Chief Executive Officer
HL7 International



Andrew Truscott
Board of Directors, Chair
HL7 International

HL7 Response

Executive Summary and Overarching Recommendations

While HL7 recognizes the value of improving the CDCREC code system including removing codes and adding much needed granularity, care must be taken to adhere to best code system change practices in order to decrease the recursive, cascading and pervasive implementation requirements on the many stakeholders related to the current CDC proposal. The downstream burden to all involved must be carefully weighed and understood.

Collecting and sharing race and ethnicity details across health care systems has received a significant amount of attention in efforts to improve health disparities. Copious resources and countless hours have been expended for health information technology systems to align and use the current codes and concepts as identified in the “Race & Ethnicity – CDC”; version 1.2 Code System

There are more than 100 million patient records saved in currently active healthcare information technology (IT) systems using this current code system’s identifiers. Every search, display, or summary records software module --tens of thousands --will have to be modified to identify the currently used code system identifier for the codes and apply changes for the new codes. This will introduce severe delays in implementation, as well as high costs. Additionally, new code system implementations can lead to errors in statistical summaries combining old data and new data.

To ensure data integrity and minimize implementer burden and downstream analytics, HL7 recommends the following:

- CDC creates a *new version* of CDCREC (rather than create a new code system) and retains the concept identifiers from CDCREC 2020 that will be equivalent in CDCREC 2022.
- CDC uses distinct concepts for every Concept Code.
- CDC supports the current distinction between Race and Ethnicity as defined by OMB and does NOT add Ethnicity concepts to the “Other Race” category.
- CDC does not include any concepts as children of “Other Race”.
- CDC engages with HL7 terminology subject matter experts to further discuss concerns before any updates are officially published.

Implementation Considerations

Below, HL7 outlines the likely spectrum of implementation burden and issues arising from different CDC CDCREC decision paths.

High Implementation Burden

- CDC chooses to create a new code system - new code system identifiers are required (E.g. CodeSystem OID, URI, URL)
- CDC chooses to create a new code system version, and *not* align with the HL7 recommendations above.

Lower Implementation Burden

- CDC chooses to create a new code system *version*, and align with the recommendations, the health IT implementation burden will still be present, but will be less than the previous two options.

Race and Ethnicity Discrepancies Introduced

- The current OMB regulations clearly define race and ethnicity. Ethnicity is identified as being either “Hispanic or Latino” or “Not Hispanic or Latino.” The 2022 CDCREC introduces the concepts of Ethnicity under the “Other Race” category. This creates what appears to be, duplicate concepts in the race hierarchy and in the

ethnicity hierarchy, blurring the lines between race and ethnicity. This would likely contradict current OMB requirements. See example below.

Table 1: Data from Race And Ethnicity Spreadsheet from Tab "Race V Mapping 2022 To 2000"

Concept Code-2022	Hierarchical Code	Concept-2022
9000-1	R6	OTHER RACE
9020-9	R6.20	OTHER RACE - HISPANIC, LATINO, OR SPANISH
9021-7	R6.20.001	Hispanic
9022-5	R6.20.002	Mexican
9023-3	R6.20.002.001	Mexican American
9024-1	R6.20.002.002	Chicano(a)
9025-8	R6.20.002.003	La Raza

- The creation of detailed values in "Other Race" includes duplicated concepts. See example below.

Concept Code-2022	Hierarchical Code	Concept-2022
6000-4	R3	BLACK OR AFRICAN AMERICAN
6053-3	R3.07	CARIBBEAN
9000-1	R6	OTHER RACE
9008-4	R6.08	Caribbean

- Note: There are 19 newly identified races not identified under one of the current 5 minimum/high level race categories. Adding concepts under "Other race" and not including these in a hierarchy that aligns to one of the OMB 5 minimum race categories creates confusion. This appears to be contradictory to the current OMB requirements.

