Proposed Criteria for Identifying Harmonization Priorities

Introduction

Data harmonization addresses differences in data that are an impediment to comparing, reusing, and sharing otherwise similar information. The purpose of harmonizing public health data is to improve the utility of the data and reduce the effort to collect and disseminate that data. The harmonization process identifies common concepts that are expressed using different terms and representations; specifies preferred terms, definitions, and representations; tracks the similarities; and where feasible, conforms to those specifications in collection, storage, and exchange instruments.

In order to maximize the resources and benefits of standards management and data harmonization, it is important to take a strategic approach for identifying high-priority harmonization areas across CDC. With input from CDC stakeholders, the Surveillance Data Platform (SDP) program developed proposed criteria for identifying harmonization priorities. The criteria are intended to provide strategic decision support as the Standards Management and Harmonization Steering Committee (SMaHSC), or another designated CDC-wide authoritative data governance body, identifies good candidates for data harmonization. The criteria offer a methodology to set initial priorities or a shortlist of harmonization opportunities that focuses on critical or common data needs across CDC.

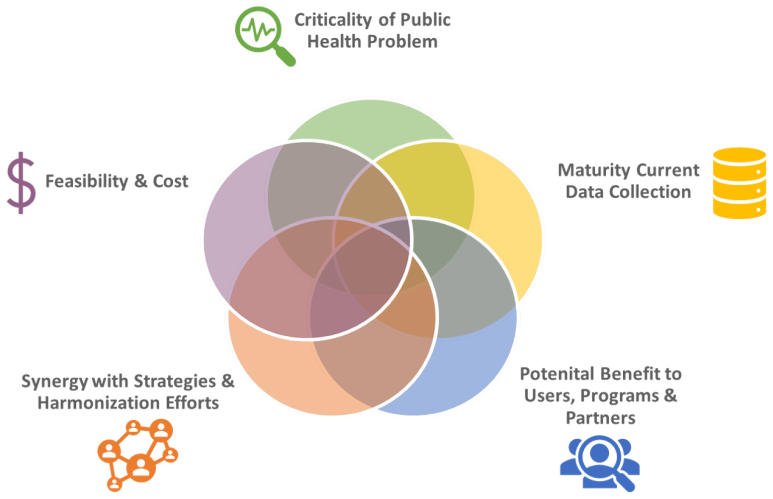
The criteria should consider critical or common data needs, likelihood of adoption in CDC and partner systems, considerations for collection and processing issues in public health surveillance and maximizing the benefit of standards management and data harmonization. Once SMaHSC identifies a public health topic as a harmonization priority, CDC leadership can make informed decisions to support future harmonization initiatives, such as sponsoring a temporary workgroup or Data Harmonization as a Service (DHaaS) task. In addition, once a topic is prioritized and selected for a data harmonization task, the same set of criteria can be used to guide the business case and set priorities *within* the execution of a harmonization task. For the execution of the harmonization task, it is helpful to document a general business case for candidate harmonization topics. The business case describes at a high-level the public health problem, the pain points caused by lack of data standardization and thus, the opportunity for data harmonization.

On this page:

- Introduction
- High-level Criteria
  - Criteria 1: Criticality of the Public Health Problem
  - Criteria 2: Maturity of Current Data Collection
  - Criteria 3: Potential Benefit to Users, Programs, & Partners
  - Criteria 4: Synergy with Strategies and Harmonization Efforts
  - Criteria 5: Feasibility & Cost
  - Summary of Proposed Criteria
- Notional Priority Levels
- Example Candidate Harmonization Topics

High-level Criteria

Five high-level criteria categories are proposed for prioritizing harmonization topics and described below.



Criteria 1: Criticality of the Public Health Problem

In this category, we consider how important or urgent the public health problem is and how much data improvements will improve the response to the problem. A harmonization priority topic may be driven by critical public health conditions, for example, childhood obesity, opioid use disorder or Zika-related birth defects. A current ongoing epidemic may be an urgent public health problem that requires immediate harmonization to support data collection and response. The scale of the health problem, which can be measured by the number of people affected, the prevalence or incidence of the health problem, is one way to characterize the importance of the topic. Additionally, the number of deaths caused by the health problem or quality of life (e.g., disability-adjusted life year) can indicate the severity of the public health problem. The type of public health surveillance information like demographics, travel history or health behaviors, can also be considered. If the data are commonly collected and cross-cutting centers, public health programs, surveillance systems and/or surveys, then harmonization efforts may provide benefit to a large area of public health areas. Another aspect of public health criticality is the urgency of the data need for this topic. For instance, if currently no data exist or are unusable for this topic, then this obstructs essential public health surveillance and response; harmonization efforts can be prioritized to improve public health.

