

# Structured Data Capture for Oncology

Alexander K. Goel, MI<sup>1</sup>; Walter Scott Campbell, PhD, MBA<sup>2</sup>; and Richard Moldwin, MD, PhD<sup>1</sup>

Lack of interoperability is one of the greatest challenges facing healthcare informatics. Recent interoperability efforts have focused primarily on data transmission and generally ignore data capture standardization. Structured Data Capture (SDC) is an open-source technical framework that enables the capture and exchange of standardized and structured data in interoperable data entry forms (DEFs) at the point of care. Some of SDC's primary use cases concern complex oncology data such as anatomic pathology, biomarkers, and clinical oncology data collection and reporting. Its interoperability goals are the preservation of semantic, contextual, and structural integrity of the captured data throughout the data's lifespan. SDC documents are written in eXtensible Markup Language (XML) and are therefore computer readable, yet technology agnostic—SDC can be implemented by any EHR vendor or registry. Any SDC-capable system can render an SDC XML file into a DEF, receive and parse an SDC transmission, and regenerate the original SDC form as a DEF or synoptic report with the response data intact. SDC is therefore able to facilitate interoperable data capture and exchange for patient care, clinical trials, cancer surveillance and public health needs, clinical research, and computable care guidelines. The usability of SDC-captured oncology data is enhanced when the SDC data elements are mapped to standard terminologies. For example, an SDC map to Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) enables aggregation of SDC data with other related data sets and permits advanced queries and groupings on the basis of SNOMED CT concept attributes and description logic. SDC supports terminology maps using separate map files or as terminology codes embedded in an SDC document.

JCO Clin Cancer Inform 5:194-201. © 2021 by American Society of Clinical Oncology

Licensed under the Creative Commons Attribution 4.0 License 

## INTRODUCTION

Interoperability, in the context of complex oncology data sets, is the ability to share and reuse data across multiple nodes without semantic, contextual, or structural loss.<sup>1</sup> Reuse of data refers primarily to secondary usage in an external data ecosystem for purposes such as patient care, cancer surveillance, research, and clinical trials. Interoperability is greatly enhanced by standardizing the structure of contextually related data fields, before capturing in an electronic health record (EHR) system.

For patient care, preservation of structure and context is critical, from the data entry form (DEF) through all downstream clinical reports. Centralized standardization of data entry fields during the data collection design process, with a focus on downstream interoperability and data reuse, has several benefits.<sup>2</sup> The design of data fields and DEF structure by centralized expert teams can make data entry more consistent and efficient, aiding in the data entry process. Standardization of data entry with consistent evidence-based data fields

helps to ensure complete collection of clinically critical data in a familiar format and enables the generation of consistent, standardized, and structured reports, regardless of EHR vendor, institution, or variations in the cosmetics of DEF and report formats.<sup>3-5</sup>

Unfortunately, this type of precapture standardization is rarely addressed by EHR vendors. Attempts to standardize and/or aggregate data fields across EHRs after the data are collected often require a significant effort in data aggregation and cleaning and often yield suboptimal results.<sup>6,7</sup> The lack of precapture semantic, contextual, and structural standardization is thus a significant barrier to the complex data analyses required in oncology investigations and is a barrier to sharing data with patients, their care teams, and other EHR systems.<sup>8</sup>

Structured Data Capture (SDC) is an open-source technical framework published by the Quality Research and Public Health committee of the standards organization Integrating the Healthcare Enterprise (IHE). SDC was designed to solve the problem of

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on December 1, 2020 and published at [ascopubs.org/journal/cci](https://ascopubs.org/journal/cci) on February 16, 2021: DOI <https://doi.org/10.1200/CCI.20.00103>

## CONTEXT

### Key Objective

Review the current state of the Structured Data Capture (SDC) initiative in the oncology data ecosystem.

### Knowledge Generated

SDC is a computer-readable information model that defines the information content of data-entry forms, supports multiple approaches for SDC data exchange, and enables secondary use of standardized data.

### Relevance

SDC templates allow clinical information to be standardized for data capture before entry into computer systems. SDC is especially valuable for the collection and exchange of rapidly versioned data elements such as those found in pathology data sets, cancer staging, and clinical trials.

precapture data standardization in an interoperable manner. SDC can be viewed as a model that specifies the structure of related data elements (DEs) and preserves their semantic and contextual integrity. Furthermore, SDC specifies the information content of interoperable DEFs so that the DEF user can capture, store, and exchange complex, context-rich data in standardized DEs.<sup>9</sup> An SDC template specifies the content of a DEF that can be rendered by any EHR vendor in a technology-agnostic manner, while maintaining an exact representation of the data definitions, allowing the captured data to be exchanged in an interoperable manner. SDC-based DEFs are particularly useful for designing and exchanging complex oncology data sets, such as those needed for anatomic pathology, biomarkers, and clinical oncology reporting.

