

How Much Clinical Content Survives C-CDA to FHIR Conversion? Evaluating an Open- Source Converter with Paired Synthetic Records

Sai Sandeep Neeli

neelisandeep369@gmail.com

Research Article

Keywords: C-CDA, HL7 FHIR, interoperability, data conversion, US Core, synthetic data

Posted Date: October 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11501897/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Objective.

Health plans facing the 2027 CMS-0057-F API deadlines often need to convert C-CDA documents into FHIR. We measured how faithfully a widely available open-source converter preserves clinical content, and whether each loss comes from the source or the converter.

Methods.

Synthea generated 331 synthetic patients, each exported both as a C-CDA document and as a native FHIR R4 bundle from the same underlying record. We converted every document with python-fhir-converter 0.3.0 and compared the native, source and converted layers event by event across eight clinical families. We also checked field-level fidelity, structural validity and selected US Core 9.0.0 requirements.

Results.

All documents converted without error. Of 276,153 source entries, 96.9% survived conversion, but only 89.8% in the element where FHIR expects them. Every value and unit was preserved. Yet no timestamp kept its time of day, no medication or condition kept its status, and none of 8,901 condition end dates survived. 8,434 assessment scores were skipped. In 1,115 entries the original code was replaced by its translation, erasing eye laterality in 292 observations. Allergens (310 of 320) and all encounter types were moved to elements standard queries do not search, and 5.8% of converted resources failed validation, including every medication. A controlled test traced the time loss to missing UTC offsets in the source.

Conclusion.

Record counts suggest near-complete conversion while hiding losses that matter to FHIR clients and quality measurement. Paired synthetic records give a cheap, repeatable way to find and attribute such defects before production.

1. Introduction

Much of the clinical data that moves between U.S. organizations still travels as Consolidated Clinical Document Architecture (C-CDA) documents. At the same time, regulation is pushing that data toward HL7 FHIR. Under the CMS Interoperability and Prior Authorization Final Rule (CMS-0057-F), Medicare Advantage organizations, Medicaid and CHIP programs and plans, and Qualified Health Plan issuers on the federal exchanges must offer Provider Access and Payer-to-Payer APIs that share USCDI data, with

most API requirements starting on January 1, 2027 [1]. For many of these organizations, the practical path is to convert the C-CDA documents they already receive into FHIR resources.

Conversion sounds mechanical, but it is not. C-CDA and FHIR model the same clinical facts differently. Status, timing, coding and context sit in different places, and some concepts have no direct counterpart. HL7's C-CDA on FHIR implementation guide now publishes consensus maps for the main clinical domains [2]. In practice, though, many teams start from open-source, template-based converters, and there is little published evidence on how much content those tools actually preserve.

Measuring that is harder than it looks. With real documents there is no reference answer to compare against, and protected health information limits what can be shared. Synthea, an open-source synthetic patient generator, offers a way around both problems [5]. It can export the same simulated patient both as a C-CDA document and as a native FHIR bundle, generated from one underlying record. The native bundle can then serve as a reference standard for whatever a converter produces from the document.

This paper uses that pairing to evaluate a widely available open-source converter. It makes four contributions:

1. A reproducible method that separates what is lost when a record is written as C-CDA from what is lost when that C-CDA is converted to FHIR.
2. Measurements across 331 synthetic patients and more than 276,000 source clinical entries.
3. A taxonomy of conversion defects, each with a count and a root cause traced to the source document or the converter.
4. Open code that lets others repeat the evaluation with other converters or newer template versions.

2. Methods

2.1 Synthetic patients

We ran Synthea (master-branch build, downloaded 24 October 2026) for a Massachusetts population of 300 living patients with ten years of history, using 20261024 as both the random seed and the clinician seed. C-CDA and FHIR R4 exports were switched on; hospital, practitioner and CSV exports were switched off. Synthea produced 344 records (300 living, 44 deceased). Of these, 337 received a FHIR export and 331 a C-CDA export. The 331 patients with both exports form the study set. Their C-CDA documents had a median size of 455 KB and a maximum of 12.6 MB.

