Journal of Advances in Developmental Research (IJAIDR)
E-ISSN: 0976-4844 • Website: www.ijaidr.com • Email: editor@ijaidr.com

Designing Interoperable Healthcare *Data* Platforms for Population Health at National Scale

Sushrut Naik

Independent Researcher
8280 164th Avenue NE Unit 407 Redmond WA 98052
sushrut.s.naik@gmail.com

Abstract:
Healthcare data across the United States continues to be siloed. Since 2009, there has been $36 billion spent in the form of EHR incentives. However, the promise of a connected health ecosystem has not materialized yet. This paper explores three core building blocks for population health platforms at a national scale: unified data architecture, FHIR-driven interoperability, and federated healthcare data networks. Using insights and data from real-world deployments involving millions of patients, it suggests that the challenges the industry faces aren't due to a lack of technology—but rather stem from deeper structural issues and misaligned incentives. To address this, the paper introduces a practical, layered framework that looks beyond just the technical infrastructure. It also emphasizes the importance of governance, outlining what's needed to ensure interoperability works effectively in real-world settings.

Keywords: Healthcare interoperability, FHIR R4, population health, federated data networks, data unification, health information exchange, clinical data integration, real-world evidence

1. Introduction

In 2024, a 43-year-old woman in rural Mississippi was admitted for diabetic ketoacidosis. She managed her type 2 diabetes across three providers in two states—none of which could see each other's medical records. The ER physician that attended to her could not view any medication history, no recent A1C numbers and no knowledge of a community health worker flagging the woman as high-risk patient two weeks earlier. The woman spent eleven days in the ICU with an estimated medical bill of $87,000. Every system involved was certified as a modern EHR, however still failed to provide the required information needed to take care and provide the right service. This was not a technology failure—it was an interoperability failure that plays out thousands of times daily [1].

Nobody in health IT leadership wants to noodle this question: Since the HITECH Act of 2009 if the industry has spent over $36 billion on EHR adoption, why does a physician in 2026 still have to fax a patient record? [1]

Journal of Advances in Developmental Research (IJAIDR)
E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com

The HITECH Act succeeded in one thing: by 2021, over 96% of non-federal acute care hospitals had adopted a certified EHR. This is an increase of 68% from 2011 [2]. The 21st Century Cures Act and the ONC Information Blocking Rule attempted to force the issue [3], yet a 2023 AHA survey found only 40% of hospitals reported routine electronic exchange of clinical summaries with outside organizations [4]. COVID-19 laid bare the consequences—fragmented data contributed to delayed responses associated with 40,000–70,000 excess deaths in the first wave alone [5]. Population health at national scale demands a fundamentally different data architecture.

2. Data Unification Architecture

Most healthcare integration projects start with a naïve assumption - that is to build enough point-to-point interfaces to solve the problem. However, this assumption fails when a 500-bed hospital that runs 100–300 discrete clinical applications [7]. The number of interfaces needed to operationalize the hospital is exponential and is also expensive to maintain. With every enhancement, vendor change, or regulatory mandates, there is cascading effect on these brittle integrations leading to more expenses and issues.

Rather than connecting systems to each other, The data unification approach connects all systems to a shared canonical data layer. Think of it as building roads between every pair of cities versus building an interstate highway system. Every consumer reads from a data source that is published to the canonical layer; Figure 1 illustrates this layered architecture, inspired by real-world implementations such as the National COVID Cohort Collaborative (N3C) (75+ sites, 18.4 million patients) [6] and the Observational Health Data Sciences and Informatics (OHDSI) network (800+ million records across 30+ countries) [20].