Since 2019, SDC has been the delivery format for the electronic Cancer Checklists (eCCs) from the College of American Pathologists (CAPs). These checklists are used by 35%-40% of North American pathologists.<sup>10,11</sup> Much of the data captured by these forms are submitted to North American cancer registries for public health surveillance.<sup>12,13</sup> Other clinical specialties (eg, radiology and surgery) are exploring the use of SDC for standardizing data entry, delivering standardized clinical reports, and facilitating downstream data usages. The eCC program is described in another paper in this issue.<sup>10</sup>

### SDC HISTORY

The SDC project was initiated in early 2013 by the Office of the National Coordinator for Health Information Technology (ONC) through its Standards and Interoperability Framework initiative.<sup>14</sup> IHE was selected as the organization to host the specification. The IHE profile for SDC was first published in October 2016, is maintained by the IHE SDC Working Group, and is regularly tested at IHE Connectathons.<sup>15,16</sup>

The ONC also sponsored an attempt to harmonize FHIR Questionnaire with IHE SDC,<sup>17</sup> to produce a hybrid, functionally equivalent FHIR SDC model. However, complete harmonization was not achieved, and the two approaches diverged because of differences in objectives and

design principles. In 2017, both IHE SDC and FHIR SDC became community-led initiatives. This paper addresses only IHE SDC.

### SDC ARCHITECTURE

SDC is an information model that describes how various types of generic clinical information should be represented for technology-agnostic data capture. The primary information type addressed by SDC is the DE,<sup>18</sup> which includes question-answer sets and fill-in questions, although SDC can also handle standard media types such as images in questions and responses. Each question and answer has a unique identifier (ID), which remains constant unless the contextual semantics of the question or answer changes. To help represent context and control the display of form parts, SDC sections and DEs may be repeated and nested to any level of depth.

The structure of SDC is defined by a set of nested eXtensible Markup Language (XML) schemas. The schemas constrain the structure of SDC XML to recurring patterns and are also used to generate programming code to create the SDC Object Model (OM). The OM is used to generate SDC XML from SDC modeling tools and may also be used to control the behavior of SDC-based DEFs. Details about the SDC Schema set may be found in the SDC Technical Reference Guide.<sup>9,19</sup>

SDC XML documents (Fig 1) that are used to generate DEFs are called Form Design Files (FDFs). An FDF may be converted to a DEF using a variety of techniques. One popular technique is to use a program (often written in eXtensible Stylesheet Language with Transformations [XSLT]) to convert the FDF into a functional web page, with JavaScript controllers to implement SDC rules and data submission functionality. However, most vendors who support SDC do not use webpages, but instead use proprietary techniques to transform the FDF into their preferred software implementation.

### Common Data Elements and Terminologies

SDC can also define common data elements (CDEs). CDEs are DEs that are common across multiple data sets and/or are shared across clinical domains and/or reused in many

```

<Question ID="53672.100004300" title="Histologic Type">
  <ListField maxSelections="0">
    <List>
      <ListItem ID="56746.100004300" title="Acinar adenocarcinoma"
        selected="true" />
      <ListItem ID="55730.100004300" title="Ductal adenocarcinoma" />
      <ListItem ID="50277.100004300" title="Small-cell neuroendocrine carcinoma"
        selected="true"/>
    </List>
  </ListField>
</Question>

```

**Histologic Type**

Acinar adenocarcinoma

Ductal adenocarcinoma

Small-cell neuroendocrine carcinoma

**FIG 1.** SDC data element. The figure shows an example of an SDC XML data element corresponding to a multiselect question, with the matching part of an eXtensible Stylesheet Language with Transformation-generated HTML DEF shown in the inset (lower right). Each of the 3 answer choices in the DEF inset matches to a ListItem in SDC XML. ListItem elements are nested under the Question element with wrapper elements in the sequence Question → ListField → List → {ListItems}, where {ListItems} represents the list of ListItem elements. As shown in the Histologic Type DEF (inset), a user has selected the first and third answer choices. In the SDC XML, a selected answer choice is expressed with the selected = “true” expression on the corresponding ListItem elements. The expression maxSelections = “0” on the ListField element indicates that the Question is multiselect. Note that each Question and ListItem has a unique ID attribute. The .100004300 part of each ID is the namespace designation for the College of American Pathologists. DEF, data entry form; ID, identifier; SDC, Structured Data Capture; XML, eXtensible Markup language; XSLT, eXtensible Stylesheet Language with Transformations. (From NAACCR Volume V,<sup>27</sup> with slight modification. Used with permission from NAACCR.)

