



**HL7 EHR Guidance:**  
**Data Consistency and Comparability, Edition 1**

May 2025

**HL7 Comment-Only Ballot**

**Sponsored by:**  
**Electronic Health Records Work Group**

Copyright © 2025 Health Level Seven International ® ALL RIGHTS RESERVED. The reproduction of this material in any form is strictly forbidden without the written permission of the publisher. HL7 and Health Level Seven are registered trademarks of Health Level Seven International. Reg. U.S. Pat & TM Off.

Use of this material is governed by HL7's [IP Compliance Policy](#).

## IMPORTANT NOTES:

HL7 licenses its standards and select IP free of charge. **If you did not acquire a free license from HL7 for this document**, you are not authorized to access or make any use of it. To obtain a free license, please visit <http://www.HL7.org/implement/standards/index.cfm>.

**If you are the individual that obtained the license for this HL7 Standard, specification or other freely licensed work (in each and every instance "Specified Material")**, the following describes the permitted uses of the Material.

**A. HL7 INDIVIDUAL, STUDENT AND HEALTH PROFESSIONAL MEMBERS**, who register and agree to the terms of HL7's license, are authorized, without additional charge, to read, and to use Specified Material to develop and sell products and services that implement, but do not directly incorporate, the Specified Material in whole or in part without paying license fees to HL7.

INDIVIDUAL, STUDENT AND HEALTH PROFESSIONAL MEMBERS wishing to incorporate additional items of Special Material in whole or part, into products and services, or to enjoy additional authorizations granted to HL7 ORGANIZATIONAL MEMBERS as noted below, must become ORGANIZATIONAL MEMBERS of HL7.

**B. HL7 ORGANIZATION MEMBERS**, who register and agree to the terms of HL7's License, are authorized, without additional charge, on a perpetual (except as provided for in the full license terms governing the Material), non-exclusive and worldwide basis, the right to (a) download, copy (for internal purposes only) and share this Material with your employees and consultants for study purposes, and (b) utilize the Material for the purpose of developing, making, having made, using, marketing, importing, offering to sell or license, and selling or licensing, and to otherwise distribute, Compliant Products, in all cases subject to the conditions set forth in this Agreement and any relevant patent and other intellectual property rights of third parties (which may include members of HL7). No other license, sublicense, or other rights of any kind are granted under this Agreement.

**C. NON-MEMBERS**, who register and agree to the terms of HL7's IP policy for Specified Material, are authorized, without additional charge, to read and use the Specified Material for evaluating whether to implement, or in implementing, the Specified Material, and to use Specified Material to develop and sell products and services that implement, but do not directly incorporate, the Specified Material in whole or in part.

NON-MEMBERS wishing to incorporate additional items of Specified Material in whole or part, into products and services, or to enjoy the additional authorizations granted to HL7 ORGANIZATIONAL MEMBERS, as noted above, must become ORGANIZATIONAL MEMBERS of HL7.

Please see <http://www.HL7.org/legal/ippolicy.cfm> for the full license terms governing the Material.

**Ownership.** Licensee agrees and acknowledges that **HL7 owns** all right, title, and interest, in and to the Materials. Licensee shall **take no action contrary to, or inconsistent with**, the foregoing.

**Licensee agrees and acknowledges that HL7 may not own all right, title, and interest, in and to the Materials and that the Materials may contain and/or reference intellectual property owned by third parties ("Third Party IP"). Acceptance of these License Terms does not grant Licensee any rights with respect to Third Party IP. Licensee alone is responsible for identifying and obtaining any necessary licenses or authorizations to utilize Third Party IP in connection with the Materials or otherwise. Any actions, claims or suits brought by a third party resulting from a breach of any Third Party IP right by the Licensee remains the Licensee's liability.**

Following is a non-exhaustive list of third-party terminologies that may require a separate license:

Terminology	Owner/Contact
Current Procedures Terminology (CPT) code set	American Medical Association <a href="https://www.ama-assn.org/practice-management/cpt-licensing">https://www.ama-assn.org/practice-management/cpt-licensing</a>
SNOMED CT®	SNOMED CT® International; <a href="http://www.snomed.org/snomed-ct/get-snomed-ct">http://www.snomed.org/snomed-ct/get-snomed-ct</a> or <a href="mailto:info@ihtsdo.org">info@ihtsdo.org</a>
Logical Observation Identifiers Names & Codes (LOINC®)	Regenstrief Institute
International Classification of Diseases (ICD) codes	World Health Organization (WHO)
NUCC Health Care Provider Taxonomy code set	American Medical Association. Please see <a href="http://www.nucc.org">www.nucc.org</a> . AMA licensing contact: 312-464-5022 (AMA IP services)

## Why a “Comment Only” Ballot?

The HL7 Electronic Health Record Work Group is assessing a broad range of issues related to the quality of health data/information. This is (in part) prompted by our work in the Reducing Clinician Burden Project where we have found that assurance of data quality is vital to address concerns regarding whether data is reliable and fit for use.