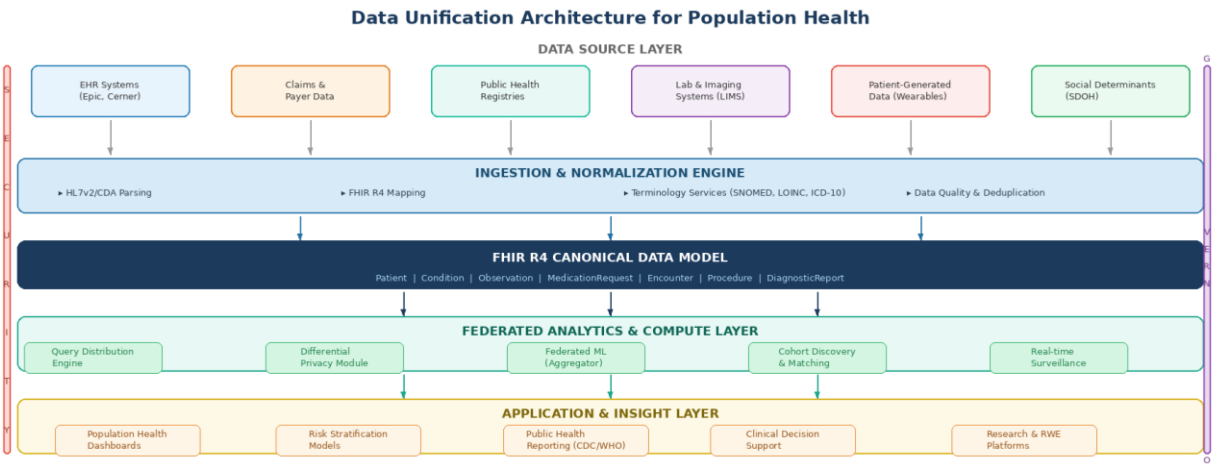

Figure 1: Layered Data Unification Architecture for National-Scale Population Health

The architecture comprises five layers: a Data Source Layer accepting EHR, claims, registry, lab, wearable, and SDOH inputs in any format (HL7v2, C-CDA, X12, FHIR); an Ingestion and Normalization Engine handling format parsing, terminology mapping across SNOMED CT, LOINC, ICD-10, and

RxNorm, plus patient matching (duplicate records exist in 8–12% of health system databases [8]); the FHIR R4 Canonical Data Model as the unified representation; a Federated Analytics Layer for distributed querying; and an Application Layer delivering dashboards, risk models, public health reporting, and research platforms.

3. FHIR-Based Interoperability

FHIR (Fast Healthcare Interoperability Resources), developed by HL7 International, was designed from the ground up for the web era: RESTful APIs, JSON/XML serialization, OAuth 2.0 security, and a resource-based model developers can actually work with [9]. HL7v2 (1987) was never designed for this; HL7v3/CDA was too complex. FHIR R4 reached normative status in 2019, and the CMS Interoperability Rule mandated FHIR R4 Patient Access APIs for Medicare Advantage, Medicaid, and CHIP payers by July 2021 [10]. All major EHR vendors now support FHIR R4 endpoints, and the SMART on FHIR framework has enabled a thriving third-party app ecosystem [11].

In our proposed architecture, FHIR R4 serves not just as an exchange format but also as the canonical data model. Core resources—Patient, Condition, Observation, Medication Request, Encounter, Procedure, DiagnosticReport are all constrained by US Core profiles for American implementations [12]. The benefit here is very real and easy to quantify. In a network of 50 organizations, you go from managing 2,450 one-off connections between systems to just 50 mappings into a shared, canonical layer. That's a 98% reduction in integration complexity—and a huge simplification in how everything connects.

FHIR is not a silver bullet. In practice, implementations still look different from one organization to another, and mapping of clinical terminology continues to be both complex and costly. On top of that, FHIR was originally built for real-time, patient-level data exchange—not large-scale population analytics [13].

Even with these limitations, it remains the strongest foundation we have today, and its adoption is only gaining speed.

4. Federated Healthcare Data Networks

It's easy to see why many organizations try to solve data fragmentation by pulling everything into a single, centralized warehouse—but in reality, that approach creates as many problems as it solves. Healthcare data isn't just any data. It's protected under laws like Health Insurance Portability and Accountability Act (HIPAA), along with various state regulations and patient consent requirements.

Moving sensitive patient data across systems and boundaries quickly introduces compliance risks and complicated governance challenges. Figure 2 highlights this tension by comparing the centralized model with a more flexible, federated approach.

