



Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2023 November 01; 117(3): 533–550. doi:10.1016/j.ijrobp.2023.05.033.

Operational Ontology for Oncology (O3) – A Professional Society Based, Multi-Stakeholder, Consensus Driven Informatics Standard Supporting Clinical and Research use of “Real -World” Data from Patients Treated for Cancer:

Operational Ontology for Radiation Oncology

CS Mayo¹, MU Feng², KK Brock³, R Kudner⁴, P Balter³, JC Buchsbaum⁵, A Caissie⁶, E Covington¹, EC Daugherty⁷, AL Dekker⁸, CD Fuller³, AL Hallstrom⁹, DS Hong¹⁰, JC Hong², SC Kamran¹¹, E Katsoulakis¹², J Kildea¹³, AV Krauze¹⁴, JJ Kruse¹⁵, T McNutt¹⁶, M Mierzwa¹, A Moreno³, JR Palta¹⁷, R Popple¹⁸, TG Purdie¹⁹, S Richardson²⁰, GC Sharp¹¹, S Shiraishi¹⁵, L Tarbox²¹, AM Venkatesan³, A Witzum², KE Woods¹⁰, J Yao¹, K Farahani⁵, S Aneja²², PE Gabriel²³, L Hadjiiski¹, D Ruan²⁴, JH Siewerdsen¹⁶, S Bratt²⁵, M Casagni²⁵, S Chen²⁵, J Christodouleas²³, A DiDonato²⁵, J Hayman¹, R Kapoor¹⁷, S Kravitz²⁵, S Sebastian²⁵, M Von Siebenthal²⁶, Y Xiao²³

⁽¹⁾University of Michigan

⁽²⁾University of California San Francisco

⁽³⁾The MD Anderson Cancer Center

⁽⁴⁾American Society of Radiation Oncology

⁽⁵⁾National Cancer Institute

⁽⁶⁾Dalhousie University

⁽⁷⁾University of Cincinnati Medical Center

⁽⁸⁾Maastricht Clinic

⁽⁹⁾Wellesley College

⁽¹⁰⁾University of Southern California

⁽¹¹⁾Massachusetts General Hospital

Charles Mayo, cmayo@med.umich.edu.

Research data are not available at this time. The O3 framework will be released into the public domain.

Author count is reflective of the extended level of expertise needed to inform and create the Operational Oncology for Oncology and all authors contributed significantly to the work.

[Conflict of Interest Statement for All Authors]

If not already noted, Conflict of Interest: None.

Disclaimer: This article represents the opinion of the authors. It does not represent the opinion or policy of the National Institutes of Health of the US Federal Government.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

- (12)James A Haley VA Medical Center
- (13)McGill University
- (14)National Institute of Health
- (15)Mayo Clinic
- (16)Johns Hopkins University
- (17)Virginia Commonwealth University
- (18)University of Alabama at Birmingham
- (19)Princess Margaret Cancer Center
- (20)Swedish Cancer Institute
- (21)University of Arkansas
- (22)Yale University
- (23)University of Pennsylvania
- (24)University of California Los Angeles
- (25)MITRE Corporation
- (26)Varian

Abstract

Purpose—The ongoing lack of data standardization severely undermines the potential for automated learning from the vast amount of information routinely archived in electronic health records (EHRs), Radiation Oncology Information Systems (ROIS), treatment planning systems (TPSs), and other cancer care and outcomes databases. The effort sought to create a standardized ontology for clinical data, social determinants of health (SDOH), and other radiation oncology concepts and interrelationships.

Methods and Materials—The American Association of Physicists in Medicine's (AAPM's) Big Data Science Committee (BDSC) was initiated July of 2019 to explore common ground from the stakeholders' collective experience of issues that typically compromise the formation of large inter- and intra- institutional databases from EHRs. The BDSC adopted an iterative, cyclical approach to engaging stakeholders beyond its membership to optimize the integration of diverse perspectives from the community.

Results—We developed the Operational Ontology for Oncology (O3) which identified 42 key elements, 359 attributes, 144 value sets, and 155 relationships ranked in relative importance of clinical significance, likelihood of availability in EHRs, or the ability to modify routine clinical processes to permit aggregation. Recommendations are provided for best use and development of the O3 to four constituencies: device manufacturers, centers of clinical care, researchers, and professional societies.

Conclusions—O3 is designed to extend and interoperate with existing global infrastructure and data science standards. The implementation of these recommendations will lower the barriers for

aggregation of information that could be used creating large, representative, findable, accessible, interoperable and reusable (FAIR) datasets supporting the scientific objectives of grant programs. The construction of comprehensive “real world” datasets and application of advanced analytic techniques, including artificial intelligence (AI), holds the potential to revolutionize patient management and improve outcomes by leveraging increased access to information derived from larger, more representative datasets.

Introduction

The construction of large, comprehensive “real world” data sets and application of advanced analytic techniques, including artificial intelligence (AI), holds the potential to revolutionize patient management and improve outcomes by leveraging increased access to information derived from larger, more representative data sets. Doing so will refine the ability of providers and researchers to identify prognostic and predictive factors, offer safe treatments, and improve relevant outcomes. Comprehensive data sets, aggregated at scale through structured integration into routine practice, could facilitate novel and meaningful observations, including meeting critical needs such as characterizing covariation of control and toxicity outcomes of anti-cancer treatments for geographic or social determinants of health (SDOH) increasing knowledge on rare cancers and permitting benchmarking of practice performance and technical radiotherapy (RT) plan quality.

However, the ongoing lack of data standardization severely undermines the potential for automated learning from the vast amount of information routinely archived in electronic health records (EHRs), Radiation Oncology Information Systems (ROIS), treatment planning systems (TPSs), and other cancer care and outcomes databases. Standardization is necessary for the scaled assemblage of artificial intelligence/analysis-ready data sets. Standardization will facilitate automating the accurate and specific extraction of data parameters from EHRs, support interoperable data exchange among multiple institutions and reduce “noise” in large-scale data sets. Only by incorporating data standardization into routine clinical practice, will the paradigm shift from “data mining” to “data farming”, with associated reduction in work needed to construct, populate, and maintain reliable databases that support clinical quality improvement and research activities. (1).

Specific domain expertise is crucial to implement large-scale data farming within a specialty, especially one as complex and technologically driven as radiation oncology. Clinicians and operators within a given specialty have the most precise understanding of which parameters are accepted as meaningful, how processes can be made feasible, and what information would be most helpful in support of quality and research efforts. Standards created by small or external groups lacking diversely represented domain expertise may omit factors crucial to clinical implementation within a specialty.(2) Development of methods validated through collective experience are material to the creation of data standards amenable to clinical and investigative use.

Prior work has demonstrated the value of an approach to creating consensus-driven information standards derived from the experience of a wide range of clinical users, professional societies, and relevant stakeholder groups. Several RT professional societies

played a critical role in the development of the American Association of Physicists in Medicine's (AAPM's) task group on Standardizing Nomenclatures in Radiation Oncology (TG-263) (3). Since its release in 2017, TG-263 standards have been adopted into clinical trials and vendor TPSs supporting inter-operable data exchange. The Global Harmonization Group recommended usage of TG-263 as part of its guidelines for organ at risk delineation for radiation therapy clinical trials. (4) Recognizing the value of RT nomenclature standardization in routine clinical practice, other professional societies, such as the Canadian Partnership for Quality Radiotherapy (CPQR), have promoted the adoption of TG-263. In 2021, CPQR released the "Guidance on the use of common nomenclature in Canadian radiation treatment programs," highlighting TG-263 as the foundation of future national RT "Big Data" efforts. CPQR aims to harmonize radiation oncology practice in Canada while facilitating large-scale data analysis of patient, diagnosis, and treatment-related factors affecting radiation oncology patient outcomes (5,6).

We report here on a stakeholder engagement method to transparently tap the combined expertise rooted in the activities and efforts of professional societies and knowledgeable parties to specify a standardized Operational Ontology for Oncology (O3), facilitating interoperable data exchange and supporting future development of large-scale AI and analysis-ready data sets. (Figure 1) The O3 development represents the single most collaborative effort in this domain to date. It has been enabled by the collaboration of the AAPM's Big Data Sub Committee (BDSC), with representatives from the American Society of Radiation Oncology (ASTRO), the Canadian Organization of Medical Physicists (COMP), the Canadian Association of Radiation Oncology (CARO), the European Society of Therapeutic Radiation Oncology (ESTRO) and clinical trials experts from NRG Oncology. Drawing from its multi-professional, multinational, and multi-institutional members' clinical, research, and informatics expertise and using a consensus-driven approach, the AAPM's BDSC identified key clinical elements, attributes, value sets, and relationships deemed necessary for standardization to support RT clinical practice and research applications. In recognition of its genesis with the professional societies cited above, the ontology was originally named the Operational Ontology for Radiation Oncology (OORO), even though the comprehensive ontology extends to all parts of cancer care. To reduce confusion in explaining the scope of the ontology, the 'R' was dropped, leaving the Operational Ontology for Oncology (O3).

A nomenclature identifies a set of standardized names to improve consistency, for example AAPM TG-263 identifies standardized names for structures segmented as part of treatment planning. A taxonomy identifies a hierarchy of categorizations, for example the Linnaeus system for subdividing categorizations of plant types.

An ontology is a philosophical representation of a body of knowledge that identifies the concepts and interrelationships. It provides the basis for a coordinated reasoning about that body of knowledge. (7) A challenge for clinicians applying the term ontology is that it tends to be abstract and difficult to provide practical examples. In this work we use the term "operational ontology" to clarify that all concepts, value sets, relationships, etc. are defined for practical application to facilitate use and improve consistency of clinical and research operations as well as linkages to other information coding systems. The operational

ontology combines concepts of nomenclature, and taxonomy with ontology developing operational standardizations. The operational ontology is presented in a form that end users can comprehend and contextualize in their own practice.

Note that this use of “operational ontology” is distinct from a “semantic-web ontology” that expresses concepts and relationships in a computer readable format that is not organized for end user development and consumption. For example, semantic-web ontologies generally do not provide standardized value sets, linkages to clinically used categorizations, coding systems, definitions, type definitions, etc. Formal semantic-web ontologies (i.e., “formal ontologies”) such as the OBO Foundry try to cover knowledge domains using formal definitions and be interoperable with coordinated use of terms.(8) Examples of semantic-web ontologies can be found at the Bioportal or OBO Foundry Web sites. (9) In practice, semantic-web ontologies leave operational gaps for clinical practice, are not developed in the context of multiple professional society efforts with peer review and public comment, nor are they developed with the express aim to make real world data interoperable. These gaps can be filled with a professional society based, operational ontology to improve consistency in creating interoperable, AI ready data sets needed to support clinical and research efforts. Development of an operational ontology for cancer care provides a strong basis subsequent development of formal semantic-web ontologies.

