International Journal of Medical Informatics 222 (2026) 106678

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

Review

Predictive analytics for health-system decision support using population health data: a global scoping review of implementation, governance and decision integration

William Dormechele [a], Pinar Guven-Uslu [a,*], Beatriz De La Iglesia [b]

[a] Norwich Business School, University of East Anglia, Norwich, UK
[b] School of Computing Sciences, University of East Anglia, Norwich, UK

ARTICLE INFO

Keywords:
Predictive analytics
Health-system decision support
Routine health data
Machine learning
Public health informatics
Workflow integration
Implementation science

ABSTRACT

Background: Predictive analytics is increasingly applied to routine and population health data, but its translation into operational health-system decisions remains uncertain. We examined implementation maturity, data integration, governance, workflow integration, uncertainty and documented decision pathways.

Methods: We conducted a global scoping review of peer-reviewed studies published from 2014 to 2025 across five databases. Eligible studies applied predictive or forecasting methods to routine healthcare or population-level data for health-system decision-making. Findings were synthesised descriptively and narratively and reported in accordance with PRISMA-ScR. An additional assessment of Overton, WHO IRIS and PAHO IRIS examined implementation evidence in grey literature.

Results: Of 2,623 screened records, 161 articles were included; 128 (79.5%) were from high-income settings. The highest documented stage was development in 139 articles (86.3%), validation in 10 (6.2%), pilot implementation in 3 (1.9%) and operational deployment in 9 (5.6%). Although 118 articles (73.3%) were positioned as relevant to resource allocation or capacity planning, only 9 (5.6%) documented an output-to-decision pathway and 14 (8.7%) reported routine workflow integration, revealing a marked claim-to-action reporting gap between stated relevance and documented action. Implementation barriers most often concerned data quality and interoperability (112; 69.6%) and validation and transportability (104; 64.6%). Only 24 articles (14.9%) described how uncertainty informed decisions. Supplementary grey literature identified three additional implementations, two operational and one pilot, supporting stroke-service planning, neighbourhood risk targeting and claims anomaly investigation.

Conclusions: Peer-reviewed evidence remains dominated by model development, while integration into decision pathways and routine workflows is infrequently documented. The three supplementary implementations from grey literature illustrate a practical role for predictive informatics in identifying system-level risks and directing planning, prevention or investigation. Advancing predictive analytics to operational decision support requires evaluation of the complete decision-integration chain, comprising decision actors, predictive outputs and associated uncertainty, delivery mechanisms, decision rules, actions, governance, workflow integration and lifecycle monitoring.

1. Introduction

Health systems increasingly rely on integrated data to support planning, service delivery, and accountability, particularly to monitor progress towards the Sustainable Development Goals (SDGs). These systems provide a foundation for evidence-informed decision-making, particularly in resource-constrained settings [1,2]. Such data can inform forecasting of service demand, identification of high-burden districts, targeting of outreach activities, and planning of workforce and commodity needs. Global policy agendas emphasise the use of data to guide health investment and service delivery [3].

Predictive modelling and machine learning have expanded rapidly in health research, alongside increasing computational capacity and greater availability of health data. In high-income settings, these

* Corresponding author at: Norwich Business School, University of East Anglia, Norwich, UK.
E-mail addresses: w.dormechele@uea.ac.uk (W. Dormechele), P.Guven@uea.ac.uk (P. Guven-Uslu), B.Iglesia@uea.ac.uk (B.D. La Iglesia).

https://doi.org/10.1016/j.ijmedinf.2026.106678
Received 29 May 2026; Received in revised form 14 August 2026; Accepted 19 August 2026
Available online 4 September 2026
1386-5056/

approaches are used for infectious disease forecasting, population risk stratification, hospital capacity planning, and resource optimisation [4–6]. However, these applications often depend on digital infrastructures and interoperable data systems that remain limited in many low- and middle-income countries [7]. As a result, studies from these settings are often research-driven modelling exercises with limited evidence of integration into routine health-system workflows.