Vocabulary Management Practices: HL7 and CDCREC

Below, HL7 outlines our use of concept and Code Systems and how HL7 standards employ CDCREC.

HL7's use of concept and code systems:

- HL7 defines [Code Systems](#) as managed collections of concepts.
- Each Code System is uniquely identified via globally unique identifiers such as OIDs and canonical URIs.
- Code Systems contain Concepts that are also uniquely identified by a code.

Where CDCREC 2020 is currently used in Key US Realm HL7 Exchange Standards:

Interoperability requires both semantic (terminologies) and syntactic (exchange) standards. Neither works without the other. HL7 standards have been using CDCREC for more than 20 years.

C-CDA and HL7 FHIR US Core must align with [United States Core Data for Interoperability](#) (USCDI) which requires CDCREC be used for exchange of Race and Ethnicity. USCDI V1 and V2 require CDC Race and Ethnicity Code Set Version 1.0 (March 2000), and USCDI V3 requires CDC Race and Ethnicity Code Set Version 1.2 (July 2021)

C-CDA

- **US Realm Header (Patient):**

- This patient SHALL contain exactly one [1..1] raceCode, which SHALL be selected from ValueSet [Race Category Excluding Nulls 2.16.840.1.113883.3.2074.1.1.3](#) DYNAMIC (CONF:1198-5322). This patient MAY contain zero or more [0..*] sdct:raceCode, which SHALL be selected from ValueSet [Race Value Set 2.16.840.1.113883.1.11.14914](#) DYNAMIC (CONF:1198-7263).
Note: The sdct:raceCode is only used to record additional values when the patient has indicated multiple races or additional race detail beyond the five categories required for Meaningful Use Stage 2. The prefix sdct: SHALL be bound to the namespace urn:hl7-org:sdct. The use of the namespace provides a necessary extension to CDA R2 for the use of the additional raceCode elements. If sdct:raceCode is present, then the patient SHALL contain [1..1] raceCode (CONF:1198-31347).
- This patient SHALL contain exactly one [1..1] ethnicGroupCode, which SHALL be selected from ValueSet [Ethnicity 2.16.840.1.114222.4.11.837](#) DYNAMIC (CONF:1198-5323). This patient MAY contain zero or more [0..*] sdct:ethnicGroupCode, which SHALL be selected from ValueSet [Detailed Ethnicity 2.16.840.1.114222.4.11.877](#) DYNAMIC (CONF:1198-32901).

FHIR US Core

- **US Core Patient**

- US Core Race Extension
 - [OMB Race Categories](#)
 - [Detailed Race](#)
- US Core Ethnicity Extension
 - [OMB Ethnicity Categories](#)
 - [Detailed Ethnicity](#)

Stakeholders Impacted by the Proposed CDCREC Changes

A diverse and impactful spectrum of stakeholders will be affected by changes to CDCREC. HL7 depicts these parties below and important implementation considerations.

Stakeholders

Minimally, the following stakeholders would be impacted with a change to CDCREC as described in the Executive Summary.

- Health IT Vendors
- Data collection systems
- Data analytics vendors
- Clinical Trial Support Systems
- Researchers across all US (Academic and Industry)
- Quality Reporting Vendors
- Quality Measure Developers
- Standards Developers
- Public Health Federal, State and Local Jurisdictions
- Precision Medicine Vendors
- Precision Medicine Implementations of EHR Vendor products at Health Systems
- Birth and Death Records
- Pregnancy Risk Assessment Monitoring System (PRAMS)
- Federal Immunization Information System (IIS)

Critical implementation issues to be weighed include:

Implementation considerations

- Initial transition
- Maintenance/historical data
 - Versioning considerations (This pertains to whether a new Code System is created (not recommended), or a new version of the existing Code System is created)
 - Hierarchical properties *must* be maintained for *retired* concepts to support comparisons/analysis of historical data and roll-up.
 - The retired concepts would be inactive for future use, but not for data analysis, historical comparisons.
 - The new roll-up path can only be used when it is clear a concept from the new code system or code system version is present. Implementations may have to perform separate roll-ups.
- By assigning new concept identifiers for concepts in the 2020 version of CDCREC and providing equivalence mappings, it appears that the concept definitions are the same.
 - Any implementation or analytics will need to map all data collected prior to adoption of the 2022 code system to discover whether the concepts are equivalent.
 - The hierarchical “code” seems to have an implied meaning and should be labeled as “hierarchical property” and not *hierarchical code*. The value exchanged should be the Concept Code.