The stakeholder engagement method was developed alongside the creation of the O3. It aimed to support incremental engagement with a wide range of stakeholders allowing for downstream optimization by addressing several clinical and research objectives shared within professional societies to expand the codified sets of disease site-specific elements. Multiple additional stakeholders were engaged, including government agencies and vendors, underscoring the value of a collaborative approach to developing comprehensive and standardized ontologies. The Operational Ontology for Oncology has been endorsed by AAPM, ASTRO, CARO, COMP and ESTRO.

The initial version of O3, described herein, includes many disease-site independent concepts common for all cancer patients and a smaller set specific for prostate cancer. These include imaging, pathology, medical oncology, surgery as well as radiation treatment. The O3 methodology is currently being applied/adapted to disease site-specific concepts for head and neck cancers and breast cancer through the AAPM BDSC, with lung and remaining sites to follow. The clinically-oriented O3 lays the foundation for informatician-oriented formalized ontology tools following the principles established by the Open Biological and Biomedical Ontologies (OBO) Foundry.(8) It supports multi-institutional federated and centralized databases and registries, by including a standardized, O3 driven database schema and tables for enumerating standardized values. This provides a professional society driven pathway toward interoperable AI ready, multi-institutional data bases.

Materials and Methods

The AAPM BDSC is composed of a diverse international group of stakeholders currently, including 38 members: physicists (n=20), physicians (n = 13), informaticians (n=3), and other representatives (n=2). The BDSC includes members of several professional societies:

AAPM (n=24), ASTRO (n=26), CARO (n=1), COMP(n=2), ESTRO(n=6), ASCO(n=2), RSNA (n=8), AMIA (n=3), as well as members of NRG Oncology (n=8) and the National Institutes of Health (NIH) (n=4), National Cancer Institute (NCI)(n=3), IROC(n=1), IHE-RO(n=2), US Veterans Administration(n=2), Southwestern Oncology Group (SWOG)(n=3) and DICOM(n=1). (10) The number of members on the committee increased over time from the initial set of 23, as professional societies appointed representatives (e.g., ASTRO, n= 3) and as additional specialty area gaps were identified with resultant need to augment the base of domain experts (e.g., technologies such as heavy particle therapy, brachytherapy, diagnostic radiology, multi-national practitioners)

Engagement with External Initiatives

The 21st Century Cures Act promoted implementation of HL7 Fast Healthcare Interoperability Resources (HL7-FHIR) as the basis for improving patient access to their health care data and interoperable exchange of data among their providers. It requires the creation of FHIR-based application programming interfaces (APIs) to align industry efforts around interoperability and provide broader access to health care data. (11,12). HL7-FHIR supersedes HL7. HL7 is a set of standards widely used by hospital systems to transmit health care data. The HL7-FHIR standard overcomes significant gaps in HL7 by providing means to extend it with medical domain specific coding. Development of medical domain specific FHIR tags is being carried out as part of the mCODE (minimal Common Oncology Data Elements) and CodeX (Common Oncology Data Elements eXtensions) projects organized by Health Level 7 (HL7) and MITRE Corporation. MITRE corporation is a not for profit-organization that works in the public interest across federal, state and local governments, as well as industry and academia.

CodeX is a community accelerator focused on interoperable data modeling and applications that are integrating and testing mCode implementation based on the HL7 Fast Healthcare Interoperability Resources (FHIR). Within CodeX, the radiation therapy treatment data (RTTD) sub-group began work on a treatment summary use case in March of 2021. RTTD members ROIS and EHR representatives worked with HL7 and IHE-RO to develop HL7-FHIR frameworks. BDSC members are leaders in both CodeX and RTTD efforts. With the strong radiation therapy focus the elements, attributes, value sets, and relationships of O3 are a superset of the elements addressed in mCODE and CodeX informing development of those standards. AAPM and ASTRO are founding members of CodeX, with BDSC representatives leading the RTTD CodeX use case project started in March of 2021 (13) BDSC representatives in RTTD successfully lobbied for IHE-RO medical physics inclusion in RTTD in addition to radiation oncologists. AAPM also sponsored the Canadian Organization of Physicists in Medicine (COMP) and Society for Imaging Informatics in Medicine (SIIM) to join as CodeX and RTTD members.

Development of O3 in AAPM BDSC intersects with other efforts. BDSC communicated with members of the Office of the National Coordinator for Health Information Technology (ONC) proposing O3 driven extensions to the United States Cord Data for Interoperability (USCDI) standards. Mappings of O3 concepts to VA-ASTRO collaborations in developing quality measures were supported with BDSC members who were part of both efforts. (14)

To promote inter-operability of O3 with other coding systems e.g., SNOMED-CT (15) and NCI Thesaurus, BDSC enumerated and mapped O3 attributes to attributes in these other systems. NCI Thesaurus is a framework coding system for medical concepts used in clinical trial research. Similar to semantic-web ontologies, gaps and challenges in structure prompted recognition that this effort carried out by the professional society-based stakeholders would be of general benefit. BDSC engaged with leadership of the NCI Thesaurus to validate mappings and homologate the O3. SNOMED-CT is a coding system for multi-lingual health concepts. It has been designated as a national standard used by ONC and in HL7-FHIR. BDSC created mappings to SNOMED-CT codes and identified gaps. It worked in conjunction with CodeX in identifying and requesting SNOMED-CT codes for missing concepts.

Intersections of O3 with DICOM were explored with DICOM leaders. Members of BDSC include leaders in IHE-RO and the AAPM DICOM Working Group. O3 was presented and discussed at meetings of the DICOM working group at the AAPM annual meeting to promote mutual understanding and development of mutually beneficial standardizations.

Development Process

AAPM's BDSC was initiated July of 2019 and initially held monthly videoconference meetings to explore common ground from the stakeholders' collective experience of issues that typically compromise the formation of large inter- and intra- institutional databases from electronic records. The Radiation Oncology Translational Research Ontology detailed in the appendix of a 2018 publication by Mayo et al from the conference proceedings of the 2017 Practical Big Data Workshop formed the starting point for the O3. (16,17) Together with the prior work of several members in creating ontologies, and clinical databases the group's experience formed a strong basis for development. (1,18–21) The group decided to focus on prostate cancer as the initial use case for identifying attributes of key elements generalizable across disease site groups (e.g., Radiation Therapy Course) as well as disease site-specific subsets (e.g., Diagnosis and Staging Components for Prostate Cancer). Prostate cancer was chosen because of its high population incidence and public health relevance, with well-defined control and toxicity outcomes, as well as to reinforce joint efforts of ASTRO and the VA on prostate treatment quality measures at that time. The BDSC then increased the frequency of meetings to bi-weekly in April of 2020, to enable rapid progress in detailing the O3.

An iterative-deliberations approach was used to incrementally construct and refine the ontology.

Versions were recorded in standardized format Word document with standardized formatting delineating all aspects of key elements, attributes, and relationships. The word document was maintained on a shared web directory and new versions with changes were emailed to members. An application was created to parse iterations in the Word document and render it into a JSON file used to iteratively populate the O3 website. This provided a different format for inspection from the Word document to facilitate narrowing focus to specific sections of interest.

The group iteratively worked through each key element, attribute, value set, relationship and prioritization asking for objections or identifying gaps or omissions. Conversation on an item continued until there were no further issues or objections raised. If issues were subsequently identified, the group returned to the discussion of the item. The group worked until consensus was reached. Members not in attendance could reintroduce topics voted on previously. For highly specific, physician centric processes, smaller subgroups were identified and carried on the same process and then brought recommendations back to the larger group.

For attributes where there was consensus among clinicians on common value sets encountered in practice, e.g., imaging modalities used for M staging, these were enumerated, and the list was identified as extensible with other values. For those without common value sets, a Delphi process was employed to obtain consensus. (22,23) Survey recipients were members of ASTRO disease site specialty groups, BDSC, and other researchers and clinicians identified as actively engaged in standards development, deployment, and use.

For final recommendations of prioritization groupings for use in recommendations for clinical focus, scores were calculated based on the prioritizations assigned for key elements and attributes. Tables were constructed grouping attributes. These were iteratively reviewed with priorities for key elements and attributes refined to accord with group consensus for areas of focus.

The BDSC adopted an iterative, cyclical approach to engaging stakeholders beyond its membership to optimize the integration of diverse perspectives from the community (Figure 2a). The group took an approach during the project years (Figure 2b) permitting the addition of new data elements, as clinical and technical concepts evolved over time to avoid obsolescence. This iterative process included collating features deemed essential for supporting most clinical and research efforts in cancer care. The ontology's concepts underwent several iterations to optimize a hierarchical framework for data capture and reasoning from clinical records. An open-access website was created to enable public comment and to support wider distribution. (30) The BDSC solicited targeted, multi-stakeholder review and feedback beyond its membership in July 2021 and again in April 2022, with expansions and refinements made based on feedback received. Targeted engagements and feedback were also solicited from AAPM and ASTRO leadership, as well as CPQR contacts at of the 46 RT centers across Canada, CARO, COMP, ESTRO, major vendors of electronic records systems, and MITRE/HL7 leadership on CodeX projects.

Results

O3 version 1.0

The first version of O3 identified 42 key elements and 359 attributes, 144 value sets, and 155 relationships. Deliberations resulted in the ranked relative importance of key elements and attributes by merging clinical significance with the likelihood of availability in electronic records or the ability to modify routine clinical processes to permit aggregation. The subset of prostate cancer-specific attributes and value sets were a small fraction of the totals, 8 and 4, respectively. The group used rankings to subgroup the attributes into 5 priority rankings

to facilitate implementation. Table 1 characterizes the first three priority groupings of 24, 22, and 65 attributes, respectively. The remaining priority groupings are tabulated in the appendix. The complete O3 is provided on the website, including links to download the complete set as a spreadsheet or in JavaScript Object Notation (JSON) format for electronic processing. (24) Figure 3 illustrates a fully detailed attribute (“Disease Site Response”) as presented on the website.

