SMART C-CDA Collaborative
Perspective & Debriefing

We review the purpose and results of the SMART C-CDA Collaborative's vendor outreach effort in this document. The Collaborative was conducted jointly by the SMART Platforms Project (www.smartplatforms.org) and the Lantana Consulting Group.

Our outreach effort involved data collection, group discussions, and individual sessions with EHR and HIT vendors, consisting of these stages:

1. Efforts were first made to reach out to a wide array of EHR and HIT vendors to provide samples of C-CDA documents for project analysis. This included individual invitations to 107 different organizations. In total, 44 organizations responded to the invitation with 31 organizations participating in Collaborative discussions. 21 organizations submitted documents to the Collaborative, although not all were in C-CDA XML format. *16 organizations submitted C-CDA documents bolstered by 6 C-CDA non-overlapping samples collected from publicly available resources*.

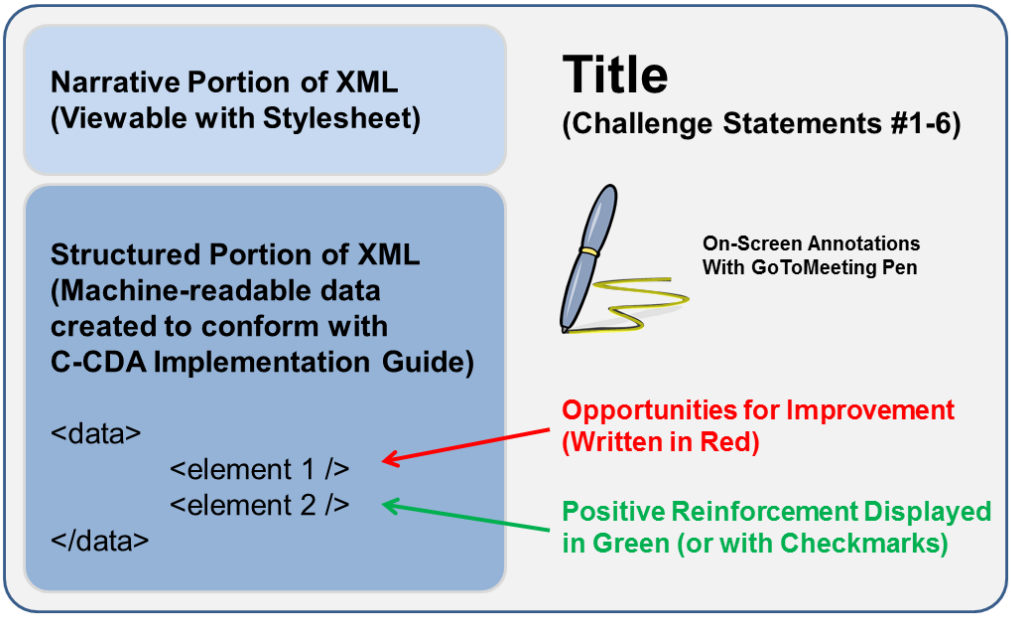
2. Group discussions were hosted as weekly web/teleconference calls from late July through early October 2013 to review observations made in the domain analysis and discuss the largest challenges faced in C-CDA implementation. 18 organizations participated in at least one of the weekly calls which focused on 7 key domains: patient demographics, problems, medications, allergies, results, vital signs, and smoking status. Feedback from these calls help inform the C-CDA Implementation Guide R2, which was submitted for September 2013 balloting by HL7. A further outcome of these discussions was a list of 6 C-CDA challenge statements. These 6 challenges became the basis for 1-on-1 sessions where the SMART and Lantana teams reviewed C-CDA findings with interoperability experts at participating vendor organization.

3. *10 organizations participated in 1-on-1 sessions to review their submitted C-CDA sample(s) in depth*. 1-on-1 Sessions lasted 60-90 minutes and included a customized walkthrough of that vendor's C-CDA document, which was also shared at the conclusion of the session. Three EHR vendors further elected to participate in a second discussion after this initial review. The explicit aim was for SMART and Lantana to illustrate interoperability issues and encourage participating vendors to improve C-CDA document quality based upon their customized review.

We recount themes, observations, and discussion points raised from the 1-on-1 Sessions, each of which followed a regular format: we evaluated a vendor's C-CDA submission in relation to **6 challenge statements** (see Appendix). The challenges were formulated during the initial stage

of the SMART C-CDA Collaborative, during which the then available vendor C-CDA files were reviewed by Lantana personnel, who identified common problems and/or heterogeneity in data representation in those segments of the C-CDA that were analyzed by the SMART C-CDA Scorecard (or ingested by the SMART C-CDA Receiver).