different FDFs.<sup>20,21</sup> SDC’s use of CDEs provides an important layer of data interoperability and reuse by pre-defining sharable DEs that are needed for the FDF clinical content.<sup>22</sup>

Ideally, CDEs should be paired with appropriate standard terminologies to optimize interoperability and encourage CDE reuse.<sup>23</sup> Terminology standards are critical to provide the semantic meaning and context of CDE components when CDEs are separated from their SDC source, when used by analysts who may not have access to the SDC or CDE definition, or when combining with data sets from non-SDC and/or non-CDE sources. Similarly, SDC DEs also benefit from being mapped to standard terminologies.

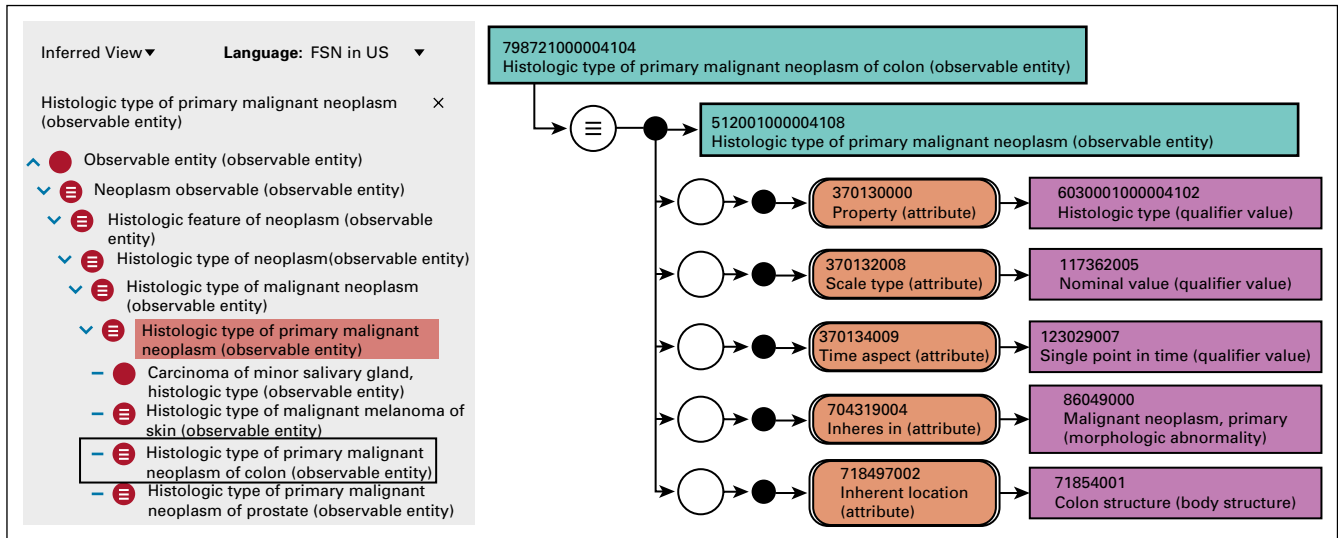
The SDC content management workflow is improved by using ancillary SDC mapping files for CDEs and terminologies, rather than placing CDE and terminology metadata directly in FDFs. External FDF maps promote centralized mechanisms for terminology management, validation, distribution, and searching for new and updated code sets, and they also enable the transmission of smaller SDC messages. However, some use cases may require the transmission of terminology codes within the FDF, and SDC supports this model as well.

### SDC + Systematized Nomenclature of Medicine Clinical Terms

Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) is the most comprehensive controlled medical terminology and is broadly adopted internationally. SNOMED CT is polyhierarchical, allowing multiple parent

nodes per clinical concept. It is composed of 19 domain hierarchies, such as body structure, observable entity, and clinical finding. Each concept may be defined (rendered computable) by specifying supertype(s) and additional defining attributes from the various domain hierarchies. Defined concepts are subjected to computer classification, which moves each concept under its logical parent concepts and creates additional logical concept relationships. The result is a robust searchable ontology that allows for granular, specific concept definitions, concept aggregations, and concept grouping by defining characteristics.<sup>24</sup> In 2014, investigators at the University of Nebraska Medical Center began development of SNOMED CT concepts specific to the eCC SDC content to address terminology deficiencies noted by the Centers for Disease Control and Prevention (CDC) for cancer reporting.<sup>25,26</sup> An example of the new SNOMED CT modeling for the eCC SDC templates is provided in [Figure 2](#).

