Assistant Secretary for Technology Policy
Office of the National Coordinator for Health Information Technology
330 C St SW, Floor 7, Washington, DC 20201

Re: USCDI+ Cancer - Clinical Trials Matching Draft Dataset

Thank you for the opportunity to submit comments to the Assistant Secretary for Technology Policy/Office of the National Coordinator for Health Information Technology (ASTP/ONC) USCDI+ Cancer – Clinical Trials Matching draft dataset.

AMIA is the professional home for more than 5,600 informatics professionals, representing frontline clinicians, researchers, public health experts, and educators who bring meaning to data, manage information, and generate new knowledge across the research and healthcare enterprise. As the voice of the nation's biomedical and health informatics professionals, AMIA plays a leading role in advancing health and wellness by moving basic research findings from bench to bedside, and evaluating interventions, innovations and public policy across care settings and patient populations.

AMIA's Public Policy Principle, Health Information Technology Data Standards & Interoperability
AMIA's Public Policy Principle, Health Information Technology Data Standards & Interoperability[1], contains several key points that underline the vision and positions to develop optimal interoperability standards for USCDI+.

Health Information Technology Data Standards & Interoperability
Technical standards enable disparate systems to communicate and are prerequisites for our health information technology (HIT) ecosystems to interoperate. AMIA's Principles and Positions describe the desired characteristics of health IT standards for care and research, and articulate the importance of governance, testing, and multistakeholder standards development.

AMIA Believes:
Clinical, research and HIT systems must be able to exchange biomedical, clinical, and health data consistently and reliably using computable, and where appropriate, standardized formats while preserving the intended meaning and inter-relationships.

Access to and reliable use of digital healthcare data at scale requires that established, consistent, published, and openly available HIT standards be used to specify the formats

[1] AMIA Public Policy Principles 2024-2029, (pg. 18-20)

and characteristics (such as data types, ranges, etc.) for biomedical, clinical, and health data.

To ensure consistency and comparability of biomedical and clinical data, HIT standards must require coordinated and collaborative development through official announcements, open public comment periods, and published meeting notes.

Whenever possible, one canonical specification should be designated as the preferred representation for each biomedical, clinical, and health data standard required for defined use-cases related to optimizing health and healthcare.

Testing of HIT systems should test both conformance to, and interoperability of standards in real world environments to ensure data consistency and reliability across a diverse spectrum of implementations and use cases.

Based on these Principles, AMIA Supports:

1. The development and management of HIT standards as a public good, operated in a non-profit, non-proprietary basis, with low barriers to review, reference, or use.[2]

2. HIT standards that leverage existing information technology stacks, such as the Internet Protocol Suite[3,4] and the Trusted Exchange Framework and Common Agreement (TEFCA)[5] that expand the functionality of existing information systems and increase the use of HIT standards by disparate systems.

3. HIT standards that are modular and substitutable, having extensible, expandable boundaries for use and application, with specifications for automated access, use, and integration with relevant data.

4. HIT standards that are simple, parsimonious, and include documentation that is comprehensive, comprehensible, readily available, actionable, and timely.

5. HIT standards that are fit for purpose within a declared domain, and clearly recognized and identifiable as the preferred standard.[6,7]

---

[2] The Office of the National Coordinator for Health Information Technology. Federal Health IT Strategic Plan_2020_2025.pdf (accessed August 17, 2024)
[3] Also known as TCP/IP (https://.ietf.org/) (accessed August 17, 2024)
[4] TCP/IP Model - GeeksforGeeks (accessed August 17, 2024)
[5] Assistant Secretary for Technology Policy/Office of the National Coordinator for Health IT. Trusted Exchange Framework and Common Agreement (TEFCA) | HealthIT.gov (accessed August 17, 2024)
[6] This criterion implies being comprehensive within a declared domain of information, purpose and context, and generating verifiable content, preserving provenance, and computer interpretable.
[7] Han L, Liu J, Evans R, Song Y, Ma J. Factors Influencing the Adoption of Health Information Standards in Health Care Organizations: A Systematic Review Based on Best Fit Framework Synthesis. JMIR Med Inform. 2020 May 15;8(5):e17334. doi: 10.2196/17334. PMID: 32347800; PMCID: PMC7260665.