Some questions are outlined for this criterion to assess the priority of a candidate harmonization topic.

Public health importance and urgency

- What is the level of urgency for harmonization in this area? Is it a pressing need for an epidemic or response?
- What is the scale of the health problem (e.g., prevalence, incidence, severity)?

Essential to routine surveillance

- Is it a topic commonly collected and essential to for routine surveillance?
- Are any data currently collected for this topic or are collected data usable for analysis and interventions?

Cross-cutting topic

- Is the topic of relevance to multiple Centers, public health programs, surveillance systems and/or surveys, health conditions?

Criteria 2: Maturity of Current Data Collection

In this category, we consider how entrenched is the existing data collection that serves the health problem is, or whether this is a burgeoning data collection need. Is there much data already collected? Are data collection instruments and systems already deployed versus being developed now? If the candidate harmonization topic will be used part to of new developing system, the use of harmonized specifications can be incorporated to ensure harmonized data collection at the start. Harmonization topics that affect many long-lasting surveillance systems may require mapping of existing data elements to harmonization specifications. Harmonization topics specific to retiring legacy systems may not be a priority.

We also consider if there are existing standards or specifications and if they are deployed in many systems or ignored in favor of many different data descriptions and representation. Data specifications describe the meaning, representation, rules, and guidelines for data for when it is collected, stored, exchanged, or processed. Some examples of data specifications are data elements, value sets, mapping guides, survey forms and questions. A harmonization topic that does not have certain specifications, such may be a good candidate for harmonization to address this gap. Even with existing specifications, the maturity of specification can be considered for prioritization. For instance, HL7 indicates different levels of stability and implementation readiness of a specification: draft, trail use, normative, informative and deprecated. Specifications under development or in draft stage may evolve versus normative specifications that are widely used and stable.

Some questions are outlined for this criterion to assess the priority of a candidate harmonization topic.

Maturity of involved systems and collection instruments involved

- What is the maturity (e.g., under development, new, existing, retiring, etc.) of the systems and data collection instruments involved with this topic? Are systems already in place or new systems under development?
- Are data collection instruments and systems already deployed versus being developed now? Will a new system or data collection instrument incorporate this topic?
- Are legacy or soon-to-be retired systems the primary systems involved?

Maturity of existing and developing data specifications

- Are existing specifications available for the harmonization focus? For example, do message mapping guides, HL7, LOINC, SNOMED standards exist for this harmonization topic?
- Are the specifications under development or adopted and used widely?

Criteria 3: Potential Benefit to Users, Programs, & Partners

Here we consider how many groups will change to incorporate the harmonized specification or the mapping in order to benefit from the harmonization efforts. There are many data sharers, data users, programs and partners involved in the public health surveillance. An estimate of the reduction of burden to external partners like state, territorial, local and tribal (STLT) health departments and public health laboratories that submit data to CDC, can be a consideration for a candidate harmonization topic. For instance, STLT health departments can potentially reduce data collection efforts for harmonized data elements across surveillance systems for the same health condition. The number of CDC centers and programs that will benefit is also a consideration; harmonization topics with broader touch points across the agency may have larger benefits.

Some questions are outlined for this criterion to assess the priority of a candidate harmonization topic.

Reduction of burden of external partners

- Does harmonization of the topic lessen STLT collection and reporting burden, e.g., fewer stand alone systems or processes?
- Does it decrease their informatics, IT coding and executable workload?

Number of CDC programs and surveillance systems

- How many estimated CDC programs and surveillance systems will benefit?
- Does it reduce the data wrangling burden of public health analysts, epidemiologists and data users?

Criteria 4: Synergy with Strategies and Harmonization Efforts

In this category, we consider whether strategies and harmonization efforts currently exist or if previous efforts target the harmonization topic. Ongoing harmonization tasks in the related area can share resources and outputs. Additionally, continuation of harmonization in related topics can help to expand specifications for particular domain area; for instance, if a previous effort has harmonized pregnancy status specifications, prenatal care specifications can be considered for prioritization to further harmonization in this maternal health domain. On the other hand, harmonization of one topic (e.g. pregnancy status) may be a prerequisite and thus, prioritized before working on another related topic (e.g., prenatal care).

Candidate harmonization topics should consider alignment with strategies like the CDC's Strategic Framework, Public Health Data Strategy or other HHS agencies. For example, if reducing opioid overdose mortality is a CDC strategic priority, harmonization of opioid use specifications can support this effort.

Some questions are outlined for this criterion to assess the priority of a candidate harmonization topic.

Combined effect with current or previous harmonization activities

- How much can this topic utilize findings and outputs from previous or current harmonization activities? Can collaboration occur with other current related efforts?