SDC IDs for answer choices change whenever the answer choice semantics change. For SDC questions, the SDC ID changes whenever any change is made to the semantics of the question or any of its child answer choices. This provides clear documentation whenever the DE’s composite semantics changes. However, this level of semantic version control in SDC can be undesirable when stable IDs are desired for querying across DE versions, where SDC IDs are used as query targets. SNOMED CT, when mapped to each SDC question and answer, solves this problem by providing stable semantic IDs for each SDC ID. Additionally, the SNOMED CT ontology provides new opportunities for analytics such as increasing or decreasing the granularity of



**FIG 2.** SNOMED CT modeling designed for use with SDC and common data element–based data analysis. The figure shows the SNOMED CT concept for *Histologic type of primary malignant neoplasm of colon*. The right portion of the figure shows the concept’s stated definition (ie, the definition provided by the concept author), which indicates that it is an observation of the histology type of a primary malignant neoplasm located in the colon made at a single point in time. The left portion of the figure, inside the grey rectangle, represents the classified concept definition (ie, the augmented definition produced from the SNOMED CT description logic classifier), which asserts that the observation is a subtype of observation of histology of primary malignant neoplasm and several other higher-level concepts. Furthermore, the concept is grouped, or aggregated, with all other types of observations of histologic types of primary malignancies regardless of organ, such as prostate and melanoma. This classified definition supports data queries such as “find any instance of adenocarcinoma in any organ” or “find all histologic types associated with primary colon tumors.” SDC, Structured Data Capture; SNOMED CT, Systematized Nomenclature of Medicine Clinical Terms.

queries through drilldowns and rollups, which would be impossible with SDC IDs alone. Finally, SNOMED CT provides a robust analytics capability that can survive minor DE version changes that alter SDC IDs.

### SDC Data Transmission

When an SDC DEF is filled out, the user’s responses may be stored inside the FDF XML, which is now called an FDF with Responses (FDF-R). The FDF-R may undergo cycles of edit-save-edit revisions before being transmitted (using standard IHE transactions) to one or more end points, such as EHRs and public health agencies. Alternatively, responses may be extracted from the DEF or FDF-R and transmitted in any suitable format, such as North American Association of Central Cancer Registries (NAACCR) Volume V,<sup>27</sup> which uses Health Level Seven International (HL7) v.2.5.1, or IHE SDC on FHIR<sup>28</sup> (discussed below). While recreating the transmitted DEF at the end point node is a trivial task if the FDF-R is transmitted intact, the question and answer responses can be extracted and reconstructed into an SDC DEF after using any of the above transmission techniques.

### SDC on FHIR

The IHE SDC working group is developing a transmission specification for IHE SDC using FHIR as the wrapper or transport mechanism for SDC forms.<sup>28</sup> This approach wraps or converts SDC forms to a variety of FHIR resources.

A FHIR resource is a reusable data structure that represents a small domain of healthcare information. Examples include patient, practitioner, claim, and location. IHE SDC on FHIR uses the resources named DocumentReference and Observation.<sup>29,30</sup> This approach provides seamless interoperability between FHIR and IHE SDC. SDC forms and data are transported in a FHIR DocumentReference wrapper, and the FDF-R question and answer content is parsed into individual FHIR Observation objects.<sup>31</sup> The SDC Observations can be processed and queried like any other FHIR data, expanding the downstream usability of the SDC data.

SDC DEs and FHIR Observations both support repurposing of SDC data for reuse in other types of data sets, eg, biospecimen annotations, clinical trial forms, reports, and rules engines. The SDC IDs and mapped terminology codes allow downstream systems to reconstruct the semantic, contextual, and structural aspects of the DEs, and if required, to trace back to the SDC form where the data originated.

### SDC ADOPTION IN CANCER PATHOLOGY

The standardization of SDC features across implementers allows accreditation organizations to support their requirements through interoperable, metadata-driven content and behavior. The extent of SDC adoption can be gauged by the number of licensed users of CAP eCCs. Currently, 45% of hospitals with > 400 beds in the United

States are licensed to use the eCC.<sup>10</sup> In addition, 92% of Ontario, Canada pathologists were using the eCCs as of 2012, and according to the Cancer Care Ontario website, 100% of Ontario pathologists are currently using the eCCs, now released only in SDC format.<sup>3,32,33</sup>