Data quality is foundational and is of vital interest to ALL stakeholders in, and for ALL purposes/uses of, health data/information:

- For the subject of data content (i.e., patient)
- For the author of data content
- For the end user of data content
- For primary use, secondary use and re-use of data content
- For ensuring integrity and effectiveness of the clinical process
- For assuring patient safety
- For data content ingested by artificial intelligence, including machine learning and (natural) language processing
- ...

[See Annex B for more expansive value propositions regarding data quality.]

In our world today, too often quantitative prowess (i.e., exchange transaction volumes) is touted over actual qualitative achievement. Is the received data trustworthy and usable? If so, what is the basis or evidence to support such confidence?

Our goal has long been focused on standards which address data quality and evidence thereof. See Annex A for Standards, including several of the HL7, ISO and ASTM Standards which are so focused.

As we consider additional topic areas, in this case “data consistency and comparability”, we agreed to reach out to the HL7 International community to gather expert guidance and real-world experience to inform our efforts before developing a standard in this space.

This “comment only” ballot includes a white paper intended to lay out the problem space as we manage health data from point of data origination to point of data end use.

Thus our inquiry for your consideration and feedback:

1. *What constructs are needed to algorithmically assess consistency and comparability of datasets (at the point of data end use)? To provide guidance to the end user?*
2. *Is it feasible to use existing standard constructs (e.g., FHIR Concept Maps, FHIR Structured Definitions)? If so, how?*
3. *Do we need to devise new standard constructs (e.g., Repositories of Data Definitions – maintained by data sources, referenced by data users)?*

## HL7 EHR Work Group – Data Quality Project Team

### Ensuring Data Consistency and Comparability – Data At Rest, in Motion and Post Exchange

General Objective: To distinguish which data elements and their content:

- Have characteristics in common with like data;
- Are “fit for use” to facilitate ready and reliable comparison (correlation/trend-ability) with like data;
- Without assurance that like Element Names always infer consistency and comparability.

Such data characteristics include similar/same context and data definition – e.g., element name(s), data type(s), range, input/display/storage format, unit(s) and scale of measure, method and purpose of capture.

Burden Reduction Objective: To reduce end user (clinician) burden by ensuring confidence/clarity in the ability to establish trends and draw comparisons between like content element occurrences over time.

Table A – Element Definition – Simple Exchange Example				
	Element Definition	Sender (Source/Originator)	→ Exchange Artifact (e.g., FHIR resource)	→ Receiver (End User/Use)
A	Element Name	DEF1 DEF1 DEF1 DEF1 DEF1	DEF1 DEF2 DEF1 DEF2 DEF2	DEF1 DEF3 DEF2 DEF2 DEF1
B	Element Description	[as per Row A]		
C	Captured by/from...	Keyed input, voice...	N/A	Exchange artifact
D1	Format when input	DEF1 DEF1	N/A	DEF1 DEF2
D2	Format when retained	[as per Row D1]		
D3	Format when transformed to/from Exchange Artifact	DEF1 (post-transformation)	DEF1	DEF1 (pre-transformation)
D4	Format when rendered to end user/use	[as per Row D1]		
E	Data Type	[as per Row A]		
F	Precision	[as per Row A]		
G	Range	[as per Row A]		
H	Unit of Measure	[as per Row A]		
I	Sampling Site	[as per Row A]		
J	Sampling Method	[as per Row A]		
K	Condition of Measure	[as per Row A]		
L	Purpose of Capture/ Purpose of Use	[as per Row A]		
Element Definition Patterns (P1-P5)				
Ideal Definition Pattern Full Equivalence				
P1. Sender = Exchange = Receiver		DEF1	DEF1	DEF1
Variant Definition Patterns Potential Mismatch				
P2. Sender ≠ Exchange ≠ Receiver		DEF1	DEF2	DEF3
P3. (Sender = Exchange) ≠ Receiver		DEF1	DEF1	DEF2
P4. Sender ≠ (Exchange = Receiver)		DEF1	DEF2	DEF2
P5. (Sender = Receiver) ≠ Exchange		DEF1	DEF2	DEF1
Key				
DEF1, DEF2, DEF3 = First, second and third order Element Definition				
N/A: Not Applicable				
Element Definition is based on data type and key clinical context elements.				