Value Set Standardization

The group composed a standardized list of minimum sets of values that could be consistently implemented across electronic record systems and improve AI readiness by reducing the need to recategorize inputs. For each attribute identified, BDSC investigated if value sets existed that were widely recognized as an existing standard in clinical or research settings. These were listed as reference systems if the group noted more than one. The group then recommended value set items drawn from these standards (e.g., American Joint Committee on Cancer for staging). Value sets for attributes were designed anticipating eventual direct use in clinical radiation oncology information system interfaces to support consistency and quality.

Relationship quantification

High priority relationships between key elements which need to be captured were categorized. The existence of child key elements were predicated upon existence of the parent key element in ChildElement-Of and ParentElement-Of categorizations. Elements that exist independently but for which the relationship should be captured if it exists were categorized as IsAssociated-with.

As current radiation oncology information systems are constructed and used in practice, there is high probability of accurately capturing parent-child relationships. IsAssociated-with relationships may be less reliably captured given current workflows and systems.

For example, a patient key element (parent) will be linked to each instance of a diagnosis and staging (child) record. A patient may have many diagnosis and staging records only some of which are associated with a radiation therapy course. Ability to quantify the relationships automatically electronically between diagnosis and staging and radiation therapy courses is dependent on the current capabilities of the EHR, ROIS and TPS systems in use and how they are implemented in clinical practice

The Delphi process was required for a small number of high priority attributes for critical concepts: disease response categorizations, and relative prioritization of provider-reported outcomes and relative prioritization of toxicities. Members of BDSC and of ASTRO disease site specialty groups were engaged in the Delphi process using iterative survey methods and discussion to identify consensus-based recommendations.

Static vs Longitudinal Key Elements

Key elements were classified as static or longitudinal. Longitudinal key elements have attributes that may change over time. For example, sex at birth is a static attribute, but

gender identity is a longitudinal attribute. For longitudinal key elements: date and time, patient age as a decimal value, or both are captured with attribute values. Recording patient age at the time of an event, relative to midnight on the date of birth, with resolution of 3, 4, and 6 decimal places allows resolving time differences of days, hours, and minutes respectively, while also meeting PHI requirements for patients younger than 89. (16) O3 can support research scenarios with a finer resolution of time differences below the millisecond range. Our approach readily scales to this use case by increasing the number of decimal places used without requiring the recording of additional fields.

Concept Definition and Consistency

Lack of consistency in specific meanings for some key concepts used in describing radiotherapy was a significant barrier in standardizing use of several oncologic concepts, especially pertaining to the number of treatment courses, dose and fractionation, and outcomes. This lack of consistency affects several aspects of our field and undermines the ability to automate electronic aggregation to create and combine curated data sets. The BDSC iterated through a wide range of clinical use cases, working closely with the CodeX RTTD to identify consistent definitions for radiation therapy course, phase, sessions, and plan fractions, including multi-modality treatments and plan adaptation.

For example, the concept of “phase” had lacked an “operational” standardized meaning implementable within radiation oncology information systems. Working in the context of multi-vendor participation the BDSC and RTTD teams removed ambiguity. Phase was defined specifically as a concept for grouping treated plans using the same technique and modality to treat the same set of target volume-dose per fraction pairings. Similarly, course was defined to group all treated plans from first delivery of radiation in the course until physician believes they have treated all body sites in need of radiation therapy. A session within a course was identified as period between when the patient enters the treatment room until they leave it and used to group all plan fractions treated during the session. Use of decimal values for sessions allowed grouping when a session is interrupted (e.g. 3.0 and 3.1 for initial and resumed treatment after a mid-treatment multileaf collimator motor failure and replacement).

Figure 4 illustrates the framework of definitions for a) prostate, b) bilateral breast with adaptation, and c) oligometastatic sites with asynchronous starts and response timing gaps. These definitions were adopted as standards and incorporated into HL7-FHIR version 2.0.0, with the cooperation of EHR, ROIS, and TPS vendors and have recently (April 2022) been included in IHE-RO interoperability testing of vendor systems.

Historically some standards systems have evolved as a patchwork of specific individual use cases rather than a comprehensive view of all information needed within a domain. Cross mapping enumerated O3 attributes highlighted gaps in several areas for SNOMED-CT concepts and in the NCI thesaurus. O3 established a hierarchy of key elements, attributes, and value sets whereas BDSC engaged with NCI to share findings and coordinate mapping efforts. O3 contained 16 key elements and 54 attributes that did not have corresponding matches in SNOMED-CT. For NCI, there were 10 unmatched key elements and 54 unmatched attributes.

Database schema

To facilitate adoption into research applications, the operational ontology was rendered as a relational database. Scripts were created to automate creation of an instance of the SQL O3 database. Key elements were encoded as tables and attributes as fields within the SQL O3 database. SQL key fields were created to quantify relationships identified in O3. To simplify adoption, only parent-child relationships were required, allowing flexibility in encoding IsAssociated-with relationships. This flexibility is important when there may be missing data from specific data sets. The key elements, attributes, value sets and coding are stored in a table of the SQL O3 database. These may be subsequently used programmatically in user interfaces to create drop down lists to facilitate used of standard values and encoding. This facilitates later refinement of standard value sets. The approach is intended to allow flexibility for aggregating data into relational databases without adding rigidity by over specifying constraints beyond the primary parent and child relationships.

Intersections with other stakeholder groups – Treatment Summaries

The RTTD use case focused on treatment summaries from a small set of items identified in previous ASTRO white papers. (19,20) Informed by the superset of elements in O3, the RTTD and BDSC groups identified a standardized operational definition for the treatment phase and enumerated values for technique and modality. The CodeX RTTD use case team successfully lobbied for resolving discrepancies and extending values in publicly available SNOMED_CT codes. (25) The ASTRO, ESTRO and AAPM-endorsed TG-263 nomenclature formed the basis of target volumes used by the BDSC and RTTD use case groups to collaboratively identify extensions to TG-263, adding certain target volume designations to handle specific radiation oncology use cases. A canonical mapping of TG-263 nomenclature terms to SNOMED_CT codes and qualifiers was developed by RTTD and incorporated into mCODE (26).

Intersections with other stakeholder groups - Social Determinants of Health

Increasing awareness of the importance of social determinants of health (SDOH), patient ability/disability, and access to care as connected to cancer outcomes has emphasized the capture of robust data amenable to analysis that can inform policy change to improve cancer health equity. (27–33) In addition to traditional SDOH measures, BDSC successfully lobbied to include Disability in the United States Core Data for Interoperability (USCDI) (34) version 2 and is lobbying for Status as Primary Care Giver of Persons with Disability in version 3.

BDSC prioritized aggregating attributes quantifying social determinants of health to support modeling effects on outcomes and guide public health policy. O3 supports ASTRO's proposal of the HEART (Health Equity Achievement in Radiation Therapy) framework to capture and address healthcare disparities in radiation oncology.(33) ASTRO recommended that the HEART assessment be included in the proposed Radiation Oncology Alternative Payment Model to provide additional funding for wraparound services to address health disparities such as patient care navigators, personal support systems, and access to transportation and housing. Data associated with episodes of care with a HEART payment could be collected and used to determine the effectiveness of HEART interventions. The

assessment and data would inevitably assist in determining what causes these disparities and understanding what interventions are most effective to narrow the gaps.

Developing the O3 in conjunction with the CodeX RTTD provided opportunities to engage with government agencies, including the American Community Survey, to identify SDOH value sets that could be implemented with minimal clinic burden. The engagement method placed a priority on clinicians identifying value sets that could reasonably be incorporated into their clinical practice. Without this engagement process, these value sets may have been determined without critical input from our professional societies

Intersections with other stakeholder groups - DICOM

By design the committee included members who were also primary leaders in design and implementation of DICOM, IHE-RO and HL7-FHIR to avoid redundancy and promote dissemination of results from O3 to support those efforts as they continue to evolve. DICOM, HL7 and HL7-FHIR operate at the transport layer operationalizing transactional communication i.e. how is a treatment record sent from one system to another. DICOM is primarily used to transport imaging and radiation dose information. HL7 and its successor HL7-FHIR are used to transport all other health record information e.g. laboratory values. O3 operates at higher, semantic layer identifying which specific items of information and value sets need to be stored and made transmittable including which inter-relationships between key elements must be quantified in those transmissions. When the transport layer is developed ahead of reference to a semantic layer developed by professional society consensus, then the transport protocols can miss important connections needed to fully convey concepts. For example, the detailed clinical, procedural, and conceptual connections between treatment plans, phases, courses, doses to target volumes provided the foundation by vendors to create the HL7-FHIR connections so that treatment summary information was correctly transported between systems. Since leadership from these groups were included in the year's long development by design to promote dissemination of the work, some O3 concepts have subsequently been incorporated into those efforts.

Discussion

An ontology represents a common framework for categorizing key elements and inter-relationships of concepts in a specific domain with a standardized vocabulary supporting sharing, reusing data and enabling AI applications and computer reasoning (7). Ontologies have a strong foothold in genomics (35–38). Narrow scope ontologies have been gradually investigated by small groups of investigators in health care (39–43). Challenges faced in achieving the multi-institutional, multi-stakeholder consensus required for comprehensive healthcare domain-specific ontology development have slowed their adoption and wider application.

A few non-professional, society-based groups are working to create standardized values sets supporting data exchange. For instance, the Observational Health Data sciences and Informatics (OHDSI) constructed the Observational Medical Outcomes Partnership (OMOP), a broad clinical data model which has been applied successfully to facilitate general clinical data science efforts across institutions. (44–48) However, OMOP and other

similar standards remain limited by their more general scope, and therefore are unable to address many narrow-scope oncology-centric problems. Lack of quantification of important relationships between key elements, undermines likelihood of design implementations in clinical systems ensure ability to correlate and quantify dependencies. These gaps leave a need for more granular, domain-specific data models in oncology and radiation oncology.

“Operational” is a key component of O3. O3 was developed by clinical and technical radiation oncology domain experts for implementation within the electronic systems used in routine practice that govern clinical workflows. Leveraging the team’s expertise in the construction and use of metrics used in quality, safety, accreditation, billing, and research, O3 was designed to support these applications. The structure and value sets were created to support the design of consistent user interfaces within electronic systems to subsequently enable automated data extractions and reduce manual effort to implement.