For each challenge, the format for discussion was to present a single slide (illustrated below). On the left side of the slide, the XML in the vendor's C-CDA under review would be split into the narrative and structured areas (depicted below in blue). Our comments would be in color, with red used to designate observations identified for improvement. Webinar tools were used to annotate "live" the screen and guide discussion. At the end of each session, we would review the SMART C-CDA Scorecard result for that C-CDA, including rerunning it live in the Scorecard to permit a review of precise scoring for each of the covered C-CDA segments.



Data Context

Vendors in the SMART C-CDA Collaborative submitted C-CDA samples conformant to Meaningful Use Stage 2 with data in 7 key domains: demographics, problems, allergies, medications, results, vital signs, and smoking status. As of December 2, 2013, a total of 6 vendors had elected to place their C-CDA into a public library so that all industry participants can make use of them in designing for interoperability.

General Observations

- **Goals in C-CDA Development.** A common goal expressed by multiple vendors was that C-CDA development was primarily motivated to pass validation, as part of Meaningful Use EHR Certification. A second goal was to demonstrate lightweight interoperability at vendor connect-a-thons. Multiple participants acknowledged that connect-a-thons do not test deep content or semantic interoperability as addressed in these 1-on-1 Sessions. In respect of real-world use, moreover, providers have not yet engaged in active, routine C-CDA exchange and consequently have not engaged their vendors on actual issues with respect to C-CDA document transport. Meaningful Use Certification as tested by NIST Transport Testing Tool (TTT) is therefore the focus of C-CDA related product development. While the upside of TTT is that it pushed adoption of key terminologies and structure from the C-CDA Implementation Guide 1.1, several downsides, detailed below, came to light in 1-on-1 Sessions. A general issued raised by vendors is that C-CDA configurations may be affected by provider (Eligible Physicians/Hospitals) database set-up, potentially compromising certified compliance demonstrated in vendor development environments. Nonetheless, despite the controlled conditions for C-CDA certification in vendor development environments, every sample we examined surfaced opportunities for improvement, which were the basis for 1-on-1 Session discussions.

- **Learning & Surprise.** The Lantana and SMART Platforms team members have significant expertise in C-CDA formatting; thus, most vendors acknowledged that they learned things their respective 1-on-1 Session, as evidenced from many questions on key points of our analysis. Vendors rarely argued for a different interpretation of the C-CDA standard for the XML segments under review, though some expressed surprise at certain issues identified in samples reviewed. The majority of identified issues fell into four categories:
    - “We did that to pass validation, even though we knew it may not be correct.”
    - “That’s a bug we did not know about or did not realize it was incorrect.”
    - “That’s a bug we know about and are working on.”
    - “That’s a hard issue since the data are not structured in our source system and therefore cannot be easily represented in structured XML.”

- **Introduction to SMART C-CDA Scorecard.** Even among those vendors which were aware of the SMART C-CDA Scorecard (which targets C-CDA document content defects rather than TTT’s schema/syntax defects) before joining the Collaborative effort, few had used it for their C-CDA development or testing. Nonetheless, upon having it presented and used live during the 1-on-1 Sessions, every vendor confirmed its utility. *Multiple vendors expressed interest in improving their performance results on the Scorecard. In fact, two vendors requested access to the Scorecard be provided as a web service in addition to web page access.* Of particular value was the Scorecard’s ability to

scan for wrong, missing, or aberrant codes used in a C-CDA document. For example:
    - The Scorecard identified when a vendor's lab results used a very rare hematology code (WBC via plateletpheresis) when a more likely and highly utilized code (WBC via blood) was appropriate.
    - The Scorecard detected when a vendor used the deprecated code for body mass index, while other vendors, in fact, used non-existent codes.