6. HIT standards that leverage prevailing security practices to protect and preserve data integrity, privacy and confidentiality.

7. Efforts to recognize and address stakeholder motivations, aims, activities, business models, and information needs in the specification of HIT standards to increase the value of their adoption by users and improve ease of implementation.

8. Standards development that incorporates implementation experience and feedback loops from real-world settings to better support an adoption pathway for HIT standards.

9. Interdisciplinary collaboration on potential standards for new modalities of biomedical data, use cases, and information technology that can evolve and mature through implementation experience before canonical specifications can be identified as the standard.

10. Interoperability testing, which tests both the sending of data using a specific standard(s) as well as receipt of data using such standard(s), and tests adherence to Postel's Principle.[8,9]

11. Adequate funding for the development, management and maintenance of HIT standards, and the SDOs that create them, due to the enormous positive impact on society HIT interoperability can have.

AMIA Recommendations and Considerations

The following sections and table are recommendations and considerations from AMIA on areas of need with potential remedies for gaps observed in the framing questions provided by ASTP.

Cancer Data Integration and Categorization

AMIA sees potential concerns about the integration of cancer prevention and treatment data, particularly with respect to documenting the distinction between what data is integrated and how in these two scenarios. Both an oncologist and an informaticist must be involved to assure that the meaning of each data element is properly understood and documented. We would also like to point out the ambiguity in some data elements, such as medications and laboratory results, and the need for clearer categorization. We understand that even partial data could be beneficial for clinical trial matching, and the

[8] Also known as Postel's Robustness Principle, stating: Be conservative in what you do, be liberal in what you accept from others (often reworded as "Be conservative in what you send, be liberal in what you accept"). Postel, Jon, ed. (January 1980). Transmission Control Protocol. IETF. RFC 761. Retrieved June, 2017.
[9] Assistant Secretary for Technology Policy/Office of the National Coordinator for Health IT. Interoperability Standards Platform (healthit.gov) (accessed August 17, 2024).

process could be streamlined to flag patients at risk, but AMIA agrees on the need for better data integration and categorization.

Improving Cancer Diagnosis Data Detail

AMIA also would like to stress the need for more detailed information in the USCDI+ – Clinical Trials Matching draft dataset, particularly regarding cancer diagnoses. We note the lack of specific information about comorbid conditions and the need to identify the primary site of cancer. AMIA wants to know more about the utility of certain data fields, such as ethnicity, and questioned the relevance. Overall, we want to understand why certain data fields were included and their potential utility.

Refining Data Elements for Oncology Trials

We agree with ASTP that the current list of 32 fields is a good starting point, it needs to be refined to better serve the needs of clinical trials. We see it necessary for more specific information about cancer treatments and concomitant medical conditions, as well as the context in which the data was collected. AMIA acknowledges that this should be a trial-and-error process and that some fields might not be specific to oncology.

Improving Patient Data Differentiation

The team discussed the need for more differentiation between patient's cancer status and other health conditions. They agreed that certain data elements, such as pregnancy status, could be improved by providing more specific information like the estimated due date. The team also discussed the challenges of populating fields with diagnosis information due to the complexity of patient conditions. They concluded that there is a need for clearer definitions and more specific data elements to avoid confusion.

Oncology Trial Criteria and Reporting

AMIA recognizes the importance of inclusion and exclusion criteria for oncology trials, emphasizing the need for diagnosis codes, histology, laterality, and the presence or absence of metastases. There is also significance to inclusion of performance status measures, often represented by ECOG or Karnofsky scores. AMIA would like confirmation of the presence of the two fields for clinical performance status data element. AMIA recommends that biomarkers, such as EGFR status, are crucial in precision oncology. We agree on the importance of pharmacogenomic data but note the challenge of reporting multiple mutated genes and suggest further exploration of how these elements can best be reported in the data.

Clinical Trial Data Elements Importance

AMIA agrees that data elements such as performance status, past clinical trials, and surgical history are crucial for eligibility determination and post-trial management. There is also a need for a unique identifier for each trial participant, which could be the FDA identifier. The concept of provenance and the importance of digital signatures or hashes for data integrity needs further exploration. AMIA recommends considering the possibility of having two different types of data collections for pre-screening and other phases of clinical trials, including final determination of eligibility.

