Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

From Data to Decisions: A Provenance Method for Auditable Digital Health Twin Claims

Michael Robert Mueller

office@futuretech.at

Bratislava University of Economics and Management (BUEM)

Method Article

Keywords: digital health, clinical decision support, provenance, health informatics, interoperability, conformance testing

Posted Date: November 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11661767/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

Background

Longitudinal digital-health systems combine measurements, personal baselines, predictive models, population evidence, and decision policies. A system may therefore present a personal deviation as clinically meaningful without retaining the additional support required for that stronger claim. Ordinary lineage records show where an output came from; they need not constrain which support authorizes a particular claim transition.

Methods

We specify an information model with typed claim nodes, support-source records, transition-targeted support edges, and transition records. We formulate rules for deviation, prediction, clinical significance, and action; compare the model with FHIR R5 Provenance, W3C PROV, FHIR Clinical Reasoning, and SACM; and implement a reference validator. Twelve synthetic specification fixtures and five single-fault stress cases test whether declared transitions pass or block.

Results

The validator matched all twelve specified outcomes. It permitted an active baseline to support deviation, a complete support bundle to support clinical significance, and a valid policy to support action. It blocked W2-only promotion, invalid mandatory sources, action without policy support, and reuse of prediction support for clinical significance. Existing standards carry provenance, rules, and assurance arguments; the method adds claim types, transition-targeted support edges, and claim-specific constraints.

Conclusion

Provenance-constrained claim transitions provide a reproducible method for testing declared support before a system releases a stronger output. Results establish structural conformance, not clinical validity, safety, utility, or patient benefit.

1 Introduction

Personalisation has changed the architecture of health information systems. Longitudinal systems can compare a new measurement with an individual baseline, combine that comparison with predictive models and population evidence, and produce a concise alert or recommendation. Individualised reference-interval research makes the baseline operation technically plausible while also emphasising its dependence on longitudinal data and unresolved implementation conditions [8,9,22]. Digital-twin

proposals make this pipeline explicit by combining sensor streams, electronic records, models, and decision-support functions within one person-specific representation [1–3,15,23,24]. The resulting integration can be useful, but it creates a representational problem: support sufficient for an earlier output can appear to authorize a later, stronger output.

Consider a resting-heart-rate measurement that is unusual relative to one person's baseline. This can support a bounded deviation claim: the current value differs from that person's prior pattern. It does not, by itself, support the stronger claim that the change is clinically significant, predicts a specified outcome, or warrants an action. Each statement has a different semantic type, intended use, and support requirement. If an interface presents them as one undifferentiated alert, clinicians and users cannot recover which relation supports which part of the apparent conclusion.

Lineage records alone do not prevent unsupported promotion. FHIR Provenance and W3C PROV can retain entities, activities, agents, versions, derivations, and responsibility [4,5]. Provenance templates and health-information-system reviews show that these facilities have been used to trace decision-support data and processes [10,11,18]. G-PROV is a closer comparator because it links guideline reasoning rules to their evidence sources [16]. It nevertheless operates at the level of guideline reasoning, rather than specifying a general, person-specific rule for authorising a transition from one typed claim to another. FHIR Clinical Reasoning can represent and evaluate executable knowledge artefacts and record a runtime guidance response [6]; standards-based CDS work shows how such logic can be coordinated with data exchange and workflow interfaces [19–21]. SACM can represent claims, evidence, and assurance arguments [7]. None of these specifications requires a health system to attach the exact support for a stronger claim to the transition that produces it. A baseline, a guideline, and a decision policy may coexist in the same system while the system does not declare which one authorized a specific promotion.

Decision provenance usefully broadens accountability beyond isolated algorithmic components to the flow of information through a decision pipeline [13]. The present method addresses a narrower informatics question within that pipeline: whether the support declared at one transition is admissible for the exact claim type that the transition releases.

The method treats release of a stronger output as a conformance question. It tests whether the support declared for a proposed transition satisfies the rule that the system itself makes explicit. It does not determine a diagnosis, select the clinically best action, or establish improvement in a deployed system.

The contribution has four parts. First, it specifies typed claim nodes, support-source records, support edges, and transition records. Second, it defines admissibility rules for transitions from measurement to deviation (T-DEV), prediction (T-PRED), clinical significance (T-SIG), and action (T-ACT). Third, it identifies how the model can be profiled over established provenance and clinical-reasoning standards. Fourth, it provides executable structural-conformance tests using non-patient-data fixtures. The evaluation is deliberately technical: it demonstrates that the reference rules permit and block specified records as designed; it is not a clinical evaluation.