## SDC RESEARCH AND DEVELOPMENT

### SDC-Based Breast Cancer Staging Calculator

One important example of new feature testing involves the implementation of an SDC-based Breast Cancer Staging Calculator (BCSC) that uses the American Joint Committee on Cancer (AJCC) staging application programming interface. New features piloted in the BCSC reference implementation included a more advanced use of skip logic (turning DEF parts of/off depending on the user type [pathologist or oncologist]) the use of surrogate codes in a format required by the staging web service, the aggregation of parameter values from selected answers and user-entered values, the sending of those parameter values to a staging web service (created by AJCC and CDC), return of the values to designated parts of the SDC DEF, generation of a full synoptic report on the basis of the user responses and values returned from the staging web service, and also transmission of that report to a CDC server using the IHE SDC SubmitForm transaction.<sup>9</sup> These features are specified declaratively inside the FDF XML, without any procedural code. In the BCSC reference implementation, small JavaScript services were used to read the FDF XML metadata and implement the above behavior when DEF buttons were pressed and when the form results were submitted to the CDC server. This pilot served as a demonstration of a multipart SDC form that is used by three different physicians in sequence to produce an integrated staging report for automating both clinical and pathological AJCC staging. Introducing these features for vendor implementation would likely require 1-2 years of additional work after the project plan is approved by the various stakeholders.

### Computable Care Guidelines

The Computable Care Guideline (CCG) technical framework reinterprets written guidelines as interoperable computer operations.<sup>34,35</sup> The technical framework is based on the FHIR Clinical Practice Guidelines Implementation Guide by the HL7 Clinical Reasoning Work Group.<sup>36</sup> In a CCG, data are collected by SDC form components, which are used to trigger FHIR-based rule blocks called Cards, on the basis of FHIR ActivityDefinition.<sup>37</sup> Cards are connected to each other using SDC-derived responses and mapped terminology codes transmitted as FHIR transactions.

For example, a cancer diagnosis or staging guideline can be converted to a set of cascading SDC forms and cards that communicate with an EHR system. As clinical results from the SDC form are saved into the patient's health record, card instructions will present EHR notifications to appropriate members of the care team with the next steps for their

patient.<sup>38,39</sup> The DEs inside the SDC forms were mapped to terminology codes that enabled coordination between the DEFs, cards, and the EHR in the demonstration.

### Computer-Assisted Reporting and Decision Support (CAR/DS)

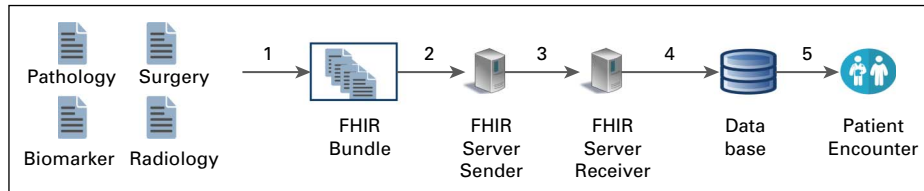
The CAR/DS framework (no relationship to CCG Cards) is a machine-readable XML-based definition format for representing radiology reporting clinical guidelines created by the American College of Radiologists.<sup>40</sup> Like SDC, CAR/DS is also designed to be an open framework for the creation of additional guidelines. The CAR/DS and SDC groups are developing a pilot to correlate radiology (CAR/DS) and pathology (SDC) cancer diagnoses through the use of FHIR transactions.<sup>41</sup> Because CAR/DS and SDC both can capture data at the point of care and then convert them to FHIR Observations, they can be made interoperable through FHIR technology. By mapping common DEs across CAR/DS and SDC templates, radiologists will be able to automate the creation of concordance reports that compare the radiologic diagnosis with the pathologic diagnosis. Concordance reports help radiologists determine the accuracy of their radiographic assessments. A pilot project between American College of Radiologists and CAP is applying this approach to the Thyroid Imaging Reporting & Data System (TI-RADS) and the Thyroid eCC SDC template.

### Registry Data Dictionaries and Form Templates

NAACCR maintains an extensive composite data dictionary, known as NAACCR Volume II, for cancer registries.<sup>42,43</sup> This data dictionary is complex, containing generic DEs applicable to all tumor types and Site-Specific Data Items applicable only to certain tumors.<sup>44</sup> Frequent NAACCR volume II changes are a challenge for the various kinds of software that must be updated and the systems that must analyze changing data sets. Furthermore, NAACCR volume II must maintain compatibility with other cancer data standards such as the eCCs, AJCC staging, and the International Classification of Diseases for Oncology, third edition (ICD-O-3) cancer classification system. The CAP and the CDC are exploring ways to represent the NAACCR Volume II DEs in SDC format, on a tumor-specific bases, with one FDF per tumor type. The tumor pathology would be drawn largely from existing eCC CDEs, which are already harmonized with AJCC staging, ICD-O-3 and other clinical standards. In this pilot model, each FDF would serve as both a data dictionary and as a DEF template for registry software vendors, potentially alleviating some of the NAACCR versioning challenges.