2.2 Converter

We evaluated `python-fhir-converter` 0.3.0, a Python port of the Liquid templates in Microsoft's FHIR-Converter [6, 7], using its CCD root template with default settings. The package pins older dependency versions that do not build on Python 3.13, so we installed it alongside newer releases (`python-liquid`

1.13.0). All 331 documents converted without errors in 160 seconds on a single process. The median time per document was 0.24 s, the 95th percentile 1.35 s, and the slowest 7.9 s.

2.3 Three layers of the same record

For each patient we compared three versions of the record:

- **Native FHIR (N):** Synthea’s own FHIR bundle, used as the reference standard.
- **C-CDA source (S):** Synthea’s C-CDA document, the input to the converter.
- **Converted FHIR (C):** the bundle the converter produced from S.

From each layer we extracted clinical events in eight families. Table 1 shows where each family comes from. Observations with components, such as blood pressure, were split into one event per component in every layer, so a panel and its parts compare like for like. Each event records its family, code, calendar day, raw timestamp, value, unit and status.

Table 1
Sources of each event family in the three layers

Family	Native and converted FHIR	C-CDA source
Condition	Condition	Problems section, problem observation value
Procedure	Procedure	Procedures section
Encounter	Encounter	Encounters section
Immunization	Immunization	Immunizations section
Medication	MedicationRequest or MedicationStatement	Medications section
Allergy	AllergyIntolerance	Allergies section, allergen participant
Observation	Observation (and components)	Results, Vital Signs, Social History and Functional Status sections
DiagnosticReport	DiagnosticReport	Result organizers

2.4 Matching

Within each patient, events were matched as multisets on family, code and calendar day. Inside one key, events with equal values were paired first. We used two matching rules:

- **Strict:** the converted code must equal the source code and sit in the element FHIR defines for it. This answers whether a standard query would find the fact.

- **Lenient:** also accepts a code the converter moved into another element, and a source translation code that replaced the source's own code. This answers whether the fact survived at all.

2.5 Measures

- **Export coverage:** share of native events found in the C-CDA source (N to S).
- **Conversion recall and precision:** share of source events found in the converted bundle, and share of converted events traceable to the source (S to C).
- **End-to-end recall:** share of native events found in the converted bundle (N to C).
- **Field fidelity:** for matched source and converted pairs, whether the time of day, numeric value, unit, code system URI and condition end date survived. For matched native and converted pairs, whether status and the link to an encounter survived.
- **Structural validity:** every native and converted resource was parsed with the fhir.resources 7.1.0 R4B models [10], and failures were counted by resource type.
- **US Core elements:** the share of converted resources that carry selected elements US Core 9.0.0 makes mandatory, such as AllergyIntolerance.code, Condition.category, Encounter.type and the laboratory category on DiagnosticReport [4].

2.6 Time zone experiment

Early results showed that converted timestamps kept only the date. To find the cause, we took one source document, added a UTC offset of -0400 to each of its 375 fourteen-digit timestamps, converted it again, and counted how many output timestamps kept the time of day.

2.7 Use of generative AI

Claude (claude-opus-5-5), was used through drafting the manuscript text from the computed results.

3. Results

3.1 How many facts survived

The converter kept almost every clinical fact it was given, but only nine in ten in the place a standard query would look. Of 276,153 source entries, 96.9% appeared in the converted bundle under lenient matching and 89.8% under strict matching. Precision was 99.996%: only 10 converted events could not be traced to the source. Across patients, lenient recall had a median of 96.3% (range 90.3% to 100%).

Only 8,444 source entries disappeared outright. Of these, 8,434 were assessment scores (for example PHQ-9 and GAD-7 totals) that Synthea writes in the Functional Status section, which the converter did

not process at all. The other 10 were drug allergies whose allergen code was dropped.