By focusing on the development of a functional ontology supporting both clinical and research objectives, we have developed an engagement method for multi-stakeholder, consensus-driven standardized ontology development, engaging experts from key professional societies. Our goal is for this to support the creation of AI and analytics data sets from routine clinical practice, real-world data. This will facilitate a wide range of uses including verification of existing hypothesis-driven studies, support for reproducible analysis (model, data, or both model/data driven), data aggregation and pattern discovery, and correlative causality analysis. O3 based data standardization within RT electronic systems could provide a framework for integrating quality measures with billing codes, facilitating efficacy and efficiency of communication with health care insurers. Automated clinical summaries reduce clinician effort but also improve consistency of communication and understanding of RT terms for those outside of radiation oncology.

The creation of standardized Operational Ontology for Oncology aims to have a direct impact on gathering quality information from databases across medical systems. Standardization of the nuances of radiation therapy delivery across facilities can not only allow for reliable pooling of data between centers but could be integrated into national EHR consistency. By having standard data elements, EHR systems could extrapolate and integrate data within and across facilities as well as to insurers and payers. Once within the EHR in a standardized format, it can also help facilitate automatic documentation, improving efficiency within the clinics. The ontology can be used to inform categorization within REDCap, a secure online web application for databases, which allows for collaborative research projects not currently possible due to the various differences in volume naming, radiation prescriptions, etc currently in practices in the United States.

The O3 has been leveraged in the development of HL7-FHIR tags in the CodeX community. For example, the Radiation Therapy Treatment Data group started with the completion note use case. It was anticipated that effort would take longer than it has. The work with O3 and multi-institutional engagement greatly accelerated RTTD development. O3 has expanded beyond the first task for RTTD. The RTTD group is now working to extend including items identified as high priority in O3 while coordinating with the larger community of CodeX users beyond radiation therapy.

Relational databases are the backbone of enterprise level information systems. The relationships identified in O3 lend themselves to the Subject-Predicate-Object form of identifying relationships that are used in triple store databases. By design the committee included members who were also primary leaders in design and implementation of triple store databases and OBO foundry-based ontologies. These members are also part of a working group operating under BDSC to create a semantic web ontology. Multi-stakeholder collaboration in creating O3 that creating operational standardizations for aggregation and dissemination, provides a foundation for subsequent use of the data in triple stores. We anticipate that using the standardizations of O3, including the database schema rendering, will help to bridge collection of data in relational databases to eventual use of large-scale data sets in inferencing applications with triple stores.

We anticipate that quantifying and standardizing the relationships, key elements, attributes, and value sets will lay a foundation for subsequent development of automated AI algorithms to traverse O3 datasets to promote new discoveries in “real-world” electronic health records data. If the standards become adopted widely, then they can enable data from multiple research studies to be combined. For example, it then becomes possible to create algorithms that can automatically investigate the combined effects of chemotherapy, radiation therapy treatment and dosimetric details on outcomes and toxicities from multiple research studies if they use the O3 standard. Subsequent professional society guided refinement of O3 and the sql rendering will be informed by experience with implementations supporting clinical practice quality improvement and research.

Health care domain-specific knowledge is needed to develop standardizations that can be implemented at point of entry to improve data fidelity. The range of ostensible data “standards” needed to support a health care specialty often fall far short of their promise and practicability, when developed and imposed from outside of the domain expertise of clinicians. For example, Kush et al recently highlighted disappointing results following creation of common data elements sets (CDEs) developed as part of siloed NIH efforts. They found that “despite the promise and promulgation of CDEs over the past two decades, most are essentially a local resource and are not suitable for wholesale adoption and global reuse.”(49) The O3 provides support for implementing the principles of Findable Accessible, Interoperable and Reusable (FAIR) data and with the goal to break down these silos. (50) By standardizing data elements and attributes and value sets, along with creating a standardized database schema, the O3 helps enable data sets to be both interoperable and reusable. By adopting the database schema, it can become possible to share queries among researchers and studies, making data more accessible. Its collaborative framework provides a strong foundation for the continued partnership of professional societies in charting common courses in data science and data-driven practice quality initiatives.

BDSC is currently working with professional societies to incrementally add disease site-specific attributes for inclusion of all ASTRO-defined disease site specialties. Updates and extensions to O3 will be published to the public website along with software tools for creating O3 based databases and interfaces. A subgroup of the BDSC is also formalizing the ontology according to OBO Foundry principles, increasing both its exposure and its interoperability with other ontologies.(51)

Ultimately, this work is intended to harvest the data generated during patient care into forms that are more useful for discovery, safety ultimately beneficial for future patients. Data that can be validated will allow testing of sophisticated and provocative hypotheses and will move the field forward more quickly. We see this work as helping patients, helping to advance clinical research, and aiding in optimal resources utilizations as we move into a future featuring personalized, adaptive, motion-managed, imaged-guided, and biologically targeted combined-modality therapies.

Recommendations

Over the course of several meetings members debated primary recommendations for clinics to engage O3 and increase the volume of high priority data available. Discrete bulleted recommendations for clinics, researchers, professional societies and manufactures are summarized in Table 2. Investment in recommendations for manufacturers and for clinics can provide benefit to the field. Support for standardized data and methods could enable automated coding which would increase compliance and reduce work associated with code capture and validation. It would also provide a framework for integrating quality measures and procedure codes to facilitate pre-treatment payor validations and minimize prior authorization burdens.

Recommendations for clinics and manufacturers can increase patient safety and increase transparency of care for patients, by improving consistency of communication and enabling automated documentation to reduce clinician effort.

Recommendations for researchers would support efforts by the NIH and NCI to assemble large databases by reducing barriers for aggregation and increasing interoperability of data across data sets. By grounding these data sets in professional society, multi-stakeholder consensus driven standardizations focus can shift to more valuable uses of data rather than repeatedly resolving differences between groups. This provides a straightforward approach to integrating clinical and imaging data respecting the domain expertise and evolving approaches of each specialty. This is also responsive to new requirements in the NIH and NCI for sharable data. Integrating the standardization in O3 using the JSON file, also recommended for manufactures, would enable creation of tools for data curation.

Together with recommendations for manufactures and clinics, recommendations for researchers and professional societies will lower the barriers for aggregation of large data sets that could be used in creating large, representative, FAIR data sets supporting the scientific objectives of grant programs. Curation is typically an intensely manual process, but implementation of the proposed recommendations would pave the way toward development of standardized, automated tools to characterize variation in practice and outcomes with reduced effort and cost.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

[Acknowledgements]

This work is only possible through the combined effort of a large number of people, working on differing components of O3 and the multi-stakeholder methodology. The author list cannot adequately reflect the contributions of each of these individuals. Several authors also spent countless hours outside of the BDSC meetings coordinating efforts with AAPM, ASTRO, CARO, CodeX and other stakeholder groups. We do wish to highlight the efforts of a few individuals who in addition to the BDSC meetings also invested substantial time on additional sub projects or may not listed as authors. We thank David Hong, Kaley Woods, Huaizhi Geng, Sang Ho Lee, Du Wang, Ying Xiao for the herculean effort of looking up mapping codes from O3 to SNOMED-CT, and NCI thesaurus. George Chang, Luybov Remennick and Lori Whiteman of NIH/NCI carefully reviewed assignment of O3 codes mapped to codes in NCI Thesaurus pointing out recommended changes. During our professional society and public review cycle in July of 2021, many individuals made valuable comments and corrections which were incorporated into this version of O3. In particular, we wish to thank Laura Cervino, Drew Hope, Colleen Lawton, Chris Guy, Marie-Chantal Gagné, Ellen Kim, Jean Moran, Alberto Traverso, Emiliano Spezi, Joanna Kazmierska, Bill Nailon, Michaël Claessens, Lakshmi Santanam, Pengpeng Zhang and Mark Phillips for their insights and suggestions. We thank Anne Hubbard for her direction on ASTRO's work with the HEART score and applications of Social Determinants of Health measures.

References

1. Mayo CS, Kessler ML, Eisbruch A, Weyburne G, Feng M, Hayman JA, et al. The big data effort in radiation oncology: Data mining or data farming? *Adv Radiat Oncol*. 2016 Dec;1(4):260–71. [PubMed: 28740896]
2. ASTRO responds to CMS Radiation Oncology Model: Implementation delay and more reforms needed - American Society for Radiation Oncology (ASTRO) - American Society for Radiation Oncology (ASTRO) [Internet]. ASTRO. [cited 2022 Jun 2]. Available from: <https://www.astro.org/News-and-Publications/News-and-Media-Center/NewsReleases/2020/ASTRO-responds-to-CMS-Radiation-Oncology-Model-Imp>
3. Mayo CS, Moran JM, Bosch W, Xiao Y, McNutt T, Popple R, et al. American Association of Physicists in Medicine Task Group 263: Standardizing Nomenclatures in Radiation Oncology. *Int J Radiat Oncol Biol Phys*. 2018 Mar 15;100(4):1057–66. [PubMed: 29485047]
4. Mir R, Kelly SM, Xiao Y, Moore A, Clark CH, Clementel E, et al. Organ at risk delineation for radiation therapy clinical trials: Global Harmonization Group consensus guidelines. *Radiotherapy and Oncology*. 2020 Sep 1;150:30–9. [PubMed: 32504762]
5. Caissie A, Brown E, Olson R, Barbera L, Davis CA, Brundage M, et al. Improving patient outcomes and radiotherapy systems: A pan-Canadian approach to patient-reported outcome use. *Med Phys*. 2018 Oct;45(10):e841–4. [PubMed: 30144083]
6. Kildea J, Brown E, Letourneau E, Mayo C, Milette MP, Milosevic M, et al. Guidance on the Use of Common Nomenclature in Canadian Radiation Treatment Programs [Internet]. Canadian Partnership for Quality Radiotherapy; Available from: <https://www.cpqr.ca/wpcontent/uploads/2021/04/UCN.2021.03.01.pdf>
7. Gruber TR. A translation approach to portable ontology specifications. *Knowledge Acquisition*. 1993 Jun 1;5(2):199–220.
8. Jackson R, Matentzoglou N, Overton JA, Vita R, Balhoff JP, Buttigieg PL, et al. OBO Foundry in 2021: operationalizing open data principles to evaluate ontologies. *Database*. 2021 Sep 29;2021:baab069.
9. Welcome to the NCBO BioPortal | NCBO BioPortal [Internet]. [cited 2022 Nov 17]. Available from: <https://bioportal.bioontology.org/>
10. AAPM Committee Tree - Big Data Subcommittee (BDS) [Internet]. [cited 2022 May 22]. Available from: https://www.aapm.org/org/structure/default.asp?committee_code=BDS
11. HITInfrastructure. Final Interoperability Rule Has Implications for APIs, FHIR [Internet]. HITInfrastructure. 2020 [cited 2022 May 22]. Available from: <https://hitinfrastructure.com/news/final-interoperability-rule-has-implications-for-apis-fhir>
12. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program [Internet]. Federal Register. 2020 [cited 2022 Jun 2]. Available from: <https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification>