Standardizing Clinical Trial Data Collection

AMIA recognizes the challenges of capturing and managing data for clinical trials, particularly in oncology. There is a necessity for a standardized approach to data collection, including medication allergies, intolerances, and other health factors. It is also important to capture data on the eligibility criteria for trials and the potential for using aggregated, de-identified real-world data to inform trial design. We emphasize the need for clarity on the purpose and source of the data, as well as the different types of information needed at various stages of the process. For example, does the data represent a single point in time, e.g. an office/clinic visit, hospital admission or treatment? Or has the data been compiled, such as a discharge summary?

Data Elements Recommendations

Table 1 addresses five data elements that would assist creating a more holistic capture of individual health information to optimize personalized treatment and clinical trial matching by developing a core dataset that improves the accuracy and efficiency of matching patient data with open clinical trial protocols.

1. Measurable Disease
2. Resectability
3. Stage
4. Stage Components
5. Biomarker

Sincerely,
Eileen Koski
Chair, AMIA Public Policy Committee

Table 1

Data Element	Description	Data Class	Well Structured?	Quality Concerns?	Missing from USCDI+?	Comments
Measurable Disease	Indication of whether or not the patient's disease is measurable per RECIST criteria.	Imaging	No	Yes	Yes	
Resectability	Indication of whether or not the patient's disease is resectable	Prior therapy	No	Yes	Yes	
Stage	Overall stage of patient's disease	Stage	No	Yes	Yes	
Stage Components	T, N, M stage components of overall AJCC stage	Stage	No	Yes	Yes	
Biomarker	Indication of whether or not the patient's tumor (or germline) harbor a particular molecular aberration	Molecular Pathology	No	Yes	Yes	
Behavior Code	Code for the behavior of the tumor being	Tumor	No	No	No	Could be useful in certain circumstances,

	reported using ICD-O-3.					particularly the central nervous system
Clinical Performance Status	A physician's assessment of the clinical performance of the patient, as measured by a rating or scale, considering disease and potential responses to therapy.	Health Status Assessments	Sometimes	Yes	No	Should be available in unstructured data but typically not structured in EMR routinely. Subjective and variable when a consideration for clinical trial matching
Clinical Performance Status Assessment Date	Clinically relevant time/time-period for the assessment.	Health Status Assessments	Sometimes	Yes	No	Same as above
Comorbid Condition Name	Medical or health condition that is concomitant or concurrent with the primary condition or disease under study.	Comorbid Conditions	Yes	No	No	
Current Address	Place where a person is located or may be contacted.	Patient Demographics/Information	Yes	No	No	

Current Clinical Status Date	Clinically relevant time/time-period for observation.	Problems	Yes	No	No	
Current Clinical Status Trend	How patient's given disease, condition, or ability is trending. EOM allowed values are: - Patient's condition improved - Patient's condition stable - Patient's condition worsened - Patient's condition undetermined - In full remission. - In partial remission - Distant metastasis present	Problems	No	No	No	
Date Medication Administered	A specific date/time or interval of time during which the administration took place (or did not take place,	Medications	Yes	No	No	

	when the 'notGiven' attribute is true).					
Date of Birth	Known or estimated year, month, and day of the patient's birth.	Patient Demographics/Information	Yes	No	No	
Date of Diagnosis	Date of first determination by a qualified professional of the presence of a problem or condition affecting a patient.	Problems	Yes	No	No	
Ethnicity	Patient's self-identification as Hispanic/ Latino or Non- Hispanic/ Non-Latino.	Patient Demographics/Information	Sometimes	Yes	No	
Gender Identity	A person's internal sense of being a man, woman, both, or neither.	Patient Demographics/Information	Yes	No	No	
Histology	The morphologic and behavioral characteristics of the cancer reported using ICD-O-3.	Tumor	Sometimes	No	No	