Table 2
Conversion recall by family (C-CDA source to converted FHIR)

Family	C-CDA entries	Strict recall	Lenient recall	What went wrong
Condition	12,066	100%	100%	Status and end date dropped
Procedure	52,519	100%	100%	Time of day dropped
Encounter	18,234	0%	100%	Type code written to Encounter.class
Immunization	4,772	100%	100%	CVX system as an OID
Medication	16,023	100%	100%	Status missing; every resource invalid
Allergy	320	0%	96.9%	Allergen moved to reaction; 10 lost
Observation	153,448	93.8%	94.5%	Functional Status section skipped; 1,109 codes replaced
DiagnosticReport	18,771	100%	100%	Category missing

The C-CDA documents themselves carried less than the native FHIR. They contained 71.9% of native events (per-patient median 67.7%, range 53.2% to 92.3%), so end-to-end recall from native to converted FHIR was 71.8%. The gap sat almost entirely in observations and reports. Synthea's C-CDA left out 88% of survey observations, 92% of social-history observations and 21% of vital signs, but only 4.5% of laboratory results. Every laboratory report reached the C-CDA. The 18,234 clinical-note reports in the native bundle did not, which is why DiagnosticReport coverage was 50.7%.

3.2 What happened to the details

Values came through intact. All 145,014 matched observation values and all 122,623 units were identical to the source.

Time of day did not. Every source entry carried a time, yet none of the 267,709 matched converted events kept one; all were reduced to a date. The cause lies in how the two systems meet. Synthea writes timestamps without a UTC offset, although the C-CDA US Realm date-time rules say a timestamp that includes a time must include an offset [3]. When no offset is present, the converter keeps only the date. In our controlled test, after adding - 0400 to the 375 timestamps in one document, all 246 timestamps in the output kept their time.

Codes were sometimes replaced. When a source code carried a translation, the converter kept the translation and dropped the original code. This affected 1,109 observations and 6 encounters:

- Oxygen saturation (LOINC 2708-6) became pulse oximetry (59408-5) 422 times.
- Body temperature (8310-5) became oral temperature (8331-1) 395 times.

- Right and left eye intraocular pressure (79892-6 and 79893-4) both became one SNOMED CT code, 41633001, in 292 observations, so the two eyes can no longer be told apart.

Immunization codes kept their values, but all 4,772 used the OID-style system

urn:oid:2.16.840.1.113883.12.292 instead of the canonical CVX URI, <http://hl7.org/fhir/sid/cvx>.

Status and timing fields were the weakest:

- **Conditions:** none of the 12,066 had a clinical status, and none of the 8,901 problems with an end date in the source kept it. The source carried no Problem Status observation, but every problem had a concern status and resolved problems had an end date.
- **Medications:** none of the 16,023 had a status, although every source entry was marked completed.
- **Encounters:** all 18,234 had status “unknown”. Here the source was at fault: it carried no encounter status, while native FHIR said “finished”.
- **Encounter links:** native FHIR linked every condition, procedure, observation, immunization, medication and report to an encounter. No converted resource did, because Synthea’s C-CDA contains no entry-to-encounter references.

3.3 Structure and US Core

Of 285,397 converted resources, 16,601 (5.8%) failed structural validation, compared with none of the 416,708 native resources. The failures were concentrated:

- All 16,023 MedicationStatement resources lacked the required status.
- 247 of 320 AllergyIntolerance resources were invalid: 156 had a reaction with no manifestation, and 91 wrote category as a single string instead of a list.
- All 331 Composition resources had an attester without the required mode.

The converted output also missed elements that US Core 9.0.0 makes mandatory [4]. No AllergyIntolerance carried a code; 310 held the allergen in reaction.substance, an element US Core does not mark Must Support. No Condition had a category, no Encounter had a type, and no DiagnosticReport had the laboratory category. Medications came out as MedicationStatement, which US Core no longer profiles; it represents medication lists with MedicationRequest. Observation category, by contrast, was present on all 145,014 observations.

