

INSURANCE DATA-DRIVEN DISEASE SURVEILLANCE

Integrating NPHIES and Health Insurance Data for Population-Level Disease Surveillance, Risk Stratification and Preventive Care in Saudi Arabia

Salah Ali Ibrahim Ramadan

Published October 2026

Abstract

Saudi Arabia's digital health transformation has created an opportunity to move from episodic, provider-centred information exchange towards population-level intelligence that supports prevention. The National Platform for Health and Insurance Exchange Services (NPHIES) can contribute to this shift because claims, eligibility, authorisation and reimbursement data describe longitudinal patterns of healthcare use across insured populations. Yet administrative data alone cannot establish clinical need with sufficient precision. This review examines how NPHIES-linked health insurance data could be integrated with electronic health records, laboratory results, pharmacy information and patient-generated data to support three connected functions: disease surveillance, risk stratification and preventive care. A structured integrative review of peer-reviewed literature published from 2020 to 2025 was undertaken, with emphasis on Saudi digital health, claims-EHR linkage, population health analytics, chronic disease prediction, interoperability and algorithmic assurance. Thirty DOI-resolvable sources were synthesised thematically. The evidence indicates that claims data are valuable for population coverage, longitudinal utilisation histories and service-cost patterns, whereas clinical data contribute physiological detail, disease severity and treatment context. Their integration can improve case finding, reveal screening and care gaps, and enable risk models that identify people likely to develop chronic disease or experience avoidable acute-care use. However, value depends on identity resolution, coding quality, semantic interoperability, calibration, subgroup fairness, privacy protection and a closed-loop process that converts risk signals into documented preventive action. A Saudi implementation framework is proposed in which NPHIES becomes a governed population-health data layer rather than only a transaction exchange.

Keywords: NPHIES; health insurance data; population health; disease surveillance; risk stratification; preventive care; interoperability; Saudi Arabia; Vision 2030

1 Introduction

Saudi healthcare is undergoing a structural transition in which digital platforms, data exchange and new models of financing are expected to support coordinated and preventive care. NPHIES is a central component of this transition because it provides a national mechanism for health and insurance information exchange and increasingly standardises interactions among payers and providers [1]. This infrastructure sits within a wider reform programme that seeks to improve access, quality, efficiency and sustainability while shifting the health system towards value, prevention and digitally enabled delivery [2]. The more consequential question is whether exchanged data can be transformed into population-level intelligence that changes care before avoidable deterioration occurs [3].

Claims, authorisations, eligibility records and pharmacy transactions create longitudinal traces of diagnoses, procedures, prescriptions, service intensity, provider contacts and cost. They can therefore reveal recurring care pathways across large populations even when individual encounters occur at different organisations. Saudi work on Internet-of-Things and big-data approaches for chronic disease monitoring has already shown the strategic importance of linking continuous or repeated health signals to proactive management [4]. More recent Saudi studies demonstrate that routinely collected clinical data can support population-oriented prediction of type 2 diabetes risk and age of disease onset [5, 6]. Emerging analyses using Saudi insurance claims have also reported proof-of-concept systems for detecting diabetes earlier and stratifying the risk of diabetic retinopathy [7, 8]. These studies do not by themselves establish readiness for national deployment, but they illustrate why claims data should be treated as an analytical asset rather than solely as a reimbursement record.

The principal limitation is that insurance data are generated for administrative purposes. A diagnosis code may reflect billing, rule-out assessment or a chronic condition carried forward; a rejected claim may represent an administrative defect rather than absence of care; and many clinically important variables, including laboratory values, blood pressure, body mass index, disease severity and social context, may be missing. Conversely, electronic health records (EHRs) are clinically richer but may be fragmented by provider, inconsistent in structure, and incomplete when patients receive care elsewhere. Patient-facing digital systems add another perspective by capturing engagement, self-reported outcomes and the practical ability to act on prevention, although adoption and sustained use vary [9–11]. The analytical problem is therefore not to choose one source but to combine complementary sources while preserving their provenance and limitations.

The objective is not only to predict an individual outcome at the point of care, but to understand distributions of need across covered populations, identify emerging disease burden, detect gaps in screening or follow-up, segment people by actionable risk, and allocate preventive resources before hospital demand materialises. Saudi research has explicitly identified opportunities for population-health data analytics, alongside persistent challenges in interoperability, governance, data quality and workforce capability [12]. International experience with national patient portals and patient-accessible records similarly indicates that digital infrastructure creates value when it supports continuity, understandable information and meaningful interaction rather than passive data availability [13–19].

This review therefore evaluates how NPHIES and health insurance data can be integrated with clinical and patient-level information to create a governed population-health intelligence capability for Saudi Arabia. The paper connects data architecture with clinical purpose: surveillance should identify where risk is accumulating; stratification should distinguish who needs which level of intervention; and preventive care should close the loop by acting on those signals. This linkage is essential because technically accurate prediction has limited public-health value when no service pathway is responsible for responding to it.

2 Aim, objectives and review questions

The aim of this review is to develop an evidence-based framework for integrating NPHIES-linked insurance and clinical data to support population-level disease surveillance, risk stratification and preventive care in Saudi Arabia. Five objectives guide the synthesis. First, the review characterises the complementary analytical value and limitations of claims, EHR, pharmacy and patient-generated

data. Second, it examines how linked longitudinal data can strengthen chronic-disease surveillance and case finding. Third, it evaluates the design and assurance requirements of models intended to predict disease onset, complications or preventable healthcare use. Fourth, it identifies mechanisms for translating risk tiers into preventive interventions rather than leaving prediction as an isolated technical output. Fifth, it proposes governance and implementation measures suitable for a national Saudi setting.