13. Radiation Therapy Treatment Data for Cancer - CodeX - Confluence [Internet]. [cited 2022 May 24]. Available from: <https://confluence.hl7.org/pages/viewpage.action?pageId=76162628>
14. Katsoulakis E, Kudner R, Chapman C, Park J, Puckett L, Solanki A, et al. Consensus Quality Measures and Dose Constraints for Head and Neck Cancer with an emphasis on Oropharyngeal and Laryngeal Cancer from the Veterans Affairs Radiation Oncology Quality Surveillance Program and American Society for Radiation Oncology (ASTRO) Expert Panel. Practical Radiation Oncology [Internet]. 2022 Jun 3 [cited 2022 Jun 8];0(0). Available from: [https://www.practicalradonc.org/article/S1879-8500\(22\)00185-0/fulltext](https://www.practicalradonc.org/article/S1879-8500(22)00185-0/fulltext)
15. SNOMED CT [Internet]. U.S. National Library of Medicine; [cited 2022 May 22]. Available from: <https://www.nlm.nih.gov/healthit/snomedct/index.html>
16. Mayo CS, Phillips M, McNutt TR, Palta J, Dekker A, Miller RC, et al. Treatment data and technical process challenges for practical big data efforts in radiation oncology. Med Phys. 2018 Oct;45(10):e793–810. [PubMed: 30226286]
17. Mayo C. Community science and reaching the promise of big data in health care. Med Phys. 2018 Oct;45(10):e790–2. [PubMed: 30307633]
18. McNutt TR, Bowers M, Cheng Z, Han P, Hui X, Moore J, et al. Practical data collection and extraction for big data applications in radiotherapy. Med Phys. 2018 Oct;45(10):e863–9. [PubMed: 29446451]
19. Christodouleas JP, Anderson N, Gabriel P, Greene R, Hahn C, Kessler S, et al. A Multidisciplinary Consensus Recommendation on a Synoptic Radiation Treatment Summary: A Commission on Cancer Workgroup Report. Pract Radiat Oncol. 2020 Dec;10(6):389–401. [PubMed: 31988040]
20. Hayman JA, Dekker A, Feng M, Keole SR, McNutt TR, Machtay M, et al. Minimum Data Elements for Radiation Oncology: An American Society for Radiation Oncology Consensus Paper. Pract Radiat Oncol. 2019 Nov;9(6):395–401. [PubMed: 31445187]
21. Kapoor R, Sleeman WC, Nalluri JJ, Turner P, Bose P, Cherevko A, et al. Automated data abstraction for quality surveillance and outcome assessment in radiation oncology. J Appl Clin Med Phys. 2021 Jul;22(7):177–87.
22. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol. 2014 Apr;67(4):401–9. [PubMed: 24581294]
23. Humphrey-Murto S, Varpio L, Gonsalves C, Wood TJ. Using consensus group methods such as Delphi and Nominal Group in medical education research. Med Teach. 2017 Jan;39(1):14–9.
24. Mayo C, Feng M, Brock K, Hong D, McNutt T, Buchsbaum J, et al. American Association of Physicists in Medicine – Operational Ontology for Radiation Oncology. 2021 Jan 5 [cited 2022 Feb 27]; Available from: <https://aapmbdsc.azurewebsites.net/Home/About?ReturnUrl=%2F>
25. HL7.FHIR.US.MCODE/Radiotherapy Modality And Technique Extension - FHIR v4.0.1 [Internet]. [cited 2022 May 24]. Available from: <https://hl7.org/fhir/us/mcode/StructureDefinition-mcode-radiotherapy-modality-and-technique.html>
26. HL7-FHIR Radiation Therapy Target Volume [Internet]. [cited 2022 May 24]. Available from: <https://hl7.org/fhir/us/mcode/StructureDefinition-mcode-radiotherapy-volume.html>
27. Mohan G, Chattopadhyay S. Cost-effectiveness of Leveraging Social Determinants of Health to Improve Breast, Cervical, and Colorectal Cancer Screening: A Systematic Review. JAMA Oncol. 2020 Sep 1;6(9):1434–44. [PubMed: 32556187]
28. Alcaraz KI, Wiedt TL, Daniels EC, Yabroff KR, Guerra CE, Wender RC. Understanding and addressing social determinants to advance cancer health equity in the United States: A blueprint for practice, research, and policy. CA Cancer J Clin. 2020 Jan;70(1):31–46. [PubMed: 31661164]
29. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: A systematic review. J Am Med Inform Assoc. 2020 Nov 1;27(11):1764–73. [PubMed: 33202021]
30. Wang M, Pantell MS, Gottlieb LM, Adler-Milstein J. Documentation and review of social determinants of health data in the EHR: measures and associated insights. J Am Med Inform Assoc. 2021 Nov 25;28(12):2608–16. [PubMed: 34549294]

31. Kwon IWG, Kim SH, Martin D. Integrating Social Determinants of Health to Precision Medicine through Digital Transformation: An Exploratory Roadmap. *Int J Environ Res Public Health*. 2021 May 10;18(9):5018. [PubMed: 34068530]
32. CDC. Disability and Health Data System (DHDS) | CDC [Internet]. Centers for Disease Control and Prevention. 2022 [cited 2022 May 24]. Available from: <https://www.cdc.gov/ncbddd/disabilityandhealth/dhds/index.html>
33. Radiation oncologists urge Congress to reverse proposed CMS cuts and create more equity in access to cancer treatments - American Society for Radiation Oncology (ASTRO) [Internet]. ASTRO. [cited 2022 May 24]. Available from: <https://www.astro.org/News-and-Publications/News-and-Media-Center/News-Releases/2021/Radiation-oncologists-urge-Congress-to-reverse-pro>
34. United States Core Data for Interoperability (USCDI) [Internet]. [cited 2022 May 22]. Available from: <https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi>
35. Barcelona V, Montalvo-Ortiz JL, Wright ML, Nagamatsu ST, Dreisbach C, Crusto CA, et al. DNA methylation changes in African American women with a history of preterm birth from the InterGEN study. *BMC Genom Data*. 2021 Sep 5;22(1):30. [PubMed: 34482817]
36. Das S, Hussey P. ContSOnto: A Formal Ontology for Continuity of Care. *Stud Health Technol Inform*. 2021 Oct 27;285:82–7. [PubMed: 34734855]
37. Hu L, Ma J, Cao M, Lin Y, Long W, Shi Z, et al. Exosomal mRNA and lncRNA profiles in cord blood of preeclampsia patients. *J Matern Fetal Neonatal Med*. 2021 Sep 1;1–11.
38. Haimel M, Pazmandi J, Heredia RJ, Dmytrus J, Bal SK, Zoghi S, et al. Curation and expansion of Human Phenotype Ontology for defined groups of inborn errors of immunity. *J Allergy Clin Immunol*. 2021 May 12;S0091–6749(21)00732–6.
39. Traverso A, van Soest J, Wee L, Dekker A. The radiation oncology ontology (ROO): Publishing linked data in radiation oncology using semantic web and ontology techniques. *Med Phys*. 2018 Oct;45(10):e854–62. [PubMed: 30144092]
40. Ju C, Zhang S. Doctor Recommendation Model Based on Ontology Characteristics and Disease Text Mining Perspective. *Biomed Res Int*. 2021;2021:7431199.
41. Michalowski M, Wilk S, Michalowski W, O'Sullivan D, Bonaccio S, Parimbelli E, et al. A Health eLearning Ontology and Procedural Reasoning Approach for Developing Personalized Courses to Teach Patients about Their Medical Condition and Treatment. *Int J Environ Res Public Health*. 2021 Jul 9;18(14):7355. [PubMed: 34299806]
42. Prieto Santamaría L, Fernández Lobón D, Díaz-Honrubia AJ, Ruiz EM, Nifakos S, Rodríguez-González A. Towards the Representation of Network Assets in Health Care Environments Using Ontologies. *Methods Inf Med*. 2021 Oct 5;
43. Steiner B, Saalfeld B, Elgert L, Haux R, Wolf KH. OnTARi: an ontology for factors influencing therapy adherence to rehabilitation. *BMC Med Inform Decis Mak*. 2021 May 11;21(1):153. [PubMed: 33975585]
44. Belenkaya R, Gurley MJ, Golozar A, Dymshyts D, Miller RT, Williams AE, et al. Extending the OMOP Common Data Model and Standardized Vocabularies to Support Observational Cancer Research. *JCO Clin Cancer Inform*. 2021 Jan;5:12–20. [PubMed: 33411620]
45. Belenkaya R, Gurley M, Dymshyts D, Araujo S, Williams A, Chen R, et al. Standardized Observational Cancer Research Using the OMOP CDM Oncology Module. *Stud Health Technol Inform*. 2019 Aug 21;264:1831–2. [PubMed: 31438365]
46. Maier C, Kapsner LA, Mate S, Prokosch HU, Kraus S. Patient Cohort Identification on Time Series Data Using the OMOP Common Data Model. *Appl Clin Inform*. 2021 Jan;12(1):57–64. [PubMed: 33506478]
47. Reinecke I, Zoch M, Reich C, Sedlmayr M, Bathelt F. The Usage of OHDSI OMOP - A Scoping Review. *Stud Health Technol Inform*. 2021 Sep 21;283:95–103. [PubMed: 34545824]
48. Observational Medical Outcomes Partnership (OMOP) [Internet]. Center for Healthcare Improvement and Medical Effectiveness — CHIME. [cited 2022 May 22]. Available from: <https://chime.ucsf.edu/observational-medical-outcomes-partnership-omop>
49. Kush RD, Warzel D, Kush MA, Sherman A, Navarro EA, Fitzmartin R, et al. FAIR data sharing: The roles of common data elements and harmonization. *J Biomed Inform*. 2020 Jul;107:103421.