- **Positive Feedback & Concern.** *Vendors expressed gratitude for the SMART/Lantana time spent identifying areas in their C-CDA documents to improve and for the opportunity to learn about how other vendors were approaching the same challenges.The 1-on-1 Sessions were especially well received.* During one group call, several vendors who had already gone through their 1-on-1 Session, articulated its value and encouraged other vendors to schedule their own sessions, with three vendors scheduling their own second (follow-up) discussions to their first sessions. Many vendors said they would use the discussion and the presentation to help muster resources within their organization to improve their C-CDA document quality, although nearly every vendor expressed that this would require the coordination of multiple teams and time before any change could be made. Two vendors raised this procedural concern: *would it be acceptable to make changes to their C-CDA document in light of their respective 1-on-1 Sessions after their product had already been Meaningful Use Stage 2 Certified?*

Specific C-CDA Observations

Changes made for EHR Certification (NIST TTT Validator)

- **Excess Precision in effectiveTime.** At least two vendors said they added some "excess" precision to effectiveTime information to comply with certification. "Excess" means that the level of precision expressed was not necessarily known. For example, only the day and month was in the source system, but the time was encoded as "20131022000000+0500" which explicitly means the event was known to have occurred at the stroke of midnight on October 22, 2013. Detecting excess timing precision represented as trailing zeroes was a feature added to the SMART C-CDA Scorecard for examination over the summer.
- **Exclusive Use of Physical Quantity in Result Observations.** One vendor reported that the NIST validator would only allow for physical quantities to be expressed for lab results, which is problematic when nominal values (e.g. "yellow" or "negative") are the logical results of certain lab test (e.g. "urine color" or "microbiology culture" respectively).

Incorrect Use of Vocabularies / Terminologies

- **LOINC.** The 1-on-1 Session and the SMART C-CDA Scorecard revealed numerous instances from multiple vendors where aberrant, deprecated or non-existent were used. These issues were typically readily acknowledged by vendors as bugs or opportunities to improve LOINC mapping.
- **RxNorm.** The 1-on-1 Session and the SMART C-CDA Scorecard revealed instances when RxNorm codes were used at the wrong level of precision or non-valid codes were used. In one instance, an RxNorm code was used but the designated codeSystem was SNOMED. In another instance, "Unknown" was used for a code (i.e. should have used nullFlavor) for a prenatal vitamin that had not been mapped to RxNorm.
- **NDF-RT.** The 1-on-1 Sessions identified opportunities to communicate drug classes as allergies, often as a larger set than a specific medication. At least two EHRs expressed that they have not developed this capability since NDF-RT wasn't a required terminology of Meaningful Use.
- **SNOMED.** The 1-on-1 Sessions and the SMART C-CDA Scorecard identified instances when SNOMED codes were used with aberrant or non-existent codes. Most frequently, these issues were readily acknowledged by vendors as bugs or opportunities to improve SNOMED mappings.
- **NCI Thesaurus.** The 1-on-1 Session revealed a couple instances when the NCI thesaurus for medication form or route of administration were used inappropriately. For example, "PO" was used but is not the code in the NCI Thesaurus for oral route of administration, although the intent is clear since this is common shorthand to denote "by mouth."

Dealing with Unstructured Data

- **Difficulty in Medication Interval Timing.** To communicate a medication prescription, one needs to know the drug, the dose timing and the amount to be administered at each dose. Most vendors expressed the drug in RxNorm with ease, but expressing the dose (as doseQuantity) and the timing (as effectiveTime xsi:type=”PIVL_TS”) was more difficult. The majority did not successfully populate these data in structured XML. A primary reason for the difficulty here was in dealing with unstructured data. Several vendors stated their systems allowed for the user option for free text sig information or that structured information was not necessarily being made available in C-CDA production. As a consequence, representing “2 pills BID” or “two tablets twice every day” (sig synonyms) were hard to reconstruct into the structured body for medications.
- **Lack of UCUM Utilization for Units.** Many vendors expressed difficulty in adopting UCUM units, which are often required as part of C-CDAs when representing physical quantities (e.g. Systolic Blood Pressure is not just “146” it is “146 mm[Hg]”) UCUM units become critical in performing automated conversion of units when multiple ways to record data are both in common use (e.g. centimeters and inches are both frequently used for height). One concern for lab results was that the source system may not have expressed the lab value in UCUM. For example, if the source lab system said “THOU/MICROLITER”(not valid UCUM) translating this information to “10*3/uL” (valid UCUM) requires extensive work and configuration for local provider environments. Finally many vendors were using unit configurations which ignored case, which is problematic since the UCUM units used in the C-CDA are case-sensitive.
- **Unstructured Example of referenceRange.** While there are no conformance criteria for structuring reference ranges, in the 1-on-1 Session we identified that the vast majority of collected samples were including these data as unstructured text within the referenceRange element (e.g. “3.8 - 10.7 10*3/uL” as a text string). When asked, most vendors stated that their system stored this data as structured elements of low, high and units, but they mapped it to text in C-CDA creation since that is how they saw the example in the C-CDA Implementation Guide. The C-CDA IG now includes revisions to illustrate how to represent this data in structured form.
- **Unstructured Reaction for Allergies.** Great divergence in how vendors approached the documentation of allergic reactions. While all understood the requirement of using codes to represent allergies, the use of SNOMED to represent reactions was only performed in approximately half of the samples. Some vendors said this information was never or rarely recorded in a structured terminology by their system. Others said that is was always or nearly always recorded as such. Given the divergence of perspective and the importance of communicating allergies, this was pointed out to vendors when not formally encoded in SNOMED.