2 Design objective and method

2.1 Design objective

The design objective is to prevent silent claim promotion in a system that processes longitudinal health information. The model targets systems in which the outputs of measurement, personal-baseline comparison, predictive modelling, clinical interpretation, and action policy may be displayed to different users or used in different workflows. A support relation must therefore be visible, time-bounded, scoped, and linked to the exact transition it authorizes.

The method is a design specification and conformance evaluation. It was developed from two representative settings: serial laboratory reference values and wearable monitoring. Neither setting is used as evidence about clinical effectiveness. Values, identifiers, sources, and results in the fixtures are synthetic. The model was compared with primary technical specifications for FHIR R5 Provenance, FHIR Clinical Reasoning, W3C PROV-DM, and SACM. The comparison identifies reusable native capabilities and distinguishes them from the new profile constraints proposed here.

2.2 Support categories

The model uses four support categories. W1 is evidence established across persons, including a validated population model or study. W2 is observation of the current person, such as a personal baseline. W3 is controlled individual response under stated conditions. W4 is a versioned health-object record identifying the capacity, activity, or justified goal on which a clinical-significance claim is asserted to bear. Model or simulation artefacts, compatibility records, and decision policies are represented as separately versioned sources.

The categories are not an evidence hierarchy. They name different roles in a transition. W2 supports that a current value is unusual for the person; it does not establish why that unusualness matters for health. A predictive model may release a typed prediction when its declared basis and validation metadata are present, but the prediction cannot be silently reused as a clinical-significance or action claim. A compatibility record states that a named evidence source bears on one exact health object within an explicit scope. A decision policy is required separately because even a valid clinical-significance claim does not select its own response.

3 Claim-transition information model

3.1 Core records

The four record types and the constraints added by the proposed profile are summarised in Table 1.

Table 1
Core records in the claim-transition information model

Record	Required function	Added constraint
Claim node	Identifies the asserted measurement, deviation, prediction, clinical-significance, or action claim	Claim type cannot silently change during downstream use
Support-source record	Identifies evidence, baseline, model, health object, compatibility assertion, or policy	Source category, version, scope, lifecycle, authority, validity interval, and conflict state are explicit
Support edge	States that one source supports one transition	The edge targets a transition record, never a claim node alone
Transition record	States the input claims, output claim, support rule, intended recipient/use, time, and validator result	Every stronger output has an inspectable support bundle

An active source is in scope, inside its validity interval, and free of unresolved conflict. A valid support edge requires both the edge and its exact referenced source to be active. This recursive condition matters for updating systems: a later expiry, withdrawal, or conflict blocks a dependent transition but does not erase an earlier claim that was valid under its own rule.

3.2 Admissibility rules

The rules are declarative rather than diagnostic. allOf requires each listed support edge; anyOf requires one valid alternative.

Table 2
Minimum admissibility rules for claim transitions

Transition	Minimum valid support	Permitted result
T-DEV	Measurement plus allOf[W2]	Personal deviation only
T-PRED	Valid input, versioned model or simulation, declared W1/W2/W3 basis, and validation metadata	Typed outcome prediction only
T-SIG	Valid input plus allOf[anyOf(W1,W3), W4, compatibility]	Clinical-significance claim within declared scope
T-ACT	Valid clinical-significance claim plus allOf[decision-policy]	Bounded action for the stated recipient and use

Two constraints make the model more restrictive than generic lineage. First, prediction support targets T-PRED; it cannot be reused for T-SIG merely because it appears in the same architecture. Second, a compatibility record must name the exact versions of the W1/W3 source and W4 health object. A code match, proximity in an evidence store, or a general statement of clinical plausibility is not sufficient.

A compatibility assertion is a governed, versioned record rather than an inference made at runtime. In the reference profile, a curator identifies the eligible W1/W3 and W4 versions, an independent reviewer may approve the record, and a withdrawal or unresolved conflict makes it unusable. The intended default is a reusable curated library; a case-specific assertion may be represented, but must still carry the same asserting authority, review state, scope, version, and validity interval. The validator checks these structural links and lifecycle conditions. It does not determine whether the underlying clinical compatibility judgment is substantively correct.