Alignment with strategies of HHS, CDC or partners

- Does the topic align with strategies of HHS, CDC or partners (e.g., CSTE)?
- Does topic support CDC's Public Health Data Strategy (2019)?

Criteria 5: Feasibility & Cost

Here, we consider how costly the harmonization of the topic may be and how complex or practical it is to undergo a harmonization initiative. The approximate time number of staff, and range of expertise (e.g., epidemiology subject matter experts, informaticists, etc.) are helpful aspects to determine the rough cost of harmonization for this

topic. Narrower topics or simple harmonization topics may require less time or resources to undertake compared to complex topics. For example, behavioral health topics with complex constructs and meanings may require a range of expertise and more effort to harmonize. Complex harmonization topics can also require higher level of engagement with a variety of stakeholders like data providers (STLTs), IT system designers, end users, external partners and agencies.

The feasibility of adoption of the harmonization specification is also important; a harmonized representation for administrative gender may be simple but owners may not be willing to adopt and change systems and collection instruments. The amount of time available to undergo harmonization of the topic also affects feasibility. Harmonization timelines may to shorter to align with program and user needs like response to emergency events, development of new surveillance systems or publication of external reports.

Some questions are outlined for this criterion to assess the priority of a candidate harmonization topic.

General cost

- What is the general cost of the harmonization task? For instance, how many staff and hours are needed?
- Are SMEs and other resources available to support this effort?
- What is rough timeline short or long (e.g., a week, a few months or a year) for the harmonization task? Is the target output for immediate value, short term or long term?

Complexity and Practically

- Is it a simple/"low-hanging fruit" effort versus a complex effort?
- How feasible is it to address this harmonization topic?
    - How complex is the harmonization topic concepts and terminology?
    - How likely is adoption of harmonization specification by system and survey owners?

Level of stakeholder engagement for harmonization

- How many stakeholder groups, like IT system owners and CSTE, need to be engaged during the actual harmonization process? How much involvement (e.g., review only or participation in clustering and analyzing existing specifications)?

Summary of Proposed Criteria

The table below summarizes the considerations of reach of the five criteria for identifying harmonization priorities.

Criteria Category		Summary of Considerations
	**Criticality of Public Health Problem**	Public health importance and urgency Essential to routine surveillance Cross-cutting topic
	**Maturity of Current Data Collection**	Maturity of involved systems and collection instruments involved Maturity of existing and developing data specifications
	**Potential Benefits to Users, Programs, & Partners**	Reduction of burden of external partners Number of CDC programs and surveillance systems
	**Synergy with Strategies and Harmonization Efforts**	Combined effected with current or previous harmonization activities Alignment with strategies of HHS, CDC or partners
	**Feasibility & Cost**	General cost Complexity and practically Level of stakeholder engagement for harmonization

Notional Priority Levels

Here we propose notional ranking and priority levels for SMaHSC to further pursue. Candidate harmonization topics can be rated using the five criteria described above. Each criterion can be rated and be weighted to distinguish importance of harmonization topics. This can help easily compare and rank topics. One option for assessment is to rate each criterion on a scale (e.g.,1 for not at all important to 5 for extremely important). Ratings can be summarized across criteria categories with higher scoring harmonization topics designated as high harmonization priorities for CDC.

Once the candidate harmonization topic is assessed by each criterion, it is helpful to designate an overall level of priority to assist strategic decisions. This can help initiate and sustain standards management and harmonization activities. For instance, the harmonization topic can be designated as high, medium or normal. High designates that harmonization is critical for this area and that active efforts should support harmonization for this topic like sponsoring a DHaaS workgroup. The importance and urgency of the high priority topic requires resource allocation to accomplish harmonization by a targeted time frame. Medium means harmonization is important for this topic and will benefit CDC. SMaHSC can encourage harmonization of the topic but the work only requires fixed resources. Normal means harmonization for this area is desirable but doesn't merit specifically committing a surge of time or resources.

Example Candidate Harmonization Topics

SMaHSC can gather candidate harmonization topics and apply the criteria to identify priorities. A list of example topics for consideration are below. This list is not exhaustive rather a sample of the range of candidate topics, including some topics that have been addressed by recent harmonization efforts or topics indicated as interests by stakeholders.

Demographics

- Race and ethnicity
- Sex and Gender

Epidemiological

- Blood pressure (for heart disease and stroke data)
- Pregnancy status
- Opioid use
- Travel history
- Dates: illness onset date, report date, specimen collection date
- Vaccination history, symptoms, food history, treatment, occupation/industry, exposure and contacts

Data specification request

- LOINC code request