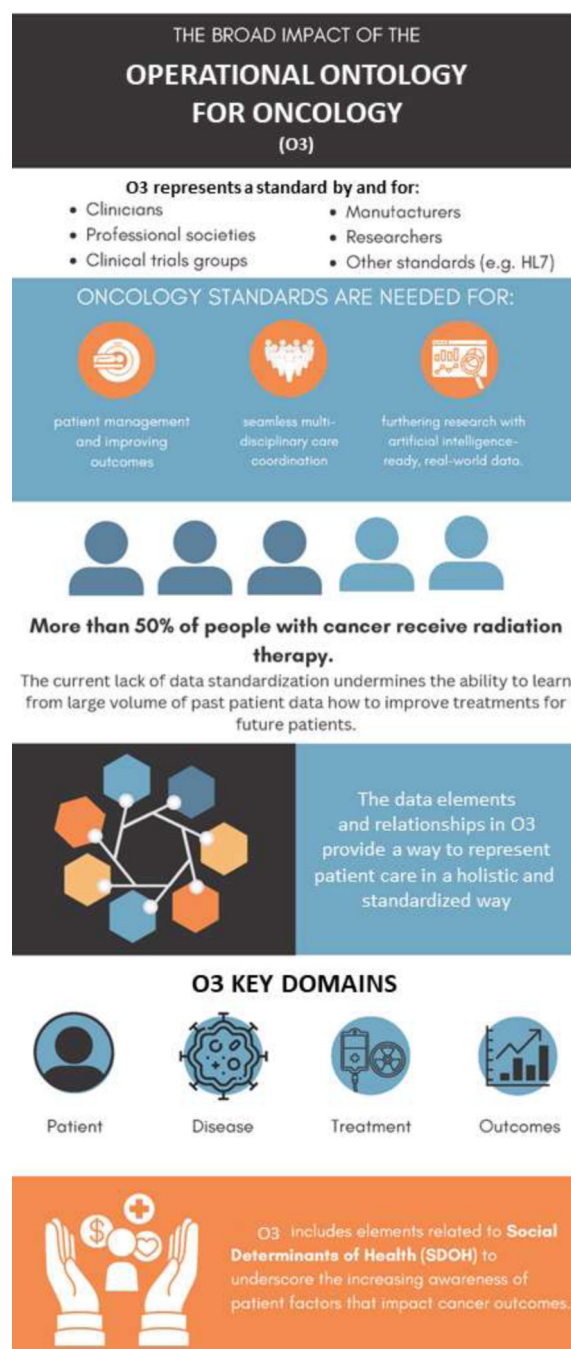
50. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016 Mar 15;3(1):160018.
51. Phillips MH, Serra LM, Dekker A, Ghosh P, Luk SMH, Kalet A, et al. Ontologies in radiation oncology. *Phys Med*. 2020 Apr;72:103–13. [PubMed: 32247963]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Figure 1.**

A professional society based, consensus driven approach was taken in developing the Operational Ontology for Oncology to address standardization gaps that undermine ability to aggregate and learn from “real-world” data, supporting clinical practice improvement and research.

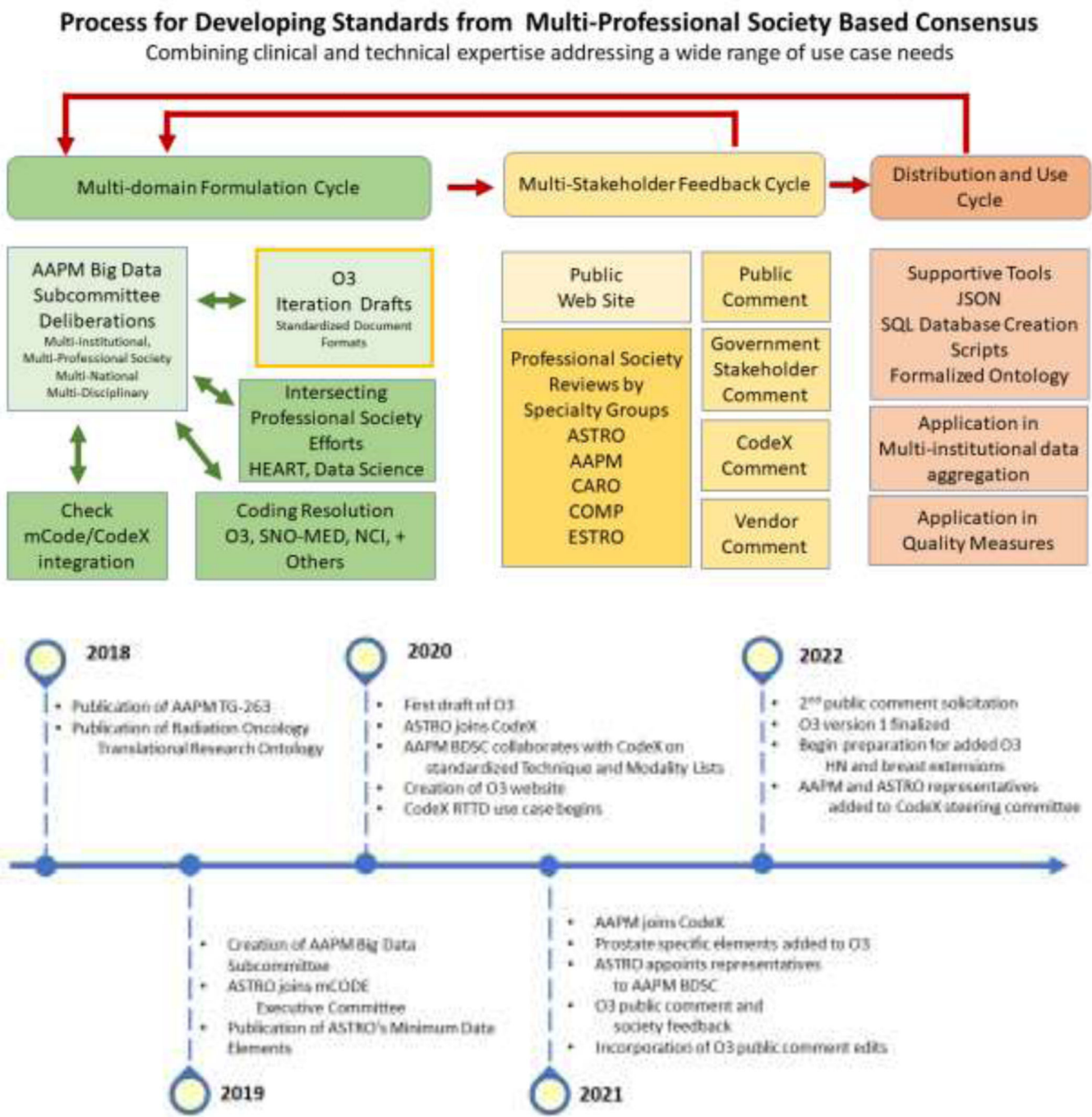


Figure 2.
A) An iterative, progressive engagement method including multiple health care electronic systems vendors, government agencies, specialty groups within professional societies, institutions and the public were engaged with a publicly accessible website for collecting comments and direct engagement with representatives to ensure incorporation of multi-stakeholder perspectives. B) building in prior work in 2018, the work was carried over multiple years with progressive engagement of stakeholders