3.3 Standards alignment

FHIR Provenance can carry source, version, agent, and activity records; W3C PROV can express generic entities, activities, agents, and derivations [4,5]. FHIR Clinical Reasoning and GuidanceResponse can evaluate a declared rule and retain its runtime output [6]. SACM can retain the higher-level assurance rationale [7]. The proposed profile does not replace these specifications. It adds three constraints: typed claim classes, support edges that target named transitions, and recursively validated support rules for claim promotion. This distinction is also different from provenance-derived data-quality trust models, which assess the trustworthiness of inputs rather than whether a specified support bundle authorises a downstream claim transition [17].

Table 3 makes the comparison operational. “Composable” means that a representation can be assembled with the standard, but the stated constraint is not required by that standard alone.

Table 3
Capability comparison and proposed profile constraints

Capability	FHIR Provenance	W3C PROV	FHIR Clinical Reasoning	SACM	Proposed profile
Entity, activity, agent, and version lineage	Native	Native	Composable	Composable	Reuses native records
Executable knowledge and runtime guidance output	Composable	Not native	Native	Composable	Reuses native records
Claim and assurance argument	Composable	Composable	Composable	Native	Reuses native records
Typed claim classes and permitted promotion	Not constrained	Not constrained	Not constrained	Not constrained	Required
Edge must target one declared transition	Not constrained	Not constrained	Not constrained	Not constrained	Required
Exact W1/W3–W4 compatibility linkage, authority, and validity	Composable	Composable	Composable	Composable	Required and recursively checked

For example, the C3 wearable fixture can retain its sources as FHIR/PROV entities, its knowledge artefact and response through Clinical Reasoning, and its rationale in SACM. The profile adds the transition identifier, typed support edges, exact-version compatibility assertion, and validator outcome.

4 Structural-conformance evaluation

4.1 Reference implementation

The reference validator is a small Python implementation that reads JSON fixtures and returns a pass or block result with reasons. It is not a production clinical-decision-support service. Its purpose is to make the admissibility semantics executable and reproducible. The complete fixture set, validator, and result files are supplied as supplementary artefacts (Online Resource 1: Discover_Informatics_Supplementary_Artefacts_v3_20261107.zip).

Twelve fixtures were specified before execution. C1 tests a bounded deviation from an active personal baseline. C2 requests clinical significance using W2 only. C3 supplies valid W1 evidence, W4 health object, and compatibility record. C4 expires a mandatory W1 source. C5 requests action without a decision policy. C6 permits a typed prediction but requests an unsupported promotion to clinical

significance. C7–C12 respectively test a withdrawn measurement input, an elapsed validity interval despite active status, an edge targeting another transition, a mismatched compatibility assertion, a withdrawn clinical-significance input despite an active policy, and a compatibility assertion without an asserting authority.

4.2 Results

Table 4 reports the specified test intention and the observed validator outcome for each fixture.

Table 4
Results of the structural-conformance fixtures

Fixture	Intended test	Observed result
C1	Valid W2 baseline may support deviation only	Pass
C2	W2 alone must not support clinical significance	Block: W1/W3, W4, and compatibility absent
C3	Complete significance bundle may pass	Pass
C4	Expiry of mandatory W1 evidence blocks significance	Block: source expired
C5	Action requires a decision policy	Block: policy absent
C6	Prediction does not authorize significance promotion	Prediction pass; promotion block
C7	Deviation requires an active measurement input	Block: measurement withdrawn
C8	Validity interval overrides active status label	Block: W2 interval elapsed
C9	Support edge must target evaluated transition	Block: W1 edge targets another transition
C10	Compatibility must match exact source versions	Block: incompatible links
C11	Action requires active clinical-significance input	Block: input withdrawn
C12	Compatibility assertion requires asserting authority	Block: authority absent

All twelve observed results, including their expected reasons, matched the specified outcomes. Five single-fault stress cases additionally altered one otherwise valid support bundle. Expired or out-of-scope W1 evidence, a conflicted W4 health object, and a withdrawn compatibility record each blocked T-SIG; adding an active decision policy to an otherwise valid action request permitted T-ACT. These results establish only that the reference implementation enforces the declared rules under constructed

conditions. They neither estimate error rates in clinical records nor demonstrate that the rules improve patient outcomes.