Three review questions follow from these objectives. RQ1 asks which data combinations and interoperability mechanisms are required to convert NPHIES transactions into reliable population-health signals. RQ2 asks how risk stratification should be designed, validated and monitored when its purpose is to prevent avoidable deterioration or hospital use. RQ3 asks how predictive outputs can be embedded in primary care, digital engagement and care-management workflows so that measurable preventive action follows from identification of risk.

3 Review methodology

A structured integrative review was selected because the topic spans health informatics, insurance analytics, population health, chronic-disease prediction, interoperability and implementation science. The evidence window was restricted to 2020-2025 to reflect current data architectures and contemporary predictive methods. Searches were organised around combinations of terms covering NPHIES, Saudi digital health, insurance claims, EHR linkage, population surveillance, chronic disease, risk stratification, preventive care, interoperability, FHIR, data quality and fairness. Scopus, Web of Science and PubMed were used as the principal scholarly sources, with backward citation checking applied to highly relevant papers.

Eligibility was based on conceptual rather than geographical similarity alone. Saudi studies were prioritised for claims about national context, health-system transformation, chronic-disease prediction and digital adoption. International evidence was retained where it addressed transferable mechanisms such as claims-EHR linkage, population risk stratification, record linkage, semantic interoperability, preventive-service surveillance or algorithmic fairness. Included sources had to be peer reviewed, published within the stated period, directly relevant to at least one review question, and traceable through a DOI. Conference abstracts were retained only where they supplied uniquely relevant Saudi claims-based evidence and were interpreted as preliminary rather than definitive [7, 8]. Thirty sources met the final relevance and traceability criteria.

Synthesis proceeded through an evidence-to-implementation matrix. Each source was coded under one or more domains: data value, surveillance, risk prediction, preventive intervention, interoperability, linkage, data quality, privacy or equity. A second coding pass identified recurring operational requirements, including longitudinal identity resolution, phenotype definition, calibration, subgroup performance, action thresholds, feedback loops and outcome measurement. Contradictory or limiting evidence was retained. In particular, the review distinguished predictive performance from clinical utility and distinguished data availability from data fitness for a specific analytical purpose. No primary patient data were analysed, and no synthetic effect estimates were calculated.

4 NPHIES as a population-health data substrate

NPHIES was designed to enable standardised exchange across the Saudi health-insurance ecosystem, but the same transactions can support a secondary population-health function if they are governed

appropriately [1]. Eligibility data define covered populations over time; claims identify diagnoses, procedures and encounters; authorisations reveal intended high-cost or specialised services; and pharmacy transactions can indicate treatment initiation, refill patterns and therapeutic change. Taken together, these records create a longitudinal administrative representation of care that is difficult to construct from isolated provider systems. For surveillance, breadth and continuity may be more valuable than clinical depth because repeated signals can expose patterns that single encounters cannot.

A person with repeated diabetes-related claims, escalating medication complexity, emergency visits and ophthalmology referrals has a recognisable trajectory even if laboratory values are absent. Likewise, geographically aggregated claims can reveal variation in screening, chronic-disease follow-up or acute-care utilisation across regions, provider networks or insured groups. Population analytics can use these patterns to identify cohorts for further verification rather than treating each code as a definitive clinical truth. This distinction is important because claims are observations of billed care, not complete observations of health.

Saudi diabetes risk models have used national health-information data and primary-care validation to identify individuals at elevated risk before diagnosis, demonstrating the feasibility of using routinely captured variables for scalable screening [5]. Other Saudi machine-learning work has shown that clinical, biochemical and behavioural features can help predict the timing of diabetes onset [6]. Insurance-derived models extend the concept by exploiting longitudinal utilisation at population scale [7, 8]. A combined architecture can therefore use claims for wide-net detection and EHR or laboratory data for refinement. The objective is not to force every source into one undifferentiated database, but to create a longitudinal patient layer in which each observation retains its source, timestamp and reliability.

Digital-health adoption studies in Saudi Arabia show that access, perceived value and professional acceptance influence whether digital tools become part of routine care [9–11]. Patient portals and accessible records can support continuity and engagement, but international evidence shows that usability, trust, workflow integration and perceived usefulness determine whether access translates into meaningful participation [13–19]. For population prevention, this matters because a risk model cannot improve outcomes if high-risk people cannot be reached, do not understand the recommended action, or encounter friction when attempting to obtain screening or follow-up.

The proposed design treats NPHIES as the population spine, not as the sole source of truth. Administrative data identify patterns at scale; clinical data establish phenotypic credibility; pharmacy and patient-generated information refine adherence and behaviour; and linkage converts fragmented observations into a longitudinal record suitable for population analysis. This architecture directly supports the Vision 2030 emphasis on integrated, proactive and digitally enabled care [2].

Table 1. Complementary data layers for NPHIES-enabled population health analytics.