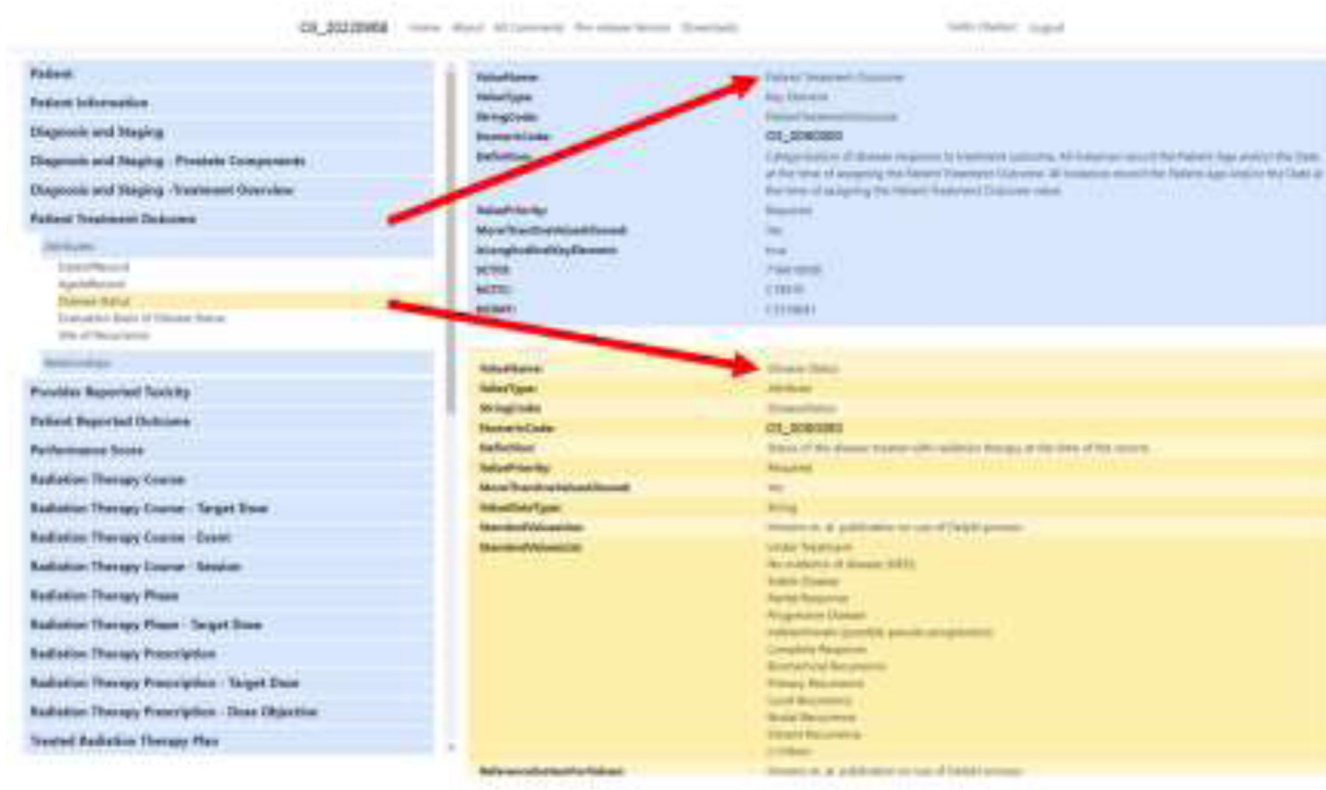
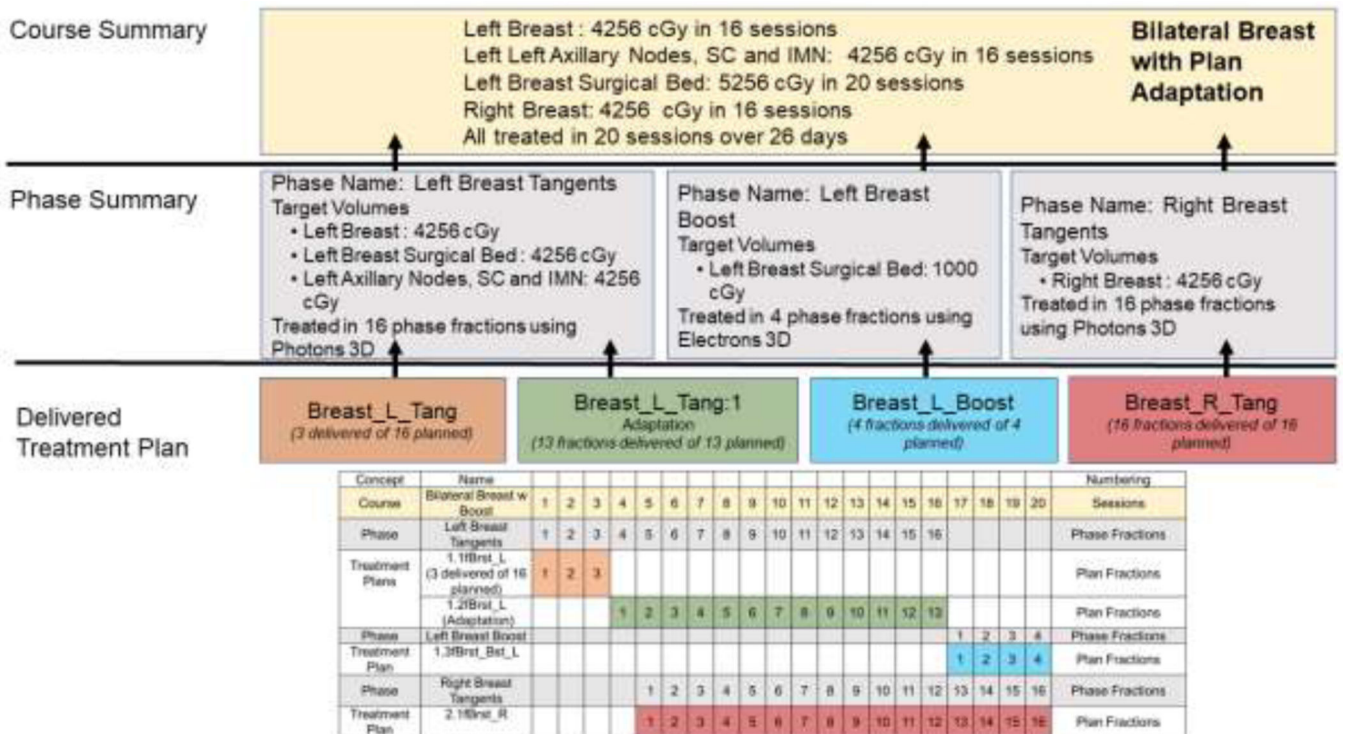
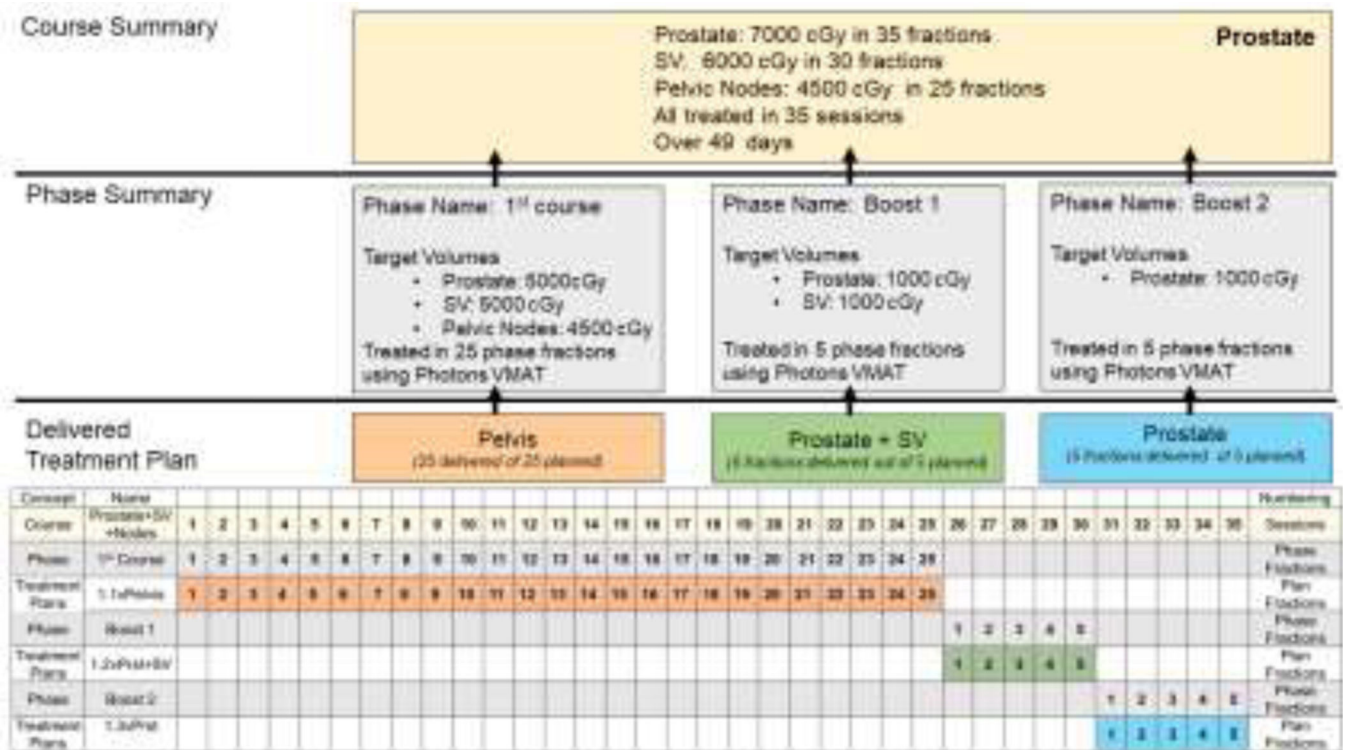


Figure 3.

An O3 website screenshot illustrating the details of one of five attributes under the Key Element of “Patient Treatment Outcome”. The O3 is formatted in a user-friendly drop-down menu with each key element listed to the left of the page (e.g., Patient Treatment Outcome), providing users the ability to access additional information through menu expansion including value name (e.g. Disease Status), value type (Attribute), definition, list of standard values with reference for such standard along with O3 codes for the value in comparison to current SCTID, NCITC and NCIMT codes. Relationships that should also be tracked are tabulated.



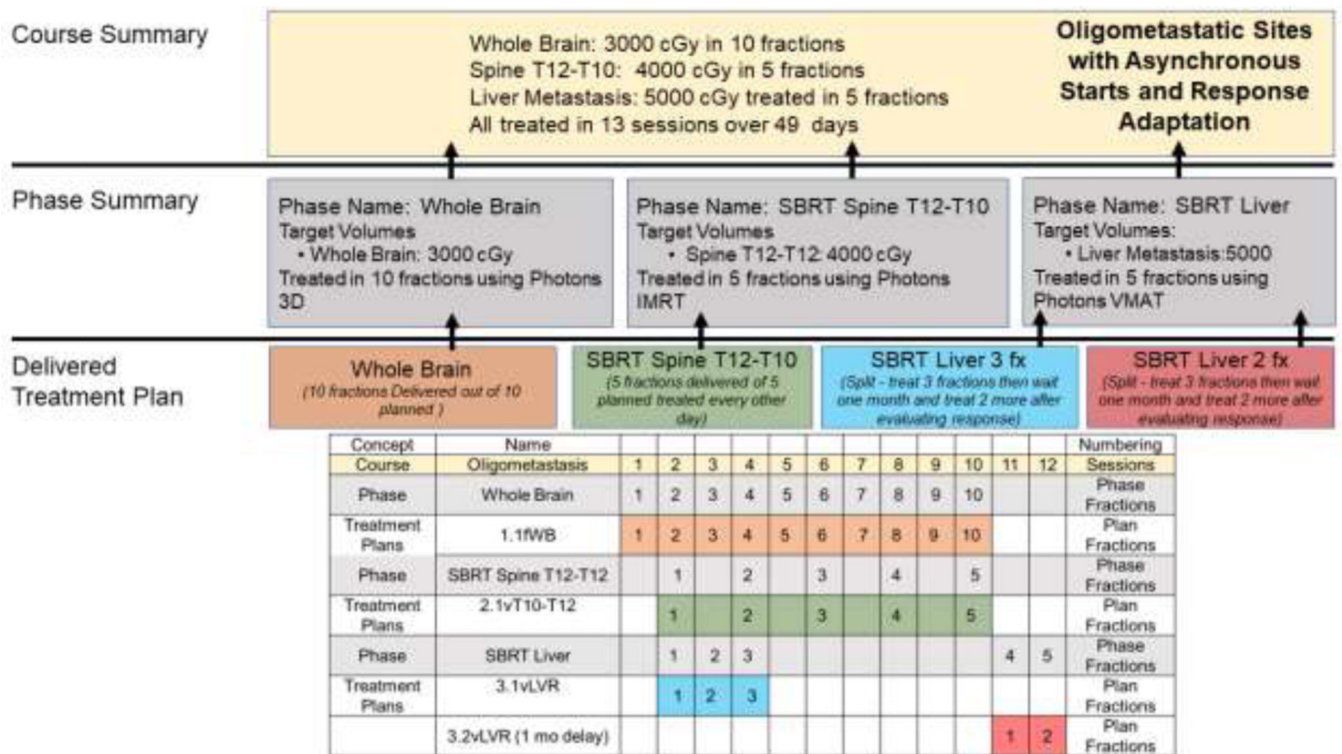


Figure 4.