5 Discussion

Personalized monitoring exposes the problem directly: a familiar baseline comparison can appear more clinically authoritative than it is. A system may retain excellent data provenance yet still let support migrate from a measurement to a stronger output. By making a transition the target of support, the method separates data origin from claim authorization and makes that otherwise implicit assumption testable.

The method is compatible with incremental implementation. An existing system can retain its FHIR resources, generic provenance graph, clinical rules, and assurance case. The additional profile attaches support categories and transition identifiers to those objects. A first deployment could use the validator in audit mode: it would block release only in a test environment, record missing support, and allow teams to inspect whether existing alerts depend on support that is only attached elsewhere. This is especially relevant where patient-generated data must be integrated into new workflows with explicit responsibility for review and management [14]. Later work could evaluate whether the method improves review, correction, or communication.

Several limitations are material. The source categories simplify the variety of clinical reasoning and do not constitute a complete ontology of evidence. The model does not determine whether a compatibility record is substantively correct; it requires that such a record exists, is versioned, and is valid. The fixtures were constructed to exercise declared boundary conditions. They are not representative of production data and cannot establish clinical safety, usability, or effectiveness. No FHIR implementation guide, deployed system, independent coding study, or human-participant evaluation is reported.

Those limitations identify a prospective validation agenda. An implementation study should test whether independent reviewers agree on claim types, source categories, scope, and compatibility. A system study should measure whether the profile identifies unsupported promotions in real workflows without creating unmanageable review burden. A clinical study would need to distinguish the accuracy of a detector, the validity of an interpretation, and the utility of a resulting policy. The present method supplies a disciplined representation for those later tests; it does not substitute for them.

6 Conclusion

Provenance-constrained claim transitions provide a compact conformance method for longitudinal digital-health systems. A personal baseline may support a bounded deviation. A stronger output must show its additional, transition- specific support rather than inherit authority from records elsewhere in the system. The reference validator demonstrates this property on predeclared, synthetic cases. Its value is not that it makes clinical decisions. Its value is that it makes the conditions for releasing a stronger claim inspectable, testable, and revisable.

Declarations

Competing interests

The author declares no competing interests.

Ethics approval and consent to participate

Not applicable. This technical study uses constructed examples and synthetic fixtures only; it includes no human participants, personal data, animals, or biological material.

Clinical trial number

Clinical trial number: not applicable.

Consent for publication

Not applicable.

Use of generative artificial intelligence

Generative AI tools were used for drafting and editorial support. The author reviewed and verified all content and accepts responsibility for the final manuscript.

Funding

No specific funding was received for this work.

Author Contribution

M.R.M. conceived the study, developed the methodology, designed and implemented the reference validator, conducted the formal analysis and investigation, prepared the visualizations, wrote the original draft, revised the manuscript, and managed the project. M.R.M. reviewed and approved the final manuscript.

Data Availability

All data generated or analysed during this study are included in this article and its supplementary information files.