Data layer	Population signals	Analytical value	Primary limitation	Quality / governance control
NPHIES claims, eligibility and authorisations	Diagnoses, procedures, encounters, covered periods, service intensity	Broad longitudinal utilisation and cohort detection across insured care	Billing purpose; coding incentives; rejected or delayed transactions	Source provenance, denominator rules, coding audits, latency monitoring

Data layer	Population signals	Analytical value	Primary limitation	Quality / governance control
Clinical EHR and laboratory data	Vitals, laboratory values, clinician diagnoses, disease severity	Phenotype confirmation and risk refinement	Fragmentation across providers; variable structure and completeness	Terminology mapping, encounter reconciliation, timestamp controls
Pharmacy information	Initiation, dispensing, refill continuity, therapeutic change	Medication pathway and adherence proxies	Dispensing does not prove consumption	Product mapping, days-supply validation, duplicate handling
Patient-facing and reported data	Symptoms, outcomes, engagement, response to outreach	Behavioural and functional context for prevention	Digital exclusion and self-report bias	Multi-channel outreach, consent rules, usability and completeness checks
Linked longitudinal record	Cross-source trajectory, screening gaps, care transitions	Surveillance phenotypes, dynamic risk tiers and outcome follow-up	Linkage error and selection bias	Privacy-preserving identity resolution, linkage-rate reporting, audit trail

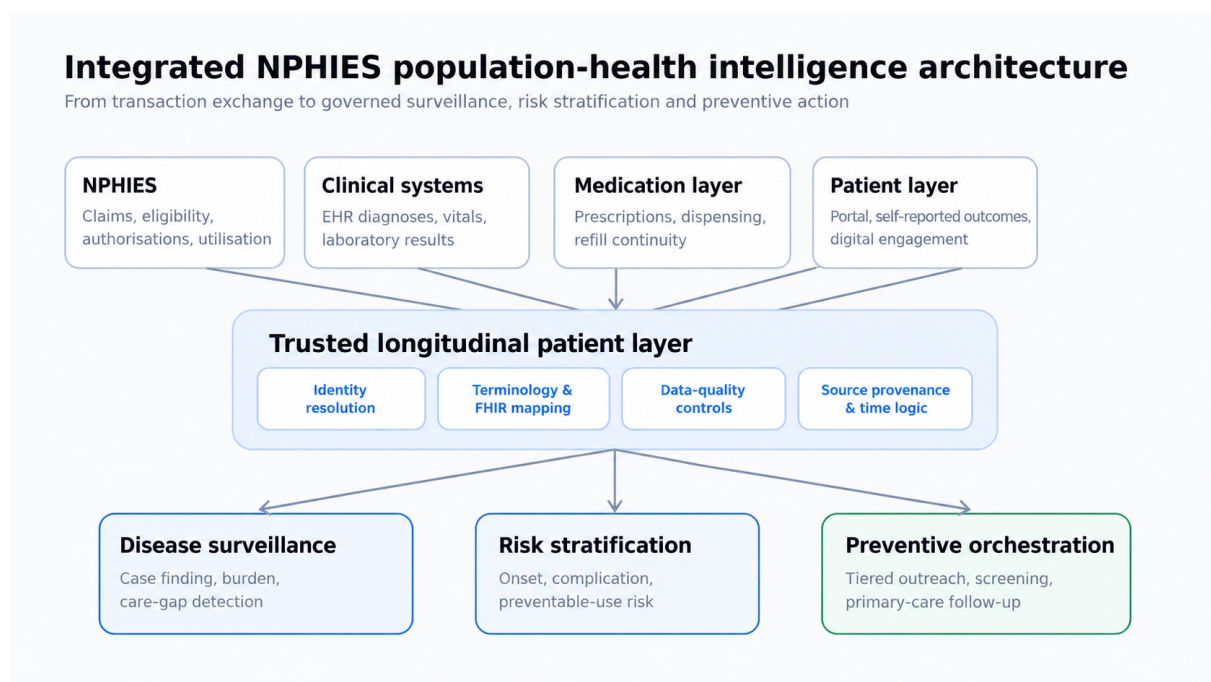


Figure 1. Integrated NPHIES population-health intelligence architecture. The design separates source systems, a governed longitudinal patient layer and three population-health functions: surveillance, risk stratification and preventive orchestration.

5 Population-level disease surveillance

Population disease surveillance traditionally relies on registries, surveys, laboratory reporting and aggregated provider statistics. These sources remain important, but linked administrative and clinical data can add timeliness and repeated observation. Large-scale EHR networks have demonstrated that routine data can identify changing patterns in preventive-service use and new chronic-disease diagnoses across millions of people [23]. The same principle can be applied to NPHIES: the platform

can support near-continuous observation of care-seeking, diagnostic coding, medication initiation, screening activity and acute-care utilisation among covered populations.

Surveillance should begin with transparent, clinically defensible phenotypes rather than opaque prediction. For diabetes, a phenotype might combine repeated diagnostic codes, antidiabetic medication, glycated-haemoglobin testing and relevant primary-care encounters. For hypertension, repeated diagnosis plus antihypertensive therapy and documented blood-pressure evidence would provide greater confidence than a single claim. Each phenotype should distinguish confirmed disease, probable disease, screening need and uncertain status. This tiering prevents administrative codes from being mistaken for ground truth while still using them to prioritise verification.

The surveillance layer should also measure care gaps. Population management literature shows that risk-stratification programmes are most useful when they are paired with targeted interventions and primary-care action [20]. Diabetes population-health initiatives likewise emphasise segmentation, multidisciplinary management and outcome tracking rather than prediction alone [22]. NPHIES data can reveal who has not completed recommended monitoring, who repeatedly uses emergency services without stable follow-up, who has medication interruptions, and where referral pathways appear incomplete. These are actionable surveillance signals because they describe failures or discontinuities in prevention, not merely prevalence.