Journal of Advances in Developmental Research (IJAIDR)
E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com

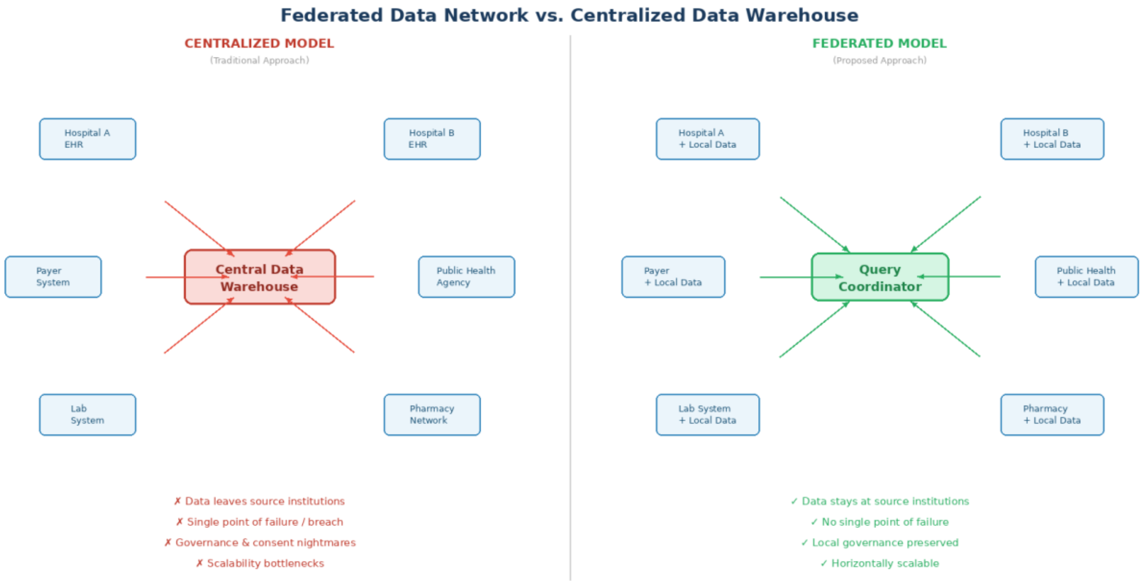

Figure 2: Centralized vs. Federated Approaches to Healthcare Data Integration

Under a federated model, data never leaves the stewardship of its home institution. Queries can be standardized so that they are dispatched to consenting nodes, run locally against FHIR-conformant stores, and only the aggregation of results needs to be returned. This is operational at scale: PCORnet has 67 sites and 110 million patients [14]; TriNetX links together more than 250 organizations globally for real-time identification of cohorts—all without pooling patient-level data. New techniques reinforce this: differential privacy introduces calibrated noise to thwart re-identification while retaining statistical relevance, and federated learning enables training of an ML model by exchanging gradients rather than sensitive data — already being implemented by companies such as Owkin for oncology research [15].

5. The Silo Problem

Despite mature technology and regulatory mandates only, 23% of the healthcare organizations ranked their interoperability as "mature" or "advanced" [16]. The reasons, unfortunately, are not a mystery: EHR vendors profit from data lock-in (the switching cost for large networks exceeds $500M), healthcare systems would rather not share with rivals, and the governance infrastructure to support sharing among multiple parties—legal agreements, consent management, breach liability—is underdeveloped. Until the news industry deals with these structural incentives, no technical innovation can bridge them.

6. Evidence Base

Fragmentation has a well-documented clinical and economic impact. In fact, HIE can save $77.8 billion per year with decreased repeated tests and adverse effects [17]. The cost of patient matching failures alone was $1,950 per admission—approximately $6 billion annually [18]. A 2023 study published in JAMIA aggregating data of 12 high-risk health systems demonstrated that FHIR-based, unified risk models outperformed single-site derived population-level models by between 18 and 24%, with the C-statistic for predicting thirty-day readmissions improving from a mean C-statistic of: 0.68 → 0.82 [19]. 4 N3C showed that pooled multisite data could reveal treatment effects obscured in any individual site analysis [6].