1. Pervasive Heterogeneity. Regardless of the Element Definition of the Exchange Artifact in the middle, between the Sender and Receiver – Element Definitions at either end are often disparate – reflecting our world of pervasive heterogeneity.
2. Element Definitions are Fundamental:
  - a. To Comparison and Correlation. Focusing on the Receiver and End User (who is often a clinician) of health data/record content, there is required knowledge of Element Definitions in order to ensure consistency in managing aggregations of data element instances from both internal and external sources, which might be purposed for comparison, correlation, graphing and trending.
  - b. To Trustworthiness. As with 2.a, trustworthiness of data content is a key factor in terms of reliance and incorporation into the Receiver's health record.
3. Sender/Receiver Partners:
  - a. One Sender will exchange with many Receiver “partners”.
  - b. One Receiver will exchange with many Sender “partners”.

Table B – Who are the Actors/Roles and What are the Activities associated with Element Definition?					
	What – Element Definition Activity	Who (Actor, Role)			When – Frequency and Repetition
		Sender	Exchange	Receiver	
Develop	Create Element Definition	Software Developer 1	SDO	Software Developer 2	Once each: 3 Definitions
	Map to/from Exchange Artifact				Once each: 2 Mappings
	Verify Mapping	Clinical Reviewer 1		Clinical Reviewer 2	Once each: 2 Mappings
Test	Test Exchange with Typical Datasets (to mirror real-world content and variations)	Tester 1		Tester 2	Once initially, then once again at each update – repeating until fully verified
	Verify Test Results	Quality Assurance Reviewer 1		Quality Assurance Reviewer 2	
Live	Production Real-Time Exchange	Source Author/Originator			End User (Clinician)

4. Running on Independent Tracks. Element Definitions are typically developed and tested independently.
5. Dependence on Proper Development and Testing. Errors, disjunctions and/or omissions will likely introduce serious consequences/risks once exchange moves into Live Production where clinical evaluation and decision making, indeed patient safety and clinical integrity, are at stake. These propagate to 100s and even 1000s of live exchange transactions, if not detected, averted and resolved.

Table C – What do the Actors know about Element Definitions?					
			Element Definitions		
	Who (Actor, Role)	Sender-side/ Receiver-side	Sender	Exchange Artifact	Receiver
Human	Software Developer 1	Sender	Knows	Knows	Rarely Knows
	Software Developer 2	Receiver	Rarely Knows	Knows	Knows
	Clinical Reviewer 1	Sender	Knows	Knows	Rarely Knows
	Clinical Reviewer 2	Receiver	Rarely Knows	Knows	Knows
	Tester 1	Sender	Knows	Knows	Rarely Knows
	Tester 2	Receiver	Rarely Knows	Knows	Knows
	Quality Assurance Reviewer 1	Sender	Knows	Knows	Rarely Knows
	Quality Assurance Reviewer 2	Receiver	Rarely Knows	Knows	Knows
	Author/Originator (Clinician)	Sender	May Know	Rarely Knows	Rarely Knows
	End User (Clinician)	Receiver	Rarely Knows	Rarely Knows	May Know
System	Source System	Sender	Knows algorithmically	Knows algorithmically	Rarely Knows
	Destination System	Receiver	Rarely Knows	Knows algorithmically	Knows algorithmically

6. What a human knows: is cognitive knowledge – as acquired by learning and understanding.
7. What a system knows: is algorithmic knowledge – as programmed in software.
8. Technical Actors. Typically (as highlighted in light green cells above):
  - a. Sender-side Technical Actors (both Human and System) know the Sender-side and Exchange Element Definitions; whereas
  - b. Receiver-side Technical Actors (both Human and System) know the Receiver-side and Exchange Element Definitions.
9. Clinical Actors. Typically (as highlighted in light blue cells above):
  - a. Sender-side Human Clinical Actors (Authors, Originators) may know the Sender-side Element Definitions but likely do not the Element Definitions of the Exchange Artifact or the Receiver; whereas
  - b. Receiver-side Human Clinical Actors (End Users) may know the Receiver-side Element Definitions but likely do not the Element Definitions of the Sender or Exchange Artifact.
10. Evaluation x 2. As noted in Comments 3 & 8, Element Definitions are typically evaluated pair-wise:
  - a. Sender to Exchange; and
  - b. Exchange to Receiver.
11. Evaluation x 3. Only when exchange involves systems residing within the same enterprise or across closely related organizations is it likely that all three Element Definitions are evaluated together: Source to Exchange Artifact to Receiver.