Other Specific Issues: Proposed CDCREC Updates

The following is a list of noted specific issues highlighted in deliberations by the HL7 Terminology Services Management Group, HL7 Terminology Authority and HL7 Terminology Infrastructure Work Group about the proposed CDCREC updates:

- CDC's proposed new content correctly assigns new concepts, new unique codes. However, a problematic issue is that identical concepts between the 2000 and 2022 releases have been assigned new identifiers.
 - If the codes *are* equivalent, new concept codes are *not* necessary and are highly problematic for implementations and downstream analysis.
 - If the (apparently) identical concepts are not unique, definitions of the 2 concepts, and map-types (narrower-than, broader than) should be provided.
 - Example: 2000 (2113-9) Irish, maps to the 2022 (8056-4) Irish
- A hierarchical change of an existing concept does not require a new code.
- Some of the mappings provided are clearly not equivalent (see 2073-5 Tobagoan and 2074-3 Trinidadian that are mapped to Trinidadian and Tobagoan 6066-5)
 - This should be characterized either as a “source narrower than target” mapping, or 2073-5 and 2074-3 should be retired.
- The Hierarchy Code should be characterized as a “Property” and not a code, to reduce potential confusion. The concept code is the value for exchange.
- There are several mistakes in the concept descriptions as well as in the column B heading. HL7 suggests CDC carefully review all concept descriptions. Examples:
 - Race2022 Tab, Code 3142-7 | Eastern Band of Cherokee Indians)
 - Race2022 Tab, column B heading - **Hierachical** Code - see above suggestion to change this to Hierarchical Property
- With the introduction of many specific race and ethnicity concepts, the single hierarchy structure isn't sufficient.
 - For example, under the "SOUTH ASIAN" concept, more specifically Sikh and Sindhi concepts. Per our understanding Sikh and Sindhi partly belong to "Asian Indian" and partly to "Pakistani" after India-Pakistan partition in 1947.

Of concern are the "Other Race" categories.

- Other Race Multiracial/Multiethnic:
 - Is this a combined race/ethnicity concept hiding in the race hierarchy?
 - How would one know whether to select or assign any of the Other Race values copied below?
 - What is the difference between *multiracial* and *multicultural* and *mixed* and *interracial*? See below.
- Health IT implementations support the collection of multiple race and ethnicity values - it is not necessary to define concepts that are in the proposed "Other Race-MULTIRACIAL/MULTIETHNIC" category.
- It is *not* recommended that data collection allow the selection of any of the proposed *specific* "Other" concepts - including the combining of some unknown race values into *mixed*.
- "Other Race" should be allowed.
- Additionally, selection of any of the R6, Other Race concepts means that none will roll-up to one of the minimum OMB race categories - violating the OMB requirements.

Table 2: Other Race Proposed Child Concepts

Concept Code-2022	Hierarchical Code	Concept-2022
9075-3	R6.21	OTHER RACE - MULTIRACIAL/MULTIETHNIC
9076-1	R6.21.001	Amerasian
9077-9	R6.21.002	Biracial
9078-7	R6.21.003	Eurasian
9079-5	R6.21.004	Interracial
9080-3	R6.21.005	Mixed
9081-1	R6.21.006	Mulatto
9082-9	R6.21.007	Multicultural
9083-7	R6.21.008	Multiracial