Illustration of radiotherapy treatment concepts of radiation therapy course, phase, plan, session, and fraction using three clinical examples ordered by increasing complexity. A) Prostate cancer, including three treatment plans delivered to increasingly focused targets of the pelvis, prostate/seminal vesicles (SV) and prostate in 3 sequential phases for a total of 35 fractions/sessions. B) Bilateral breast cancer, including three treatment plans delivered to left (L) breast/axilla, boost site of L breast and right (R) breast. In this case there is a fourth plan due to adaptation required of the L breast/axilla plan which had only 3 fractions delivered of the 16 planned. While the L breast/axilla and L breast boost phases were treated sequentially, the phase of R breast plan was treated concurrently (in the same sessions) with those of the contralateral treatments. C) Oligometastatic disease, including a total of 4 treatment plans delivered over a total of 12 sessions. The two plans within the liver phase were treated asynchronously. Following evaluation of treatment response, the second liver plan of 2 fractions was delivered one month after the first liver plan of 3 fractions, which had been treated concurrently with the whole brain (10 fraction) and spine (5 fraction) plans

Table 1

Attributes recommended as a) first, b) second and c) third priority for efforts to improve systematic aggregation for all patients. Attribute values constrained by patient health information (PHI, ♦) and those supporting measurement of social determinants of health (SDOH, ⌘) are indicated. Values judged implementable with electronic (★) or manual (☆) extraction or likely missing (⊙) from current clinical record systems are indicated. Values for Longitudinal elements (✓) record date and time as well as high precision patient age. Fourth and fifth priority tables are provided in the appendix.

Table 1A		First priority for aggregation efforts			
Key Element	Attribute				
Patient	Patient MRN	♦		★	
	Anonymized Patient Identifier			⊙	
	Patient Identifier Data Set			★	
	Date of Birth	♦		★	
	Year of Birth			★	
	Date of Death	♦		★	
	Age At Date of Death			★	
	Sex at Birth			★	
Patient Information	Vital Status			★	✓
	Postal Code of Domicile	♦	⌘	★	✓
Diagnosis and Staging	Staging System			☆	✓
	ICD Version			★	✓
	ICD Code			★	✓
	Histology			☆	✓
	Staging-T Category (e.g. Clinical, Pathologic)			★	✓
	Staging-T			★	✓
	Staging-N			★	✓
	Staging-M			★	✓
Diagnosis and Staging – Treatment Overview	Had Radiation Therapy As Part of Treatment			★	✓
	Had Systemic Therapy As Part of Treatment			★	✓
	Had Surgical Procedure As Part of Treatment			★	✓
Diagnosis and Staging - Prostate Components	Had Prostatectomy			☆	✓
	Prostatectomy Margin Status			☆	✓
	Disease Status			☆	✓
Table 1B		Second priority for aggregation efforts			
Key Element	Attribute Name				
Patient	Cause of Death			⊙	
	Cause of Death Attributable to Treatment			⊙	
	Patient Reported Race		⌘	★	
	Patient Reported Ethnicity		⌘	★	
	Gender Identity		⌘	⊙	✓

Table 1A	First priority for aggregation efforts				
Key Element	Attribute				
	Has tobacco use history			★	✓
	Height			★	✓
	Weight			★	✓
	Address of Domicile	❖		★	✓
	Address of Primary Care Physician	❖		☆	✓
	Insurance		⌘	⊙	✓
	Disability		⌘	⊙	✓
	Caregiver of disabled person status		⌘	⊙	✓
	Education Level		⌘	⊙	✓
	Primary Language Spoken at Home		⌘	⊙	✓
Diagnosis and Staging	Pathology Molecular Test Name			☆	✓
	Pathology Molecular Test Result			☆	✓
Diagnosis and Staging - Prostate Components	PSA Value			★	✓
	Primary Gleason Score at Biopsy			★	✓
	Secondary Gleason Score at Biopsy			★	✓
Patient Treatment Outcome	Evaluation Basis of Disease Status			☆	✓
	Site of Recurrence			☆	✓
Table 1C	Third priority for aggregation efforts				
Key Element	Attribute Name				
Provider Reported Toxicity	Toxicity Coding System			☆	✓
	Toxicity Measure			☆	✓
	Toxicity Value			☆	✓
Patient Reported Outcome	PRO Instrument			⊙	✓
	PRO Question ID			⊙	✓
	PRO Question			⊙	✓
	PRO Question Response			⊙	✓
Performance Score	Scoring System			☆	✓
	Performance Score Value			☆	✓
Radiation Therapy Course	Radiation Therapy Course Facility			★	✓
	Radiation Therapy Course Facility Postal Code	⌘		★	✓
	Number Of Sessions in Course			★	✓
	Course Involves Reirradiation			☆	✓
Radiation Therapy Course - Target Dose	Target Volume			☆	
	Target Volume Dose			☆	
	Dose Unit			☆	
Radiation Therapy Prescription	Name of Radiation Therapy Prescription			☆	✓
	Number of Treatment Sessions			★	✓
	Number of Sessions Per Day			★	✓

Table 1A	First priority for aggregation efforts				
Key Element	Attribute				
	Number of Sessions Per Week			★	✓
Radiation Therapy Prescription - Target Dose	Target Volume			☆	
	Target Volume Dose			☆	
	Dose Unit			☆	
Radiation Therapy Prescription - Dose Objective	Structure Type			⊙	
	Structure Name			⊙	
	Dose Objective			⊙	
	Objective Condition			⊙	
	Value			⊙	
Treated Radiation Therapy Plan	Plan Name			★	
	Number of Fractions Intended			★	
	Number of Fractions Treated			★	
	Adaptation Or Revision			☆	
	Modality			★	
	Dose Delivery Category			☆	
	Is plan representing cumulative phase			☆	
	Is plan representing cumulative course			☆	
	Is Plan Sum			★	
Radiation Therapy Treated Plan - Target Dose	Target Volume			☆	
	Target Volume Dose			☆	
	Dose Unit			☆	
Treatment Plan Details - External Beam Radiation Therapy	Technique			★	
	Delivery Device			★	
	Energy			★	
Treatment Plan Details - Brachytherapy	Technique			★	
	Delivery Device			★	
	Energy			★	
Treatment Plan Details - Radiopharmaceutical	Technique			☆	
	Energy			☆	
Treatment Plan Details - Hadrons	Area Coverage Method			☆	
	Robust Optimization			☆	
	Relative Biological Effectiveness (RBE) Basis			☆	
	Relative Biological Effectiveness (RBE)			☆	
	Used Repainting			☆	
Treatment Plan - Field Details - Hadrons	Field Name			☆	
	Range Shifter			☆	
Systemic Therapy Course	Systemic Therapy Course Facility			★	✓
	Systemic Therapy Course Facility Postal Code		⌘	★	✓

Table 1A	First priority for aggregation efforts				
Key Element	Attribute				
	Systemic Therapy Course Type			☆	✓
Systemic Therapy Cycle	Systemic Therapy Cycle Name			★	✓
	Cycle Status			★	✓
Systemic Therapy Cycle Drugs Used - Chemotherapy	Drug Name			★	
Interventional Procedure	Interventional Procedure Facility			★	✓
	Facility Postal Code		⌘	★	✓
	Interventional Procedure Type			☆	✓
	Is Pre-Irradiation			★	✓

Table 2

Recommendations to clinics, manufacturers, granting agencies and professional societies for leveraging O3 to meet mutual objectives for quality, safety, improved workflow and FAIR data sets.

Table 2 Recommendations	
1	A. For manufacturers of Electronic Record Systems, Radiation Oncology Information Systems and Treatment Planning Systems applications should implement OORO in user interfaces ordered by the prioritized tables. The standardized JSON formatted OORO file can be used to automatically populate drop down lists in user interfaces including mappings to other coding systems (SNOMED-CT, NCI, etc) and allow periodic update as OORO is updated. This structure would support interoperability among multi-vendor systems, and design of clinic customizable “forcing functions” based on professional society-based standards to ease data collection at part of routine practice. B. Use the OORO relationships to ensure database recording methods enable ability to reliably extract high value related items and relationships (e.g., from diagnosis to treatment course and outcomes). Currently even the highest priority elements and relationships, that are routinely needed by physicians are not well optimized for entry, retrieval, and connection.
2	A. For radiotherapy centers OORO should be embedded withing the programs quality framework that focuses on Radiation Therapy data standardizations. Centers may choose to implement OORO in a stepwise fashion guided by prioritization of clinical relevance and ease of implementation. To start full implementation of table 1a should be achievable within 1–2 years for many clinics and within 3 years for the vast majority. B. Clinics should examine and adjust clinical processes to assure highest priority attributes are captured in existing EHR and ROIS systems as part of routine use in a fashion supporting accurate, automated electronic capture. C. Many data elements in tables 1a and b are currently required by accrediting bodies (e.g ACR, APEX, COC, ASCO). We recommend use of OORO as an organizing reference for accreditation bodies to increase consistency and interoperability of quality metrics.
3	A. For researchers, OORO should form the basis of RT outcomes research with inclusion in clinical trials and grant applications. B. Use the attribute tables and standardized value sets in clinical quality and research designs. This reduces effort and the resulting interoperability supports ability for re-use of data sets. Promoting use of the standardized OORO including value sets and mappings to other systems (e.g., NCI-Thesaurus) supports government granting agencies for creating large scale interoperable data sets while synchronizing common objectives with professional societies. C. Use the freely available, OORO based database schema as the basis for creating learning health systems in radiation oncology. This increases ability to interoperable share data and enables development of shared tools for data extraction and analysis to reduce development effort by local staff. D. Encourage NIH and NCI to integrate the standardized OORO JSON data file into data curation applications such as the Pearl Opens Source DICOM Archive (POSDA) to create standardized curation tools supporting large scale aggregations of clinical data with imaging data. E. Increase support for consolidation standardizations efforts among groups developing coding systems, taxonomies and ontologies.
4	A. For professional societies, continue support for collaborative leadership and effort among professional societies in defining standards that support mutual objectives for clinical practice and research. B. Engage with government regulatory bodies to promote the role of professional societies in defining standardizations used as the basis for public health and reimbursement. C. Continue and expand engagement with stakeholders outside of radiation oncology’s traditional boundaries e.g., CodeX, USCDI to promote data policies defined by outside groups that are well aligned to the expertise of professional societies. D. Advocate for standards through groups such as the Global Harmonization Group – Global Quality Assurance of Radiation Therapy Clinical Trials