### Minimal Common Oncology Data Element

The minimal Common Oncology Data Elements (mCODE) project curates a standardized set of oncology data represented as FHIR resources. mCODE resources are designed to promote data transfer between EHRs and other systems such as registries and clinical trials software. mCODE began as a collaboration between ASCO and the



**FIG 3.** SDC on FHIR. The figure shows submission of several SDC forms with SDC on FHIR, using FHIR DocumentReference and Observation in a Bundle. Arrow 1 shows a group of SDC forms being processed for inclusion in an FHIR bundle. Arrow 2 shows the use of SDC on FHIR to submit the FHIR bundle from a sending server to a receiving server. Arrow 3 shows submission of the SDC on FHIR bundle to a receiver. Arrow 4 shows submission to a database end point. Arrow 5 shows the extraction, transformation, and transfer of that data to permit viewing by end users. SDC, Structured Data Capture.

MITRE Corporation.<sup>45</sup> mCODE data may be collected using a variety of tools, including SDC-based DEFs. Current work is exploring ways to transmit SDC-collected data as mCODE resources. Future work may include mapping from specific SDC DEs to mCODE FHIR resources. Additionally, mCODE DEs may be combined with SDC data using FHIR bundles to create a more comprehensive report.

### Other Ongoing SDC Work

Additional ongoing SDC work involves testing of SDC schema enhancements, improving documentation, and piloting new approaches to data bundling, transmission, and downstream use. Examples include pilots using SDC + FHIR bundles containing arbitrarily complex SDC and FHIR content,<sup>43</sup> tools to extract and display the transmitted data, bulk data transmissions to facilitate large data sets in public health use cases (Fig 3),<sup>46,47</sup> sharing patient data with multiple members of a distributed care team in diverse locations, and automated trigger-based clinical decision support.

### AFFILIATIONS

<sup>1</sup>Cancer Protocols and Data Standards, College of American Pathologists, Northfield, IL

<sup>2</sup>Department of Pathology/Microbiology, University of Nebraska Medical Center, Omaha, NE

### CORRESPONDING AUTHOR

Richard Moldwin, MD, PhD, Cancer Protocols and Data Standards, College of American Pathologists, 325 Waukegan Rd, Northfield, IL 60093; e-mail: rmoldwin@cap.org.

### SUPPORT

Supported by a CDC Cooperative Agreement (5 NU58DP006457-03-00).

### AUTHOR CONTRIBUTIONS

**Conception and design:** All authors

**Collection and assembly of data:** All authors

**Data analysis and interpretation:** All authors

**Manuscript writing:** All authors

**Final approval of manuscript:** All authors

**Accountable for all aspects of the work:** All authors

In conclusion, IHE SDC is a technology designed to meet both data capture and interoperability needs. Complex, frequently changing, oncology data sets were a major design focus. SDC synergizes with technologies such as FHIR and SNOMED CT to greatly increase the capabilities of standardized structured reporting. Some capabilities include improving guideline adherence, clinical decision support, public health surveillance, and data aggregation for research and clinical trials.

Although most vendors who use SDC opt for desktop application DEF technology, SDC-enabled DEF software is also straightforward to implement using standard web technologies. An open-source, web-based SDC reference implementation is available on the IHE-SDC-WG GitHub.<sup>48,49</sup> Because of its flexible and interoperable features, SDC could play an important role in data capture for healthcare by creating centrally standardized DEFs for diseases like cancer and could be leveraged for emerging diseases such as COVID-19.

### AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to [www.asco.org/rwc](http://www.asco.org/rwc) or [ascopubs.org/cci/author-center](http://ascopubs.org/cci/author-center).

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](http://Open Payments)).

#### Walter Scott Campbell

**Travel, Accommodations, Expenses:** GenomOncology

No other potential conflicts of interest were reported.

### ACKNOWLEDGMENT

The authors are grateful to Eric Daley, MS, PA (ASCP)CM for helpful comments and review of the manuscript.