Table D – Simple Example – Systolic Blood Pressure			
	Element Definition	Example 1	Example 2
A	Element Name	Systolic Blood Pressure	Systolic Blood Pressure
B	Element Description	Measures the force the heart exerts on artery walls each time it beats (a.k.a., the “top number”) Source: Mayo Clinic	Measures the force the heart exerts on artery walls each time it beats (a.k.a., the “top number”) Source: Mayo Clinic
C	Input/capture method	Keyed input	Keyed input
D1	Format when input	Integer	Integer
D2	Format when retained	Integer	Integer
D3	Format when transformed to/from exchange artifact	Integer	Integer
D4	Format when rendered to end user/use	Integer	Integer
E	Data Type	Integer	Integer
F	Precision	Whole number	Whole number
G	Range	0-300	0-300
H	Unit of Measure	mmHg	mmHg
I	Sampling Site	Upper left arm w/cuff	Upper left arm w/cuff
J	Sampling Method	Auscultation via stethoscope using inflatable cuff	Auscultation via stethoscope using inflatable cuff
K	Condition of Measure	• Patient Position = <b>standing</b> • Patient is resting	• Patient Position = <b>supine</b> • Patient is resting
L	Purpose of Capture/ Purpose of Use	Vital Sign Measurement/ Evaluation	Vital Sign Measurement/ Evaluation

12. Key Clinical Parameter. Condition of Measure/Patient Position is key to ensure comparability of Systolic (and Diastolic) Blood Pressure measurements.
13. Potential Points of Failure:
  - a. The Diastolic Blood Pressure Element Definition of the Sender, Exchange Artifact or Receiver do not all include Condition of Measure/Patient Position.
  - b. The exchange instance did not convey Condition of Measure/Patient Position from Source to Exchange to Receiver.
  - c. The Receiver did not recognize the variance in Patient Position value(s) in its aggregation of Diastolic Blood Pressure measurements over time before offering a view to the Clinician which included a comparison, correlation or trend.
14. Fail-safe Actor. *Which Actor is best able to catch the failure before incompatible data is offered to End Users (including clinicians)?*
  - a. Human Actors: No because, as per Table C, they will rarely “know” the Element Definitions of all three: Sender, Exchange Artifact, Receiver.
  - b. System Actors, in one case: No, because, as per Table C, they will rarely “know” the Element Definition of their Exchange Partner, as Sender or Receiver.
  - c. System Actors, in second case: Yes, because they will have captured/analyzed the Element Definition of their Exchange Partner and can thus algorithmically compute variances in the Definition and also in data values exchanged (including metadata).
15. Per 14.c: Described is the need for transparency of Element Definition across Source and Receiver systems. *How might this be achieved?*
  - a. By assigning each Element Definition a Globally Unique ID (GUID) – unique to the source or receiver system and each element exchanged? And/or...
  - b. By exchanging each Element Definition initial configuration/binding, and periodically (as necessary) thereafter, in or out of band of production exchange? And/or...
  - c. By establishing one or more repositories of Element Definitions – maintained by data sources, referenced by data users? And/or...
  - d. Other alternatives?

More examples follow...

Table E – Simple Example – Serum Creatinine			
	Element Definition	Example 1	Example 2
A	Element Name	Serum Creatinine	Serum Creatinine
B	Element Description	Creatinine concentration [mass/volume] in serum LOINC #2160-0	Creatinine concentration [mass/volume] in serum LOINC #2160-0
C	Input/capture method	Transmitted electronically	Transmitted electronically
D1	Format when input	Decimal	Decimal
D2	Format when retained	Decimal	Decimal
D3	Format when transformed to/from exchange artifact	Decimal	Integer
D4	Format when rendered to end user/use	Decimal	Decimal
E	Data Type	Decimal	Decimal
F	Precision	Two decimal places	Two decimal places
G	Range	0.57 – 1.27*	45 – 104*
H	Unit of Measure	mg/dL UCUM code mg/dL = 'mg/dL'	umol/L UCUM code umol/L = 'umol/L'
I	Sampling Site	Blood	Blood
J	Sampling Method	Venipuncture	Venipuncture
K	Condition of Measure	Not applicable	Not applicable
L	Purpose of Capture/ Purpose of Use	Renal function evaluation	Renal function evaluation

\* Although these values are typical, the normal range for serum creatinine will vary between different healthcare organizations using different blood chemistry analyzers.