Geographical and organisational comparison requires caution. Differences in claim rates may reflect genuine disease burden, provider access, coding practice, benefit design or care-seeking behaviour. Risk-adjusted rates should therefore be accompanied by data-quality indicators and denominator definitions. A region with apparently low diabetes prevalence may instead have lower screening intensity or incomplete data capture. Surveillance dashboards should display uncertainty and source completeness so that operational leaders can distinguish a possible epidemiological signal from an information artefact.

The greatest value arises when surveillance is cyclical. New claims and clinical observations update cohort membership; public-health and primary-care teams investigate unexpected changes; verified findings improve phenotype definitions; and interventions are measured against subsequent utilisation. This produces a learning surveillance system in which data collection and service improvement reinforce one another. It also creates the foundation for risk stratification because reliable phenotypes and longitudinal histories are prerequisites for meaningful prediction.

6 Risk stratification for chronic disease and preventable utilisation

Its purpose is not to label people as permanently high or low risk, but to estimate an actionable probability within a defined time horizon. Reviews of population risk-stratification tools show wide variation in predictors, target populations and interventions, with stronger programmes linking risk tiers to specific primary-care responses [20]. A separate review of algorithms predicting healthcare utilisation cautions that many tools are developed in settings that differ from their intended deployment and may not demonstrate adequate transportability or clinical impact [21].

“Chronic disease risk” is too broad for deployment. Models should predict outcomes such as new type 2 diabetes within twelve months, diabetic retinopathy within two years, emergency admission for heart failure within ninety days, or repeated potentially avoidable admission within six months. Target definition determines which variables are legitimate, when prediction should occur and which intervention remains feasible. Saudi diabetes studies illustrate the difference between identifying

high-risk undiagnosed individuals [5], estimating age of onset [6], detecting diabetes from claims [7] and predicting a specific complication from claims [8].

Claims provide history of service use, diagnoses, procedures, pharmacy activity and cost; EHRs contribute laboratory values, vital signs and clinician-documented conditions; patient-reported information can add behaviour, symptoms and barriers. Research linking EHR and claims data shows that combination can improve information completeness, although linkage may reduce sample size or introduce selection differences that must be quantified [29]. The objective should be parsimonious predictive value rather than maximal feature volume. Variables that are unstable across payers, susceptible to coding incentives or unavailable at the intended decision time can make a model appear stronger during development while weakening deployment.

Area under the receiver-operating characteristic curve is useful, but population prevention also depends on calibration, positive predictive value at realistic thresholds, subgroup performance and decision consequences. A model can rank patients correctly yet systematically overestimate absolute risk, causing excessive outreach and unnecessary testing. Conversely, a highly specific threshold may miss people who would benefit from inexpensive preventive action. Thresholds should therefore be chosen with service capacity and intervention burden in mind. Low-cost digital reminders can justify a lower threshold than an intensive nurse-led care-management programme.

Algorithmic prediction can reproduce disparities when access, coding and historical utilisation differ by population group [28]. Reviews of fairness reporting in clinical risk models show that subgroup assessment remains inconsistent [27]. NPHIES models should therefore report calibration and error rates by sex, age, region, insurance characteristics and other legally and clinically appropriate groups, while avoiding crude assumptions that equal error rates necessarily represent equal need. Under-recorded populations may require additional clinical verification rather than statistical adjustment alone.

Coding practices, benefit structures, provider networks and population behaviour change, producing drift. Monitoring should track input shifts, prevalence shifts, calibration, intervention yield and downstream outcomes. Risk status should also expire or be recomputed because a patient successfully treated after an alert should not remain indefinitely classified by an obsolete prediction.

7 Translating risk into preventive care

Prediction becomes clinically meaningful only when it changes an intervention. This principle is evident across population-health management literature: risk stratification is most defensible when different risk levels trigger proportionate, evidence-based responses [20, 22]. In the Saudi context, NPHIES can help create that linkage because it contains signals about encounters and services while digital and primary-care channels can deliver follow-up. A population-health programme should therefore define the preventive pathway before deploying the model.

A low-risk tier receives routine prevention, education and standard screening reminders. An emerging-risk tier receives targeted digital outreach, confirmation of modifiable risk factors and accelerated screening. A high-risk tier receives primary-care review, laboratory assessment, medication reconciliation and personalised prevention planning. A complex-risk tier, characterised by multimorbidity or repeated acute-care use, receives multidisciplinary case management and coordinated specialty input. Movement between tiers should be dynamic and based on new information rather than administrative labels.

Studies of patient portals and mHealth show that willingness to use digital services depends on perceived usefulness, trust and integration with care [9–19]. Preventive outreach should therefore offer multiple routes: secure messaging, telephone contact, primary-care appointment, pharmacy engagement or case-manager follow-up. A non-response to an app notification is not evidence that prevention is unnecessary. It may indicate digital exclusion, low health literacy, competing priorities or uncertainty about the recommendation.

A model output should create a task with an accountable owner, a recommended action, a completion status and an outcome. For example, a high diabetes-risk signal may trigger laboratory testing; the result then confirms, downgrades or refines the risk state. A predicted retinopathy risk may trigger retinal screening, with attendance and findings feeding back into the model. Claims data can subsequently show whether preventive services occurred and whether emergency use or avoidable admissions changed, but clinical confirmation remains necessary when outcomes cannot be inferred reliably from billing data.