## REFERENCES

1. HIMSS: What is Interoperability? HIMSS, 2016. <https://www.himss.org/library/interoperability-standards/what-is-interoperability>
2. Sluijter CE, van Lonkhuijzen LRCW, van Slooten HJ, et al: The effects of implementing synoptic pathology reporting in cancer diagnosis: A systematic review. *Virchows Arch* 468:639-649, 2016
3. Srigley J, Lankshear S, Brierley J, et al: Closing the quality loop: Facilitating improvement in oncology practice through timely access to clinical performance indicators. *J Oncol Pract* 9:e255-e261, 2013
4. Donahoe L, Bennett S, Temple W, et al: Completeness of dictated operative reports in breast cancer—the case for synoptic reporting. *J Surg Oncol* 106:79-83, 2012
5. Temple WJ, Chin-Lenn L, Mack LA: Evaluating population-based breast cancer surgical practice in real time with a web-based synoptic operative reporting system. *Am J Surg* 207:693-697, 2014
6. den Broeck JV, Cunningham SA, Eeckels R, et al: Data cleaning: Detecting, diagnosing, and editing data abnormalities. *PLOS Med* 2:e267, 2005
7. Windle T, McClay J, Windle J: The impact of domain knowledge on structured data collection and templated note design. *Appl Clin Inform* 04:317-330, 2013
8. Staes C: Solor—The Simple Healthcare Terminology Solution. 2020. [http://solor.io/wp-content/uploads/2017/12/White-paper\\_Achieving-semantic-data-interoperability.pdf](http://solor.io/wp-content/uploads/2017/12/White-paper_Achieving-semantic-data-interoperability.pdf)
9. IHE Quality Research and Public Health: Structured Data Capture Technical Framework. 2016. [http://ihe.net/uploadedFiles/Documents/QRPH/IHE\\_QRPH\\_Suppl\\_SDC.pdf](http://ihe.net/uploadedFiles/Documents/QRPH/IHE_QRPH_Suppl_SDC.pdf)
10. Torous VF, Simpson RW, Balani JP, et al: The College of American Pathologists Cancer Protocols: From optimizing cancer patient care to facilitating interoperable reporting and downstream data use. *JCO Clin Cancer Informatics* 5:47-55, 2021
11. Simpson RW, Berman MA, Foulis PR, et al: Cancer biomarkers: The role of structured data reporting. *Arch Pathol Lab Med* 139:587-593, 2014
12. Rollison DE, Howlander N, Smith MT, et al: Epidemiology of myelodysplastic syndromes and chronic myeloproliferative disorders in the United States, 2001-2004, using data from the NAACCR and SEER programs. *Blood* 112:45-52, 2008
13. North American Association of Central Cancer Registries: Pathology Laboratory Electronic Reporting, Volume V. 2020. <https://www.naacr.org/pathology-laboratory-electronic-reporting/>
14. The Office of the National Coordinator for Health Information Technology: Standards & Interoperability (S&I) Framework | HealthIT.gov. Off Natl Coord Health Inf Technol, 2019. <https://www.healthit.gov/topic/interoperability/standards-interoperability-si-framework>
15. IHE: IHE Connectathon Overview | IHE USA. IHE. 2019. <https://www.iheusa.org/ihe-connectathon-overview>
16. O'Donnell K: Radiation exposure monitoring: A new IHE profile. *Pediatr Radiol* 41:588, 2011
17. HL7: FHIR Structured Data Capture. HL7 FHIR, 2020. <https://build.fhir.org/ig/HL7/sdc/>
18. International Organization for Standardization: ISO/IEC 11179-7:2019 Information technology—Metadata Registries (MDR)—Part 7: Metamodel for Data Set Registration. ISO, 2019. <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/87/68766.html>
19. Moldwin R (ed): IHE Structured Data Capture (SDC) Technical Reference Guide (TRG). 2020. [https://github.com/IHE-SDC-WG/SDC-Schema-Packages/blob/master/Documentation/CAP%20SDC%20Technical%20Reference%20Guide\\_v0.42%20\(Feb%202020\).pdf](https://github.com/IHE-SDC-WG/SDC-Schema-Packages/blob/master/Documentation/CAP%20SDC%20Technical%20Reference%20Guide_v0.42%20(Feb%202020).pdf)
20. Crichton C, Davies J, Gibbons J, et al: Metadata-driven software for clinical trials, in 2009 ICSE Workshop on Software Engineering in Health Care, Vancouver, BC, Canada, IEEE, 2009. <https://ieeexplore.ieee.org/document/5069600>
21. Hruby GW, Hoxha J, Ravichandran PC, et al: A data-driven concept schema for defining clinical research data needs. *Int J Med Inf* 91:1-9, 2016
22. Sinaci AA, Laleci Erturkmen GB: A federated semantic metadata registry framework for enabling interoperability across clinical research and care domains. *J Biomed Inform* 46:784-794, 2013
23. Sheehan J, Hirschfeld S, Foster E, et al: Improving the value of clinical research through the use of Common Data Elements. *Clin Trials Lond Engl* 13:671-676, 2016
24. Campbell WS, Pedersen J, McClay JC, et al: An alternative database approach for management of SNOMED CT and improved patient data queries. *J Biomed Inform* 57:350-357, 2015
25. Campbell JR, Talmon G, Cushman-Vokoun A, et al: An extended SNOMED CT concept model for observations in molecular genetics. *AMIA Annu Symp Proc* 2016:352-360, 2017
26. Campbell WS, Karlsson D, Vreeman DJ, et al: A computable pathology report for precision medicine: Extending an observables ontology unifying SNOMED CT and LOINC. *J Am Med Inform Assoc* 25:259-266, 2018
27. Jones S, Mazuyrk J: Standards for Cancer Registries Volume V: Pathology Laboratory Electronic Reporting, Version 5. 2020. <https://www.naacr.org/pathology-laboratory-electronic-reporting/>
28. Canada Health Infoway, Ontario Health (Cancer Care Ontario): IHE SDC on FHIR. SIMPLIFIER.NET. 2019. <https://simplifier.net/sdconfhir>
29. HL7: DocumentReference—FHIR v4.0.1. HL7 FHIR, 2019. <https://www.hl7.org/fhir/documentreference.html>
30. HL7: Observation—FHIR v4.0.1. HL7 FHIR, 2019. <https://www.hl7.org/fhir/observation.html>
31. Francis J: IHE-SDC-WG/IHE-SDC-FHIR-Parser. IHE Structured Data Capture Working Group, 2020. <https://github.com/IHE-SDC-WG/IHE-SDC-FHIR-Parser>
32. Canadian Partnership Against Cancer: Evaluation of the Electronic Synoptic Pathology Reporting Initiative. *Can Partnersh Cancer*, 2020. <https://www.partnershipagaincancer.ca/topics/synoptic-pathology-reporting/>
33. Cancer Care Ontario: Synoptic Pathology Reporting. *Cancer Care Ont*, 2017. <https://www.cancercareontario.ca/en/guidelines-advice/treatment-modality/pathology-laboratory-testing/synoptic-pathology-reporting>
34. IHE Quality Research and Public Health: Computable Care Guidelines—IHE Wiki. 2020. [https://wiki.ihe.net/index.php/Computable\\_Care\\_Guidelines](https://wiki.ihe.net/index.php/Computable_Care_Guidelines)
35. Shah H, Allard RD, Enberg R, et al: Requirements for guidelines systems: Implementation challenges and lessons from existing software-engineering efforts. *BMC Med Inform Decis Mak* 12:16, 2012
36. HL7: FHIR Clinical Guidelines. HL7 FHIR, 2020. <http://build.fhir.org/ig/HL7/cqf-recommendations/index.html>
37. HL7: ActivityDefinition—FHIR v4.0.1. 2019. <https://www.hl7.org/fhir/activitydefinition.html>
38. Boxwala AA, Rocha BH, Maviglia S, et al: A multi-layered framework for disseminating knowledge for computer-based decision support. *J Am Med Inform Assoc* 18:i132-i139, 2011
39. Francis J, Goel AK, Ibrahim M, et al: Lung Diagnosis Computable Care Guideline, in *Computable Care Guidelines*, Cleveland, OH, IHE Connectathon, 2020. <https://simplifier.net/guide/LungCancerDiagnosisComputableCareGuideline/Demo>
40. Alkasab TK, Bizzo BC, Berland LL, et al: Creation of an open framework for point-of-care computer-assisted reporting and decision support tools for radiologists. *J Am Coll Radiol* 14:1184-1189, 2017