Table 3
Defects found, with counts and root cause

Defect	Affected	Root cause	Effect
Time of day dropped	267,709 events	Source omits UTC offset; converter keeps date only	Ordering within a day lost
Status missing	16,023 medications	Converter	Resources invalid
Status and end date missing	12,066 conditions	Converter	Resolved problems look unresolved
Section skipped	8,434 observations	Converter	Assessment scores lost
Type in wrong element	18,234 encounters	Converter	Visit type not queryable
Allergen in wrong element	310 allergies	Converter	Allergen not queryable by code
Allergen dropped	10 allergies	Converter	Allergy content lost
Translation replaces code	1,115 entries	Converter	Original code lost; laterality lost in 292
Category missing	30,837 resources	Converter	US Core non-conformance
Encounter status unknown	18,234 encounters	Source	Visit status lost
No encounter links	All entries	Source	Visit context lost

4. Discussion

The headline numbers are reassuring and misleading at the same time. Judged by record counts, the converter kept 97% of what it received and changed none of the values. Judged by what a downstream FHIR client can use, the picture is different: every medication fails validation, no allergy can be found by its allergen code, every encounter has lost its type, and every timestamp has lost its time. A pipeline that checks only row counts or reconciliation totals would report success and pass all of these problems on.

4.1 Why this matters for payers

Under CMS-0057-F, payers must expose USCDI data through Provider Access and Payer-to-Payer APIs by January 2027 [1], and in practice those APIs are built on US Core. Converted data with no allergy code, no condition category and no encounter type will not meet US Core without further work. The errors also compound when one payer's API feeds another's. A resolved condition with no end date and no status

looks active to the next organization, and an assessment score that was never converted cannot be shared at all.

Quality measurement is exposed too. Assessment instruments such as PHQ-9 feed depression screening and follow-up measures, yet the whole Functional Status section, which held 8,434 of these scores, was skipped. Day-level timestamps are usually enough for annual measures, but they break rules that depend on order within a day, such as which reading came last.

4.2 Practical recommendations

The findings point to a short list of checks for any team putting a C-CDA converter into production:

1. **Validate against US Core, not just base FHIR.** Base structural checks caught the medication and allergy errors, but only profile checks reveal missing categories, codes and types.
2. **Keep every coding.** When a source code has translations, carry all of them as codings rather than choosing one. This would have preserved laterality and the original LOINC codes.
3. **Fix time zones at the boundary.** Ask senders for UTC offsets, as C-CDA requires, or apply a documented default offset per source before conversion.
4. **Derive status from what the source does say.** The C-CDA on FHIR maps fall back to the concern status when no Problem Status observation exists, map the end date to abatement, set category from the section, and map medication statusCode to status [2]. The converter tested here did none of these.
5. **Check section coverage.** List every section in incoming documents and confirm that each one has a mapping, so content is never dropped silently.
6. **Test with paired synthetic data first.** The method used here costs little, needs no PHI, and isolates whether a fault lies with the sender or the converter.

4.3 Related work

Earlier work showed that C-CDA documents from certified EHRs vary widely in how they use templates, codes and optional elements [8]. More recent work has built CDA-to-FHIR transformations with the FHIR Mapping Language and evaluated them for usability with domain experts [9]. HL7's C-CDA on FHIR guide supplies consensus maps but no measurements of particular tools [2]. Randmaa et al. recently proposed measuring the quality of CDA-to-FHIR transformations with text-mining techniques [11]. Our study complements that line of work with a different design: it compares a converter's output, event by event and field by field, against a reference standard generated from the same underlying record, and it attributes each defect to either the source document or the converter.

5. Limitations

Synthetic documents are cleaner and more uniform than real ones. Synthea's C-CDA exporter has its own habits, such as omitting time zone offsets and encounter links and placing assessment scores in the Functional Status section, and real documents from different EHRs will trigger other failures. The export coverage figures describe Synthea's exporter, not the limits of C-CDA as a standard.

We tested one converter, one version and default settings. The Python port may behave differently from the upstream Microsoft templates, and newer releases may fix some defects. We also ran it with newer dependency versions than it pins.