These include earlier diagnosis, improved screening completion, better disease control, fewer emergency visits, fewer avoidable admissions and reduced disparity in access. Predictive accuracy should be reported as an enabling metric rather than the final endpoint. If a model identifies high-risk patients accurately but the health system lacks capacity to intervene, the programme has created information without prevention.

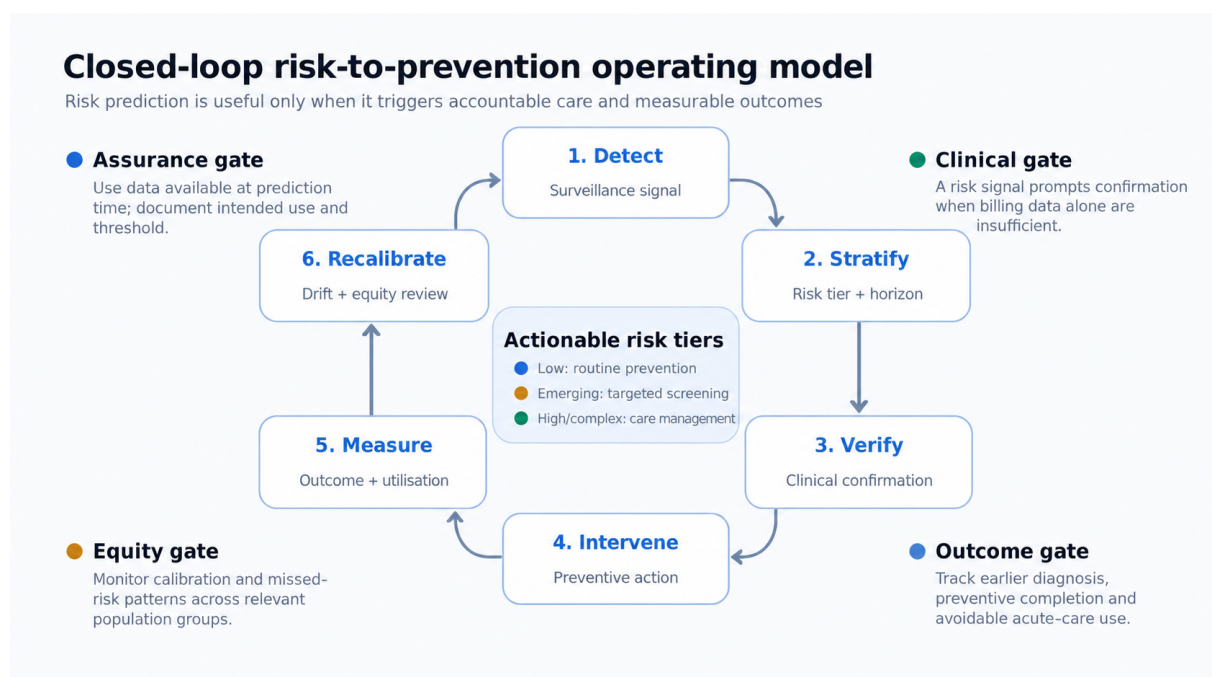


Figure 2. Closed-loop risk-to-prevention operating model. Prediction is treated as one step in a cycle that requires clinical verification, accountable intervention, outcome measurement, recalibration and equity assurance.

Table 2. Staged implementation and assurance framework for Saudi population-health analytics.

Stage	Core capability	Assurance gate	Key measures	Preventive-care output
1. Trusted data foundation	Identity resolution, provenance, terminology and quality controls	Coverage and linkage quality acceptable for intended use	Linkage rate, latency, missingness, duplicate identities	Reliable cohort denominator and longitudinal history

Stage	Core capability	Assurance gate	Key measures	Preventive-care output
2. Population surveillance	Validated phenotypes and care-gap detection	Clinical review of definitions and false-positive patterns	Prevalence, screening gaps, medication continuity, acute-care patterns	Priority cohorts for verification and outreach
3. Risk stratification	Condition-specific prediction with defined horizon and threshold	Calibration, subgroup performance and external/prospective validation	Calibration, PPV, sensitivity, drift, subgroup error	Dynamic low, emerging, high and complex risk tiers
4. Preventive orchestration	Routing of risk to primary care, digital outreach or case management	Named owner, task closure and escalation logic	Time to contact, completion, screening uptake, treatment initiation	Documented preventive action proportional to risk
5. Learning and equity assurance	Outcome feedback, model refresh and service redesign	Evidence that benefit persists across groups and time	Avoidable admissions, emergency use, outcome gaps, cost per prevented event	Continuously improving prevention pathway

8 Data integration, interoperability and linkage

The technical foundation of this model is reliable linkage across heterogeneous systems. Population health cannot be reconstructed safely by matching names or identifiers ad hoc. A governed identity-resolution service should link NPHIES transactions, provider EHRs, laboratory data, pharmacy records and patient-facing information while preserving source provenance. Privacy-preserving record-linkage research shows that administrative claims and EHR databases can be connected without unnecessary exposure of direct identifiers when linkage architecture is designed explicitly [25].

Interoperability must operate at both syntactic and semantic levels. FHIR-based bulk-data approaches demonstrate how standardised population extraction can support large-scale analytics rather than one-patient-at-a-time exchange [24]. More recent work on modular FHIR transformation pipelines shows how heterogeneous real-world clinical data can be converted into reusable, standard representations while retaining mapping logic and validation [30]. For NPHIES, a similar approach would allow claims concepts to coexist with clinical observations without pretending that they are semantically identical.