Table F – Simple Example – Metformin 500 mg oral tablet			
	Element Definition	Example 1	Example 2
A	Element Name	Metformin	Metformin
B	Element Description	Drug used to lower blood glucose in treatment of diabetes	Drug used to lower blood glucose in treatment of diabetes
C	Input/capture method	Keyed input	Keyed input
D1	Format when input	Text String	Text String
D2	Format when retained	Alphanumeric Drug Code	Alphanumeric Drug Code
D3	Format when transformed to/from exchange artifact	Alphanumeric Drug Code	Alphanumeric Drug Code
D4	Format when rendered to end user/use	Text String	Text String
E	Data Type	Alphanumeric Drug Code	Alphanumeric Drug Code
F	Precision	Not Applicable	Not Applicable
G	Code	861007	A10BA02
H	Coding System	RxNorm RxCUI	WHO ATC Classification
I	Dose	2 Tablets	2 Tablets
J	Interval	Twice daily with food	Twice daily with food
K	Condition of Measure	Not Applicable	Not Applicable
L	Purpose of Capture/ Purpose of Use	Patient Medication List	Patient Medication List



## **ANNEX A – Key Reference Sources and Standards** – which focus on data quality and evidence thereof

- [HL7 EHR WG/Data Quality Confluence Page](#)  
Collateral and Reference Material for the EHR WG/Data Quality Project
- [ISO 21089:2018, Health Informatics – Trusted End-to-End Information Flows](#)  
Requirements Standard for management of Record Entries – across multiple systems  
Record Entry Lifespan – from point of data origination to point of data access/use or point of data archive or deletion  
Record Entry Lifecycle Events – occurring within lifespan
- [HL7/ISO 10781:2023, Health Informatics – Electronic Health Record System Functional Model R2.1](#)  
Record Infrastructure Section  
Requirements Standard for management of Record Entries – within an EHR system
- [HL7/ISO 16527:2023, Health Informatics – Personal Health Record System Functional Model R2](#)  
Record Infrastructure Section  
Requirements Standard for management of Record Entries – within a PHR system
- [HL7/ASTM 2147:2018, Standard Specification for Audit and Disclosure Logs for Use in Health Information Systems](#)  
Requirements Standard for health record/data audit
- [HL7 FHIR R5 Record Lifecycle Event Implementation Guide \(2023\)](#)  
Implementation Standard for FHIR, based requirements specified in ISO 21089, HL7/ISO 10781/16527 and HL7/ASTM E2147

## ANNEX B – Reference Sources

### Value Proposition for Data Quality – X/Grok (9 March 2025):

Data quality is important because it directly impacts the reliability, accuracy, and usefulness of any analysis, decision-making, or system that depends on that data. Here's why it matters:

1. **Better Decisions:** High-quality data—accurate, complete, consistent, and up-to-date—ensures that decisions based on it are sound. Poor data can lead to misjudgments, like a business investing in the wrong market because sales figures were off.
2. **Efficiency:** Clean data saves time and resources. If you're sifting through duplicates, errors, or missing values, you're wasting effort that could be spent on actual insights or action.
3. **Trust:** People and systems rely on data they can believe in. If the data's shaky, confidence erodes—whether it's a scientist studying climate trends or a company tracking customer behavior.
4. **Cost Reduction:** Bad data can get expensive fast. Think of marketing campaigns targeting the wrong people due to outdated addresses, or inventory mismanagement from faulty stock numbers. Fixing errors after the fact costs more than getting it right up front.
5. **Performance of Tech:** AI, machine learning, and analytics thrive on good data. Garbage in, garbage out—models trained on messy datasets spit out unreliable predictions or recommendations.

Real-world example: In healthcare, poor data quality (like incorrect patient records) can lead to misdiagnoses or wrong treatments—literally life-or-death stakes... Without it, you're building on sand.

### Value Proposition for Data Quality – Microsoft Co-Pilot (9 March 2025):

The value proposition for data quality is multi-faceted and significant, especially in fields like health IT where precision and compliance are paramount. Here are some key benefits:

1. **Improved Decision Making:** High-quality data ensures that decisions are based on accurate and reliable information, reducing the risk of errors and improving outcomes.
2. **Operational Efficiency:** Clean, consistent data reduces the time spent on data correction and reconciliation, allowing for more streamlined operations and better resource allocation.
3. **Regulatory Compliance:** Accurate data is crucial for meeting regulatory requirements and avoiding penalties. This is particularly important in health IT where compliance with regulations like HIPAA is mandatory.
4. **Enhanced Patient Care:** In the health IT sector, high-quality data translates to better patient outcomes by ensuring that medical decisions are based on accurate patient information.
5. **Cost Savings:** Reducing the costs associated with data errors, such as incorrect billing or unnecessary diagnostic tests, can lead to significant financial savings.
6. **Reputation Management:** Maintaining high data quality helps in building trust with stakeholders, including patients, partners, and regulatory bodies, enhancing the organization's reputation.