Incorrect Usage of nullFlavor

A frequently observed mistake in C-CDA documents is inappropriate utilization of nullFlavor. In some instances, a nullFlavor should have been specified when a code value was unknown. Instead, vendors utilized an invalid code or in one instance SNOMED code 261665006, which means “Unknown.” Commonly, the wrong nullFlavor was used, with many vendors over-utilizing nullFlavor = “UNK”. This nullFlavor is defined to be used when “a proper value is applicable, but is not known” in the C-CDA IG. Its use in sample documents revealed two mistakes. First, “UNK” was used when a translation or non-coded value was readily available. In these circumstances, a more appropriate nullFlavor would be “OTH”. Second, “UNK” was used when not applicable, such as effectiveTime/high for an active problem. A more appropriate nullFlavor may be “NI” or “NA”, although another common conventions omit such XML elements altogether. Other less frequent errors related to nullFlavor were also observed and discussed.

Commonly Omitted Elements

- **Medication Timing Interval.** Given the common unstructured format of this data, a majority of vendors are omitting medication timing information, even when known based on medication sig information. This is permitted in the C-CDA IG since it is only a SHOULD conformance requirement, but it omits a key part of of prescription information
- **doseQuantity.** Some vendors did not include doseQuantity on any medication or included the XML element but never populated with actual values. doseQuantity is only a SHOULD conformance requirement since rateQuantity would be used when no discrete dose is administered (e.g. for continuous IV infusion), but this omits a key part of of prescription information.
- **interpretationCode.** Many vendors included this on every lab and some included on vital signs as well. A few vendors omitted interpertationCode on any result observation or only included it on a minority of cases (i.e. when not normal). Some vendors also used this element in different ways (i.e. one used “A” for any abnormal results, where most used “L” for low and “H” for high when physical quantities were represented).
- **methodCode or targetSiteCode.** In some instances, it would be useful to know positioning data for vital signs, such as blood pressure measured while sitting, although this data was generally omitted in C-CDAs by most vendors.
- **referenceRange.** In a few cases, reference ranges were omitted from common lab results where they would be expected to be available
- **author.** A major change to guidance in the revised C-CDA IG was to increase the prominence of authoring information. This may be a required component for Meaningful Use since documenting source and last modified date are requirements in document extracts. Generally, 1-on-1 Sessions did not identify author omission as an issue, but it was occasionally discussed in relation to timing and clinical information reconciliation.

Not Understanding XML Element Definition

- **act/statusCode.** At least two vendors had difficulty in managing the act/statusCode for active problems and/or allergies. There is a clear relationship of this information to problem status such that it is currently checked in the SMART C-CDA Scorecard. In addition, this point of confusion was addressed in detail in the revisions to C-CDA IG.
- **act/effectiveTime.** Nearly all vendors showed confusion in how to populate the act/effectiveTime information within allergy and problem sections of the C-CDA. This point of confusion was addressed in detail in the revisions to C-CDA IG, which provide new explanations for the same intended data elements. Most were simply replicating the biological onset and resolution dates as recorded in the observation element.
- **doseQuantity.** Multiple vendors were not correctly using units within doseQuantity as it relates to the level of medication specified by the RxNorm code. Once explained, most vendors understood the relationship and the necessity of coordinating units in a UCUM conformant manner.
- **administrationUnitCode.** One vendor used administrationUnitCode in a understandable, albeit technically incorrect way to specify additional drug form information. This element is only needed in the rare instances where a medication is described as an entire package (e.g. bottle of eye drops) but the unit of administered is smaller (e.g. an individual eye drop).
- **Result observation/value.** There were multiple questions from vendors on how to appropriately use UCUM for units and also in how to represent results which did not come as a specific value (e.g. less than 0.01).
- **Allergy Code for Drug Allergy.** At least one vendor shared a sample when the allergy was presented at the level of a pill (e.g. loratadine 10mg oral tablet) rather than the logical ingredient level (e.g. loratadine). While not controlled via conformance criteria, ingredient level information for allergies would likely make decision support and drug check easier to implement.