Metric	Value	Source
Annual cost of data fragmentation	$77.8 billion (U.S.)	Walker et al. [17]
Cost per matching failure	$1,950 per admission	Pew Trusts [18]
Hospitals w/ mature interop.	23%	KLAS Research [16]
Readmission prediction gain	18–24% (C: 0.68→0.82)	Rodriguez et al. [19]
N3C network coverage	75+ sites, 18.4M patients	Haendel et al. [6]
PCORnet network coverage	67 sites, 110M patients	Fleurence et al. [14]

7. Conclusion

Such a platfrom could be based of this paper framing example. Each of the three pillars is individually mature and collectively sufficient. The layered architecture has been proven in real world deployments such as N3C (18.4M patients), PCORnet (110M patients), and OHDSI (800M+ records) [6, 14, 20]. The main finding is blunt: the technology has existed for years. This ongoing massive failure is a governance and economic incentives problem dressed in technical clothing.

There are three developments that will determine the path ahead. As TEFCA matures operationally, national governance will form for this structure where Qualified Health Information Networks serve as trusted intermediaries. Genomic and multi-omic integration will require FHIR Genomics extensions and petabyte scale compute. It will also significantly reduce the manual effort driving bottlenecks in interoperability projects by applying large language models to clinical data normalization. The question is not whether national-scale platforms are technically possible. They are. The question is whether the industry has the collective will to do so.

REFERENCES:

[1] Blumenthal, D. (2010). Launching HITECH. New England Journal of Medicine, 362(5), 382–385.
[2] Office of the National Coordinator for Health IT (ONC). (2022). Non-Federal Acute Care Hospital Health IT Adoption and Use: 2008–2021. HealthIT.gov Data Brief.
[3] 21st Century Cures Act, Pub. L. No. 114–255 (2016); ONC Information Blocking Final Rule, 85 Fed. Reg. 25642 (2020).
[4] American Hospital Association. (2023). 2023 AHA Survey on Health Information Exchange. Chicago, IL: AHA.
[5] Woolf, S. H., Chapman, D. A., & Lee, J. H. (2022). Excess deaths from COVID-19 and other causes in the US. Health Affairs, 41(4), 531–540.
[6] Haendel, M. A., Chute, C. G., Bennett, T. D., et al. (2021). The National COVID Cohort Collaborative (N3C). JAMIA, 28(3), 427–443.
[7] KLAS Research & CHIME. (2022). Clinical IT Integration: Application Inventory Benchmarking. Orem, UT: KLAS.
[8] ONC. (2021). Patient Matching in Health Information Exchanges. ONC Issue Brief.
[9] Bender, D., & Sartipi, K. (2013). HL7 FHIR: An agile and RESTful approach to healthcare information exchange. IEEE CBMS, 326–331.

[10] CMS. (2020). Interoperability and Patient Access Final Rule (CMS-9115-F). Federal Register, 85(85), 25510–25640.
[11] Mandel, J. C., et al. (2016). SMART on FHIR: A standards-based, interoperable apps platform for EHRs. JAMIA, 23(5), 899–908.
[12] HL7 International. (2023). US Core Implementation Guide (STU 6.0). HL7 FHIR.
[13] Mandl, K. D., & Mandel, J. C. (2022). FHIR Bulk Data Access: Enabling population-level data access. J Biomed Inform, 132, 104133.
[14] Fleurence, R. L., et al. (2014). Launching PCORnet, a national patient-centered clinical research network. JAMIA, 21(4), 578–582.
[15] Rieke, N., et al. (2020). The future of digital health with federated learning. NPJ Digital Medicine, 3(1), 119.
[16] KLAS Research. (2024). Interoperability Maturity: Where Healthcare Organizations Stand Today. Orem, UT: KLAS.
[17] Walker, J., et al. (2005). The value of health care information exchange and interoperability. Health Affairs, 24(Suppl1), W5-10–W5-18.
[18] Pew Charitable Trusts. (2022). Enhanced Patient Matching Is Critical to Achieving Full Promise of Digital Health Records. Pew Issue Brief.
[19] Rodriguez, F., et al. (2023). Multi-site FHIR-based analytics platform for risk stratification. JAMIA, 30(7), 1182–1191.
[20] Hripcsak, G., et al. (2015). OHDSI: Opportunities for observational researchers. Stud Health Technol Inform, 216, 574–578.