References

1. Björnsson B, Borrebaeck C, Elander N, et al. Digital twins to personalize medicine. *Genome Medicine*. 2020;12:4. https://doi.org/10.1186/s13073-019-0701-3
2. Corral-Acero J, Margara F, Marciniak M, et al. The digital twin to enable the vision of precision cardiology. *European Heart Journal*. 2020;41:4556–4564. https://doi.org/10.1093/eurheartj/ehaa159
3. Kamel Boulos MN, Zhang P. Digital twins: from personalised medicine to precision public health. *Journal of Personalized Medicine*. 2021;11:745. https://doi.org/10.3390/jpm11080745
4. Health Level Seven International. *FHIR Release 5: Provenance*. 2023. https://hl7.org/fhir/R5/provenance.html
5. Moreau L, Missier P, eds. *PROV-DM: The PROV Data Model*. W3C Recommendation; 2013. https://www.w3.org/TR/prov-dm/
6. Health Level Seven International. *FHIR Release 5: Clinical Reasoning Module*. 2023. https://hl7.org/fhir/R5/clinicalreasoning-module.html
7. Object Management Group. *Structured Assurance Case Metamodel, Version 2.3*. 2023. https://www.omg.org/spec/SACM/2.3
8. Coskun A, Sandberg S, Unsal I, Serteser M, Aarsand AK. Personalized reference intervals: from theory to practice. *Critical Reviews in Clinical Laboratory Sciences*. 2022;59:501–516. https://doi.org/10.1080/10408363.2022.2070905
9. Roshan D, Ferguson J, Pedlar CR, et al. A comparison of methods to generate adaptive reference ranges in longitudinal monitoring. *PLOS ONE*. 2021;16:e0247338. https://doi.org/10.1371/journal.pone.0247338
10. Curcin V, Fairweather E, Danger R, Corrigan D. Templates as a method for implementing data provenance in decision support systems. *Journal of Biomedical Informatics*. 2017;65:1–21. https://doi.org/10.1016/j.jbi.2016.10.022
11. Sembay MJ, Macedo DDJ, Pioli Júnior L, Braga RM, Sarasa-Cabezuelo A. Provenance data management in health information systems: a systematic literature review. *Journal of Personalized Medicine*. 2023;13:991. https://doi.org/10.3390/jpm13060991
12. Taber P, Radloff C, Del Fiol G, Staes C, Kawamoto K. New standards for clinical decision support: a survey of the state of implementation. *Yearbook of Medical Informatics*. 2021;30:159–171. https://doi.org/10.1055/s-0041-1726502
13. Singh J, Cobbe J, Norval C. Decision provenance: harnessing data flow for accountable systems. *IEEE Access*. 2019;7:6562–6574. https://doi.org/10.1109/ACCESS.2018.2887201
14. Sittig DF, Boxwala A, Wright A, et al. Patient-centered clinical decision support challenges and opportunities identified from workflow execution models. *Journal of the American Medical Informatics Association*. 2024;31:1682–1692. https://doi.org/10.1093/jamia/ocae138
15. Riahi V, Diouf I, Khanna S, Boyle J, Hassanzadeh H. Digital twins for clinical and operational decision-making: scoping review. *Journal of Medical Internet Research*. 2025;27:e55015.

https://doi.org/10.2196/55015

16. Agu NN, Keshan N, Chari S, et al. G-PROV: provenance management for clinical practice guidelines. In: *Proceedings of the 2019 Workshop on Semantic Web for Medical Data and Applications*. CEUR Workshop Proceedings; 2019:68–75.
17. Stoldt J-P, Weber JH. Provenance-based trust model for assessing data quality during clinical decision making. In: *2021 IEEE/ACM 3rd International Workshop on Software Engineering for Healthcare*. IEEE; 2021. https://doi.org/10.1109/SEH52539.2021.00012
18. Gierend K, Waltemath D, Ganslandt T, Siegel F. Traceable research data sharing in a German medical data integration center with FAIR principles: proof-of-concept study. *JMIR Formative Research*. 2023;7:e50027. https://doi.org/10.2196/50027
19. Strasberg HR, Rhodes B, Del Fiol G, et al. Contemporary clinical decision support standards using HL7 FHIR. *Journal of the American Medical Informatics Association*. 2021;28:1796–1806. https://doi.org/10.1093/jamia/ocab070
20. Thiess H, Del Fiol G, Malone DC, et al. Coordinated use of HL7 standards to improve clinical decision support implementation: a case study. *International Journal of Medical Informatics*. 2022;162:104749. https://doi.org/10.1016/j.ijmedinf.2022.104749
21. Isabelle M, Ip IK, Bakhtin M, et al. Assessment of three standards-based clinical decision-support tools: a proof-of-concept study. *JAMIA Open*. 2025;8:ooaf085. https://doi.org/10.1093/jamiaopen/ooaf085
22. Blatter TU, Witte H, Fasquelle-Lopez J, Nakas CT, Raisaro JL, Leichtle AB. The BioRef infrastructure: a federated approach to personalised reference intervals. *Journal of Medical Internet Research*. 2023;25:e47254. https://doi.org/10.2196/47254
23. Armeni P, Polat I, De Rossi LM, et al. Digital twins in healthcare: is it the beginning of a new era of evidence-based medicine? A critical review. *Journal of Personalized Medicine*. 2022;12:1255. https://doi.org/10.3390/jpm12081255
24. Calcaterra V, Guardamagna L, Gatti A, et al. Digital twins in healthcare: a systematic review of current applications, frameworks, and future directions. *Digital Health*. 2026;12. https://doi.org/10.1177/20552076261415934

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DiscoverInformaticsSupplementaryArtefactsv320261107.zip