A common data layer should standardise core entities such as patient, encounter, diagnosis, procedure, medication, laboratory test, authorisation and organisation. Terminology mapping must be version controlled, because a code change can alter apparent incidence or model features. Temporal logic is equally important: the system should distinguish service date, claim submission, authorisation, result date and data-availability date so that models are trained only on information that would have been available at prediction time.

Linkage also introduces selection bias. Claims-EHR studies show that linked cohorts can differ from source populations because not every record can be matched and not every patient has comparable data density [29]. Data-quality dashboards should therefore report linkage rate, missingness, latency, duplicate identity rate and source coverage. These measures should accompany model performance so that a drop in apparent disease risk is not mistaken for improvement when it actually reflects loss of data capture.

9 Governance, privacy, data quality and equity

Insurance data can reveal sensitive patterns about chronic illness, reproductive health, mental health and high-cost treatment. Patient trust cannot be assumed simply because data are held within a national platform. Research on EHR sharing indicates that privacy-protection measures influence willingness to participate [17], while patient-access studies repeatedly show that transparency and useful functionality shape engagement [13–19].

Biases in EHR-derived real-world evidence can arise from missingness, differential documentation, care access and coding processes [26]. Claims introduce additional incentives because information is recorded within reimbursement workflows. Each analytical variable should therefore have a documented definition, source, expected completeness, refresh interval and known limitation. Critical phenotypes should be validated against clinical evidence in representative samples. Changes in coding rules, payer policy or provider systems should trigger impact assessment before model refresh.

Population surveillance rarely requires analysts to see directly identifying information. Tokenised identifiers, role-based access, minimum-necessary datasets, protected linkage services and separated re-identification functions can reduce exposure. Record-linkage frameworks show that privacy-preserving methods can support claims-EHR integration while maintaining analytical utility [25]. Governance should also distinguish operational care use from secondary research use because purpose, legal basis and access expectations may differ.

Historical utilisation is partly a record of who could reach care, not only who needed it. A patient with fewer claims may be healthy, underinsured, geographically constrained or disengaged. Models trained mechanically on utilisation can therefore learn access patterns as if they were biological risk. Fairness analysis should examine whether predictions and resulting interventions distribute benefit appropriately across groups, and whether false negatives cluster in populations with poorer data capture [27, 28]. Where data incompleteness is structural, the correct response may be active outreach or alternative data collection rather than purely mathematical correction.

Data engineers can certify pipeline integrity, analysts can validate models, clinicians can confirm phenotype and intervention logic, payers can contribute utilisation knowledge, and population-health leaders can own outcomes. No group should be able to deploy a material model unilaterally. High-impact uses should require documented approval, intended-use boundaries, external validation where possible, monitoring plans and retirement criteria. This transforms governance from a compliance layer into the mechanism that connects data reliability, clinical legitimacy and accountable action.

10 Implementation framework for Saudi Arabia

Stage one establishes the trusted data foundation: common definitions, longitudinal identity resolution, source-provenance rules, data-quality metrics and controlled access. NPHIES transactions should first be used for descriptive cohorting and utilisation analysis before more complex prediction is introduced.

Stage two develops surveillance phenotypes for a small number of high-burden conditions such as diabetes, hypertension and cardiovascular disease. Each phenotype should be clinically reviewed and validated against EHR or laboratory evidence. Dashboards can then report disease burden, screening gaps, medication continuity and potentially avoidable acute-care patterns. Stage three introduces risk

models only after the target population, outcome and intervention are defined. Saudi diabetes prediction research provides a credible starting point, but claims-derived models remain early and require prospective and external validation before national decision use [5–8].

Alerts should route to primary-care teams, digital outreach or case management according to intervention intensity. Every alert should have a completion pathway, and non-response should trigger escalation rules appropriate to clinical urgency. Stage five creates a learning system in which outcomes, model drift, subgroup performance, intervention uptake and resource use are reviewed together. Preventive success should feed back into model recalibration so that declining risk after treatment is recognised rather than interpreted as prediction error.

Technical measures include linkage rate, data latency, missingness, calibration and drift. Clinical measures include screening completion, newly confirmed disease, treatment initiation and control of key risk factors. Population measures include emergency utilisation, avoidable admissions, regional variation and equity gaps. Operational measures include time from risk detection to contact, intervention completion and workload per prevented event. This balanced scorecard prevents the programme from optimising model accuracy while overlooking service capacity or patient outcomes.

Population-health analytics requires collaboration among clinicians, health-information specialists, actuaries, public-health practitioners, data scientists and primary-care operations. Building reusable definitions, validation protocols and intervention templates creates institutional knowledge rather than dependence on isolated models. Such capability is consistent with the broader reform objective of using digital infrastructure to improve efficiency while strengthening preventive and patient-centred care [2, 3, 12].

11 Discussion and research gaps

The evidence supports a strong case for integrating NPHIES and insurance data with clinical information, but it also exposes a gap between technical feasibility and demonstrated population benefit. Claims offer longitudinal scale; EHRs offer clinical depth; patient-facing systems offer an engagement channel; and standards-based linkage makes these sources analytically compatible. The distinctive contribution of an integrated model is the ability to observe population trajectories, prioritise risk and verify whether action occurred.