Lack of Illustrative Examples

During 1-on-1 Sessions and afterwards, vendors asked questions about how to communicate a certain condition. Many expressed concern that not enough illustrative examples existed in the public domain on how to accomplish this. As part of the SMART C-CDA Collaborative, a list of potential annotated examples has been created (see Trello), although the XML for these samples has not been created. In addition, the revised C-CDA Implementation Guide markedly increased the scope and complexity of examples included in the balloted document. HL7 has also brought together a volunteer group to assist in drafting examples, although no examples had been formally approved as of December 1, 2013.

- **First Documented vs. Last Modified Timing.** For example, when a problem was last modified vs. when first documented as a concern.
- **Value of "Less Than" or "Greater Than" for Lab Result.** For example, reported lab value of less 0.01. This would be reported when outside range limits.
- **Microbiology Results.** Many vendors thought this would be a helpful example since it is a complicated relationship of multiple components. One vendor did attempt to structure in XML, although most others did not believe they could accomplish this in current released technology.
- **Medication with Change in Dose Frequency.** For example a medication which requires two pills to be consumed on first dose only.
- **Medication with Continuous Administration** For example a continuous medication infusion with a participant of normal saline.
- **Coordination of Tobacco Use and Smoking Status.** For example how to document a specific amount of cigarette consumption and coordinate between two separate templated in social history to coordinate this information.

APPENDIX: Six C-CDA Challenges

The challenges are re-produced from the groupware site used for the project (Trello).

1. Coordinating Problem Status and Timing Information

Description

Problem status and relevant author information are key meta-data around problems. Many technologies currently demonstrate variation in how they encode problem status and timing. The intent of `act/statusCode` is to represent whether the concern is being followed and the use of `observation/effectiveTime/high` represents the resolution date of a problem. This challenge is to create two problem entries which accurately use these XML elements.

Challenge

1.
Create an active problem that shows:
a. `act/statusCode` = "active"
b. `act/effectiveTime/low` = when the problem was authored, precise to the minute or second
c. `observation/effectiveTime/low` = biological start date of problem, precise to the day
d. `observation/effectiveTime/high` = not shown OR nullFlavor of "NI" or "NA"

2.
Create a resolved problem which is no longer a concern that shows
a. `act/statusCode` = "completed"
b. `act/effectiveTime/low` = when the problem was authored, precise to the minute or second
c. `act/effectiveTime/high` = when the concern was no longer being followed, precise to the minute or second
d. `observation/effectiveTime/low` = biological start date of problem, precise to the day
e. `observation/effectiveTime/high` = biological resolution date of problem, precise to the day

Reach Goal

1. Add a qualifier to the problem within `observation/value` (not seen in any submitted sample). For example:

```
<value xsi:type="CD" code="233604007"
  codeSystem="2.16.840.1.113883.6.96"
  displayName="Pneumonia">
    <qualifier>
      <name code="363698007"
            displayName="Finding site"/>
      <value code="41224006"
             displayName="Left lower lobe of lung"
    </qualifier>
</value>
```

2. Coordinating Medical Interval, Dose and Code (Pre-Coordinated)

Description

Medication dose, timing and coding are key elements in patient care although they have several complex relationships. Interval timing should be structured (while it often is not) and dose needs to be related to both the interval and the code being used for pre-coordinated drugs. Many technologies have done this differently and this challenge is to create an example of a common antibiotic dose and timing.

Challenge

Create a single medication with the following information:

1.
5-day course, twice a day
a. Use `effectiveTime xsi:type="IVL_TS"` to show 5 days course, precise to the day
b. Use `effectiveTime xsi:type="PIVL_TS"` to show administered twice a day

2.
Use a pre-coordinated dose
a. Code = "309309" for "Ciprofloxacin 500 MG Oral Tablet"
b. Use the appropriate `doseQuantity@value` element to represent 1 pill at every dose.
c. Use the appropriate `doseQuantity@unit` to represent unitless (i.e. not shown) or “{tablet}” for UCUM denotation of unitless metric

Reach Goal

Add a partial dispensing history (# of pills + date + place where dispensed)

Edit the card description.