Matching on code and calendar day can pair the wrong events when the same test is repeated on one day; pairing equal values first reduces, but does not remove, that risk. Structural validation used the fhir.resources R4B models rather than the official HL7 validator, and the US Core checks covered selected mandatory elements rather than full profile validation.

6. Conclusion

An open-source C-CDA to FHIR converter preserved nearly every clinical fact and every value in 331 synthetic patients, yet it lost every time of day, every medication and condition status, every condition end date, and an entire section of assessment scores. It also put allergens and encounter types where standard queries will not find them. Counting records would hide all of this. Paired synthetic records offer a cheap and repeatable way to find such problems, and to tell a sender's faults from a converter's, before real patient data is at stake.

Declarations

Competing interests.

The author declares no competing interests. The author works in healthcare data interoperability. This study was carried out independently, using only public software, public standards and synthetic data; it uses no employer or client data, systems or proprietary methods, and the views expressed are the author's own.

Ethics approval.

Not applicable. The study used only computer-generated synthetic patient records from Synthea. No human participants, patient data or identifiable information were involved, so ethics committee approval was not required.

Consent to participate and consent to publish. Not applicable.

Use of generative AI.

Described in Section 2.7.

Funding.

No funding was received for this work.

Author Contribution

S.S.N. conceived and designed the study, developed the software, performed the analysis, and wrote and reviewed the manuscript. S.S.N. is the sole author and takes full responsibility for the work. A generative AI tool assisted with drafting manuscript, as described in Section 2.7.

Data Availability

No real patient data were used. The synthetic dataset can be regenerated exactly with Synthea using the seeds and settings in Section 2.1

Code availability.

All code is available at <https://github.com/neelisandeep/ccda-fhir-fidelity> under the MIT License.

References

1. Centers for Medicare & Medicaid Services. CMS Interoperability and Prior Authorization Final Rule (CMS-0057-F) fact sheet. 17 January 2024. <https://www.cms.gov/newsroom/fact-sheets/cms-interoperability-and-prior-authorization-final-rule-cms-0057-f>
2. HL7 International. C-CDA on FHIR Implementation Guide, STU 2 (version 2.0.0). <https://hl7.org/fhir/us/ccda/en/mappingBackground.html>
3. HL7 International. Consolidated CDA: US Realm Date and Time (Point in Time). <https://hl7.org/cda/us/ccda/StructureDefinition-USRealmDateTime.html>
4. HL7 International. US Core Implementation Guide, STU 9 (version 9.0.0). <https://hl7.org/fhir/us/core/>
5. Walonoski J, Kramer M, Nichols J, et al. Synthea: an approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record. J Am Med Inform Assoc. 2018;25(3):230–238. <https://doi.org/10.1093/jamia/ocx079>
6. Microsoft. FHIR-Converter. GitHub repository. <https://github.com/microsoft/FHIR-Converter>
7. python-fhir-converter, version 0.3.0. GitHub repository chaseastewart/fhir-converter. <https://github.com/chaseastewart/fhir-converter>

8. D'Amore JD, Mandel JC, Kreda DA, et al. Are Meaningful Use Stage 2 certified EHRs ready for interoperability? Findings from the SMART C-CDA Collaborative. *J Am Med Inform Assoc.* 2014;21(6):1060–1068. <https://doi.org/10.1136/amiajnl-2014-002883>
9. Bossenko I, Randmaa R, Piho G, Ross P. Interoperability of health data using FHIR Mapping Language: transforming HL7 CDA to FHIR with reusable visual components. *Front Digit Health.* 2024;6. <https://doi.org/10.3389/fdgth.2024.1480600>
10. fhir.resources, version 7.1.0. GitHub repository [nazrulworld/fhir.resources](https://github.com/nazrulworld/fhir.resources). <https://github.com/nazrulworld/fhir.resources>
11. Randmaa R, Bossenko I, Piho G, Ross P. Measuring the quality of CDA to FHIR transformations using text mining techniques. *Stud Health Technol Inform.* 2026. <https://doi.org/10.3233/SHTI260294>