National clinical data have supported diabetes-risk models with external validation [5], and recent research has explored broader prediction of disease onset [6]. Claims-based diabetes and retinopathy systems demonstrate feasibility but are currently reported as conference abstracts, so their generalisability, calibration stability, subgroup performance and prospective impact remain uncertain [7, 8]. This is precisely why a national architecture should separate experimentation from operational deployment and require increasing levels of evidence as decision consequence rises.

Future research should evaluate whether NPHIES-linked risk stratification reduces preventable utilisation compared with standard care, not merely whether it predicts it. Pragmatic cluster trials or phased implementation studies could compare intervention strategies, while interrupted time-series designs could evaluate population outcomes where randomisation is impractical. Cost-effectiveness is also essential because high-sensitivity surveillance can generate substantial follow-up workload. Finally, Saudi research should examine fairness across regions and insurance contexts, and should quantify the information gained or lost when claims are linked to provider EHR data, following approaches used in other real-world linked cohorts [29].

High utilisation can indicate severe disease, poor access to appropriate primary care, fragmented care or simply proximity to services. Intervention design should therefore incorporate clinical review and service-pathway knowledge. Population analytics is most valuable when it helps distinguish modifiable system failures from non-modifiable risk, allowing prevention resources to be directed where action can plausibly change outcomes.

12 Conclusion

NPHIES creates an important opportunity to extend Saudi health-data exchange from transaction processing to population-health intelligence. Insurance claims can provide longitudinal visibility of service use at scale, but their preventive value increases substantially when they are linked to EHR, laboratory, pharmacy and patient-level information. The resulting data layer can support transparent disease phenotypes, dynamic risk stratification and targeted preventive pathways for chronic conditions.

The central design principle is a closed loop from signal to action. Surveillance must produce clinically credible cohorts; risk models must be calibrated, fair and monitored; interventions must have accountable owners; and outcomes must feed back into both service design and model assurance. Success should therefore be measured in earlier diagnosis, completed prevention, reduced avoidable acute care and narrower equity gaps rather than predictive accuracy alone. With disciplined interoperability, governance and prospective evaluation, NPHIES can become a foundation for a more proactive, learning and prevention-oriented Saudi health system.

References

- [1] Almazrou SH. Unveiling the National Platform for Health and Insurance Exchange Services: data exchange era within the healthcare system in Saudi Arabia. *Saudi Journal of Health Systems Research*. 2025;5(1):1–2. <https://doi.org/10.1159/000542561>
- [2] Suleiman AK, Ming LC. Transforming healthcare: Saudi Arabia's Vision 2030 healthcare model. *Journal of Pharmaceutical Policy and Practice*. 2025;18(1):2449051. <https://doi.org/10.1080/20523211.2024.2449051>
- [3] Binkheder S, Aldekhyel R, Almulhem J. Health informatics publication trends in Saudi Arabia: a bibliometric analysis over the last twenty-four years. *Journal of the Medical Library Association*. 2021;109(2):219–239. <https://doi.org/10.5195/jmla.2021.1072>
- [4] Qaffas AA, Hoque R, Almazmomi N. The Internet of Things and big data analytics for chronic disease monitoring in Saudi Arabia. *Telemedicine and e-Health*. 2021;27(1):74–81. <https://doi.org/10.1089/tmj.2019.0289>
- [5] Alkattan A, Al-Zeer A, Alsaawi F, Alfaleh A, Alabdulkareem K, Radwan N, et al. The utility of a machine learning model in identifying people at high risk of type 2 diabetes mellitus. *Expert Review of Endocrinology & Metabolism*. 2024;19(6):513–522. <https://doi.org/10.1080/17446651.2024.2400706>
- [6] Al-Hussein F, Tafakori L, Abdollahian M, Al-Shali K, Al-Hejin A. Predicting type 2 diabetes onset age using machine learning: a case study in KSA. *PLOS ONE*. 2025;20(2):e0318484. <https://doi.org/10.1371/journal.pone.0318484>
- [7] Ghosheh G, Ghosheh M, Mohamed MA, Afzal W, Kheir A. 1076-P: real-time early diabetes detection using Saudi insurance claims data – development and validation of a predictive system. *Diabetes*. 2025;74(Supplement 1):1076-P. <https://doi.org/10.2337/db25-1076-P>
- [8] Ghosheh G, Ghosheh M, Afzal W, Mohamed MA, Kheir A. 473-P: predicting diabetic retinopathy using Saudi insurance claims data – development and validation of a real-time risk stratification system. *Diabetes*. 2025;74(Supplement 1):473-P. <https://doi.org/10.2337/db25-473-P>
- [9] Alzghaibi H. Barriers to the utilization of the Sehaty mobile health application in Saudi Arabia: a cross-sectional survey of non-healthcare users. *Frontiers in Medicine*. 2025;12:1554078. <https://doi.org/10.3389/fmed.2025.1554078>