Laboratory Results: Date and Timestamps	Date and timestamps associated with the completion of laboratory results, that are meta data associated with laboratory results.	Laboratory	Yes	No	No	
Laterality	Code for the side of a paired organ, or the side of the body on which the reportable tumor originated. This applies to the primary site only. These codes are use in mCODE: - 51440002 Right and left (qualifier value) - 399488007 Midline (qualifier value) - 24028007 Right (qualifier value) - 7771000 Left (qualifier value)	Tumor	No	No	No	
Medication Class	A code that identifies the major functional	Medications	No	No	No	Not clear how frequently this is structured as

	or pathway classification to which a drug belongs.					a distinct element but can be normalized with relative ease
Medications	Pharmacologic agent used in the diagnosis, cure, mitigation, treatment, or prevention of disease.	Medications	Yes	No	No	
Metastasis Anatomic Site	The anatomic site of the patient's cancer condition that is the main reason for evaluation and treatment.	Tumor	Sometimes	No	No	
Personal Medical History Procedure Name	Activity performed for or on a patient as part of the provision of care that is not related to the primary cancer condition.	Personal Medical History	Sometimes	Yes	No	Can be difficult to establish for referral patients
Personal Medical History Procedure Performance Date	Time and/or date a procedure not related to the primary cancer condition is performed.	Personal Medical History	Sometimes	Yes	No	Same as above

	Examples include but are not limited to vaccine or medication administration times, surgery start time, and time ultrasound performed.					
Pregnancy Status	State or condition of being pregnant or intent to become pregnant. Examples include but are not limited to pregnant, not pregnant, and unknown.	Health Status Assessments	Sometimes	No	No	
Primary Site	The anatomic site of the patient's cancer condition that is the main reason for evaluation and treatment reported using ICD-O-3.	Tumor	Yes	No	No	
Race	An individual's response to the race question based upon self-identification.	Patient Demographics/Information	Sometimes	Yes	No	

Radiation Therapy Received Indicator	Indicator of whether or not the subject has received radiation therapy, including chemoradiation. Example values: Yes, No, Unknown	Radiation Therapy	Yes	Yes	No	Perceived integration difficulty between source clinical systems
Recurrence Anatomic Site	Anatomic site where a cancer has recurred (come back), usually after a period of time during which the cancer could not be detected.	Tumor	No	Yes	No	Typically not needed to know granular site of disease
Recurrence or Relapse Clinical Status	Recurrence is the return of a solid tumor cancer after a clinically disease-free interval (even after a previous relapse); this includes local or regional recurrence. The term relapse is used to describe the return of a leukemia,	Problems	Sometimes	No	No	Mentioned in unstructured notes; unsure how frequently it's actually structured

	lymphoma, or other hematopoietic malignancy that was not previously clinically apparent or symptomatic. Status may include: "active," "recurrence," "relapse," etc. EOM's allowed values: - Yes (Active Recurrence or Active Relapse) - No (Inactive Recurrence or Inactive Relapse)					
Result Reference Range	Upper and lower limit of quantitative test values expected for a designated population of individuals.Usage note: Reference range values may differ by patient characteristics, laboratory test manufacturer,	Laboratory	Yes	No	No	

	and laboratory test performer.					
Result Unit of Measure	Units of measurement for the lab test	Laboratory	Yes	No	No	
Sex Parameter for Clinical Use	Category based upon clinical observations typically associated with the designation of male and female. Usage note: There may be multiple instances of this data element for a single person, based on how the clinical observations (e.g., anatomic characteristics, recent hormone levels, or genetic testing) relate to or affect the clinical uses such as laboratory tests and results, diagnostic imaging, or preventive	Observations	Yes	No	No	

	screening measures. Context specific values should be associated with these clinical uses.					
Smoking Status	Assessment of a patient's smoking behaviors. Examples include but are not limited to pack-years and current use.	Health Status Assessments	No	Yes	No	
Tests	Analysis of specimens derived from humans which provide information for the diagnosis, prevention, treatment of disease, or assessment of health.	Laboratory	Yes	No	No	If this is exclusive to laboratory values, no concerns. If it extends to molecular tests, such as hematoxylin and eosin stains as indicated on the element page, frequently unstructured and have quality concerns. The

						linked data element is to vital signs, which are not a consideration for cancer trial matching.
Values/Results	Documented findings of a tested specimen including structured and unstructured components	Laboratory	Yes	No	No	Same as above