41. Kuo MD, Yamamoto S: Next generation radiologic-pathologic correlation in oncology: Rad-path 2.0. *Am J Roentgenol* 197:990-997, 2011
42. Thornton M: Data Standards & Data Dictionary, Volume II. 2018. <https://www.naaccr.org/data-standards-data-dictionary/>
43. North American Association of Central Cancer Registries: Steering Committees and Subgroups. NAACCR, 2018. <https://www.naaccr.org/steering-committees/>
44. Ruhl J, Ward E, Hofferkamp J: Site-Specific Data Item (SSDI) Manual. 2019. <https://www.naaccr.org/SSDI/SSDI-Manual.pdf?v=1596228380>
45. mCODEI Minimal Common Oncology Data Elements. 2020. <https://mcodeinitiative.org/>
46. HL7: Bundle—FHIR v4.0.1. 2020. <https://www.hl7.org/fhir/bundle.html>
47. HL7, Boston Children's Hospital: FHIR Bulk Data Access (Flat FHIR). 2020. <https://hl7.org/fhir/uv/bulkdata/>
48. IHE-SDC-WG/SDC-Conformance-Tool. IHE Structured Data Capture Working Group, 2017. <https://github.com/IHE-SDC-WG/SDC-Conformance-Tool->
49. IHE-SDC-WG/SDC-Conformance-Tool. IHE Structured Data Capture Working Group, 2017. <https://sdcapp.azurewebsites.net/SDCApp/>