-
- [10] Alsahli S, Hor SY, Lam MK. Physicians' acceptance and adoption of mobile health applications during the COVID-19 pandemic in Saudi Arabia: extending the unified theory of acceptance and use of technology model. *Health Information Management Journal*. 2025;54(3):255–267. <https://doi.org/10.1177/18333583241300534>
- [11] Alsahli S, Hor SY, Lam M. Factors influencing the acceptance and adoption of mobile health apps by physicians during the COVID-19 pandemic: systematic review. *JMIR mHealth and uHealth*. 2023;11:e50419. <https://doi.org/10.2196/50419>
- [12] Alshaikh H, Alghamdi H, Aldosari H, Alanazi A, Aldosari B. Data analytics applications in population health in Saudi Arabia: challenges and opportunities. *Studies in Health Technology and Informatics*. 2025;328:501–504. <https://doi.org/10.3233/SHTI250770>
- [13] Simola S, Hörhammer I, Xu Y, Bärkås A, Fagerlund AJ, Hagström J, et al. Patients' experiences of a national patient portal and its usability: cross-sectional survey study. *Journal of Medical Internet Research*. 2023;25:e45974. <https://doi.org/10.2196/45974>
- [14] Abd-Alrazaq A, Alalwan AA, McMillan B, Bewick BM, Househ M, Al-Zyadat AT. Patients' adoption of electronic personal health records in England: secondary data analysis. *Journal of Medical Internet Research*. 2020;22(10):e17499. <https://doi.org/10.2196/17499>
- [15] van Mens HJT, Duijm RD, Nienhuis R, de Keizer NF, Cornet R. Towards an adoption framework for patient access to electronic health records: systematic literature mapping study. *JMIR Medical Informatics*. 2020;8(3):e15150. <https://doi.org/10.2196/15150>
- [16] Wang P, Li T, Yu L, Zhou L, Yan T. Towards an effective framework for integrating patient-reported outcomes in electronic health records. *Digital Health*. 2022;8:20552076221112152. <https://doi.org/10.1177/20552076221112152>
- [17] Guo S, Dang Y, She B, Li Y. Sharing intention of electronic health records in online health communities: patients' behavioral decisions in the context of privacy protection measures. *Frontiers in Psychology*. 2022;13:1047980. <https://doi.org/10.3389/fpsyg.2022.1047980>
- [18] Kujala S, Hörhammer I, Väyrynen A, Holmroos M, Nättiäho-Rönnholm M, Hägglund M, et al. Patients' experiences of Web-based access to electronic health records in Finland: cross-sectional survey. *Journal of Medical Internet Research*. 2022;24(6):e37438. <https://doi.org/10.2196/37438>
- [19] Liu J, Gong X, Weal M, Dai W, Hou S, Ma J. Attitudes and associated factors of patients' adoption of patient accessible electronic health records in China – a mixed methods study. *Digital Health*. 2023;9:20552076231174101. <https://doi.org/10.1177/20552076231174101>
- [20] Golinelli D, Pecoraro V, Tedesco D, Negro A, Berti E, Camerlingo MD, et al. Population risk stratification tools and interventions for chronic disease management in primary care: a systematic literature review. *BMC Health Services Research*. 2025;25(1):526. <https://doi.org/10.1186/s12913-025-12690-0>
- [21] Oddy C, Zhang J, Morley J, Ashrafian H. Promising algorithms to perilous applications: a systematic review of risk stratification tools for predicting healthcare utilisation. *BMJ Health & Care Informatics*. 2024;31(1):e101065. <https://doi.org/10.1136/bmjhci-2024-101065>
- [22] Geurten RJ, Struijs JN, Bilo HJG, Ruwaard D, Elissen AMJ. Disentangling population health management initiatives in diabetes care: a scoping review. *International Journal of Integrated Care*. 2024;24(1):3. <https://doi.org/10.5334/ijic.7512>
- [23] Jackson SL, Lekiachvili A, Block JP, et al. Preventive service usage and new chronic disease diagnoses: using PCORnet data to identify emerging trends, United States, 2018–2022. *Preventing Chronic Disease*. 2024;21:E49. <https://doi.org/10.5888/pcd21.230415>
- [24] Mandl KD, Gottlieb D, Mandel JC, et al. Push button population health: the SMART/HL7 FHIR Bulk Data Access Application Programming Interface. *npj Digital Medicine*. 2020;3:151. <https://doi.org/10.1038/s41746-020-00358-4>
- [25] Kiernan D, Carton T, Toh S, et al. Establishing a framework for privacy-preserving record linkage among electronic health record and administrative claims databases within PCORnet, the National Patient-Centered Clinical Research Network. *BMC Research Notes*. 2022;15:337. <https://doi.org/10.1186/s13104-022-06243-5>
- [26] Al-Sahab B, Leviton A, Loddenkemper T, Paneth N, Zhang B. Biases in electronic health records data for generating real-world evidence: an overview. *Journal of Healthcare Informatics Research*. 2024;8(1):121–139. <https://doi.org/10.1007/s41666-023-00153-2>
-

-
- [27] Rountree L, Lin YT, Liu C, et al. Reporting of fairness metrics in clinical risk prediction models used for precision health: scoping review. *Online Journal of Public Health Informatics*. 2025;17:e66598. <https://doi.org/10.2196/66598>
 - [28] Paulus JK, Kent DM. Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *npj Digital Medicine*. 2020;3:99. <https://doi.org/10.1038/s41746-020-0304-9>
 - [29] Geddes J, Katz J, Asimwe A, et al. Quantification of information gained by linking claims data to an electronic health record cohort of patients with metastatic breast cancer. *Pharmacoepidemiology and Drug Safety*. 2025;34(10):e70213. <https://doi.org/10.1002/pds.70213>
 - [30] Marfoglio A, Nardini F, Arcobelli VA, Moscato S, Mellone S, Carbonaro A. Towards real-world clinical data standardization: a modular FHIR-driven transformation pipeline to enhance semantic interoperability in healthcare. *Computers in Biology and Medicine*. 2025;187:109745. <https://doi.org/10.1016/j.combiomed.2025.109745>