3. Creating Drug & Class Allergies with Coded Reactions

Description

Coding a drug or drug class allergy are essential to reduce adverse prescribing events and create clinical decision support / surveillance tools. There appears to be difficulty in some technologies to encode drug classes and allergic reaction in a normalized format with appropriate status. This challenge is to create two allergies for both a drug and its class.

Challenge

1.
Create a drug allergy to pencillin with coded reaction of “hives”
a. `act/statusCode` = “active”
b. `act/effectiveTime/low` = date when the allergy was documented
c. `observation/value@code` = "416098002" ("drug allergy")
d. RxNorm code of 7980, which is the highest form level for “penicillin”
e. `entryRelationship` for reaction observation (2.16.840.1.113883.10.20.22.4.9) with coded value of hives (SNOMED code = “126485001”)

2.
Create a drug class allergy to penicillins with coded reaction of “hives”
a. `act/StatusCode` = “active”
b. `act/effectiveTime/low` = date when the allergy was documented
c. `observation/value@code` = "416098002" ("drug allergy")
d. NDF-RT code of C18318
e. `entryRelationship` for reaction observation (2.16.840.1.113883.10.20.22.4.9) with coded value of hives (SNOMED code = “126485001”)

Edit the card description.

4. Creating a Highly Structured Lab Result

Description

For common lab results representing a physical quantity with a `referenceRange` an `interpretationCode`, a best practice is to represent all this data in a structured format within Result Observation. This challenge is to create a common lab with structured values that use UCUM and have accompanying information.

Challenge

Create a white blood cell count that uses appropriate values and a structured reference range:

1.
`observation/code` of white blood cell count (e.g. LOINC code = "26464-8")

2.
`observation/value` which represents an abnormal high reading:
a. `xsi:type="PQ"`
b. `value = "18.2"` (or any value above 11)
c. units from UCUM (e.g. 10*3/uL)
d. translation codes, if needed, to represent a varying source format (e.g. `<originalText>6.7 billion per liter</originalText>`)

3.
Inclusion of a `referenceRange`
a. That is optionally structured but at a minimum includes units. For example:

```
<referenceRange>
  <observationRange>
    <value xsi:type="IVL_PQ">
      <low value="4.0" unit="10*3/uL"/>
      <high value="11.0" unit="10*3/uL"/>
    </value>
  </observationRange>
</referenceRange>
```

Reach Goal

Represent a structured microbiology lab panel for urine culture + sensitivities.

Edit the card description.

5. Complete Vital Sign Set with BMI

Description

Vital signs are a key element of patient data. Some variation has been shown in what vital signs are encoded. For example, some technologies are not structuring body-mass index. In addition, many technologies are not using the appropriate UCUM units for vital signs. This challenge is to create a full set of vital signs that would be commonly recorded on an adult.

Challenge

Create a full set of vital signs that all structured values (type = "PQ") and use appropriate LOINC code and UCUM units:
1. Height
2. Weight
3. BMI (this can be logically derived from patient height & weight)
4. Respiratory Rate
5. Pulse Rate
6. Temperature
7. Systolic Blood Pressure
8. Diastolic Blood Pressure
9. Pulse Oximetry (optional)

Edit the card description.

6. Coordinating Current Smoking & Tobacco Use

Description

Smoking status is a key piece of information for quality measures and care planning. Representing smoking status, however is complex, especially in representing the current status vs. known history. Some significant changes have been suggested to the C-CDA R2 IG to represent smoking status. This challenge is to use those changes in a way which accurately communicates relevant data on a patient's smoking history.

Challenge

1.
Create a patient with current smoking status (2.16.840.1.113883.10.20.22.4.78) of "former smoker"
a. Use `observation/effectiveTime@value` to represent the author time of the observation, precise to minute or second
b. Use the appropriate vocabulary set for former smoker (SNOMED code = "8517006")

2.
Create a complementary tobacco use entry (2.16.840.1.113883.10.20.22.4.85) to give additional information on when the patient smoked:
a. Use `observation/effectiveTime/low` to represent the biological start date of smoking, precise to the year, month or day
b. Use `observation/effectiveTime/high` to represent the biological end date of smoking, precise to the year, month or day
c. Uses the appropriate vocabulary for when the patient was smoking (e.g. SNOMED code = "160604004" for 10-19 cigarettes per day)

Edit the card description.