Evidence from low- and middle-income settings shows that predictive approaches can be applied to vaccination coverage, service utilisation, and health-seeking behaviour [8–10]. Yet, routine health information is not consistently used to guide planning and management decisions in many health systems [11]. Previous literature highlights challenges in integrating heterogeneous data sources for decision support, including the burden of harmonisation and missing data [8,12,13].

Existing reviews have described growing interest in artificial intelligence, predictive analytics, interoperability and clinical data infrastructures, while also identifying persistent challenges related to infrastructure, governance and integration [14–20]. However, they have given less attention to whether predictive outputs are embedded in decision pathways, governed by accountable arrangements, integrated into workflows, or communicated with usable uncertainty. These features are central to medical informatics because the operational value of predictive models depends on their embedding within sociotechnical systems that connect data infrastructure, intended users, organisational workflows, governance arrangements, implementation processes and accountable decision-making [21–24].

This scoping review therefore examines how predictive modelling, machine learning, and statistical forecasting have been applied to routine healthcare and population health data to support health-system decision-making. Specifically, we map data sources, integration approaches, modelling families, predicted outcomes, decision levels, implementation maturity, governance reporting, uncertainty communication and actionability. By shifting the focus from model performance alone to the conditions required for operational use, the review identifies where current predictive analytics research aligns with, and falls short of, medical informatics priorities for real-world implementation.

2. Methods

2.1. Study design and framework

This study was conducted as a scoping review following the framework of Arksey and O'Malley [25] with later refinements by Levac and colleagues [26], aligned with Joanna Briggs Institute guidance [27]. The review was reported in accordance with the PRISMA-ScR guidelines [28] (Supplementary Appendix A).

The review was guided by the question: *How have predictive modelling, machine learning and statistical forecasting methods using routine or population health data been developed, integrated, governed and operationalised for health-system decision support?* We treated predictive analytics as an informatics intervention rather than merely a modelling exercise, meaning that coding focused not only on analytical methods but also on data integration, workflow integration, governance, communication of uncertainty, and links to decision-making.

2.2. Search strategy and selection criteria

A structured search strategy identified empirical studies using predictive modelling, machine learning or statistical forecasting with routine or population health data for population-level or health-system decision support. XX and YY developed the strategy with input from a librarian and co-author ZZ.

We searched Scopus, Web of Science, PubMed, IEEE Xplore and the ACM Digital Library. The prespecified primary review was restricted to peer-reviewed journal articles to characterise the indexed peer-reviewed evidence base. An additional grey-literature assessment was undertaken to examine implementation-stage evidence reported outside journals. The five bibliographic database searches were updated on 22 December 2025. Publication filters covered 1 January 2014 to 31 December 2025, with the searchable evidence window ending on the final search date. The 2014 start date was a pragmatic boundary chosen to capture a contemporary 12-year period during which routinely collected digital health data and applied predictive analytics expanded substantially, while retaining sufficient scope for detailed coding of implementation, governance and decision integration. It was not tied to a single technological breakthrough or policy change. Search terms combined concepts relating to predictive analytics (e.g., 'predictive model', 'forecast', 'machine learning'), routine and population health data (e.g., 'health information systems', 'administrative data', surveillance, HDSS), and integration mechanisms (e.g., 'data linkage', 'interoperability', 'data harmonisation'). Full strategies are provided in Supplementary Appendix B. Search strategy development followed published methodological guidance from Khalil and Tricco [29], and Pollock and colleagues [30].

Studies were eligible if they were peer-reviewed journal articles published between 1 January 2014 and 22 December 2025, with full text available in English. Eligible studies analysed routine healthcare or population-level data and applied predictive modelling, machine-learning methods or statistical forecasting. Eligibility was determined by the intended decision actor and action rather than by the unit of prediction. Studies producing individual-level predictions were included only when the article identified, explicitly or through a clearly described application, an organisational or public-health actor and a system-level action such as service planning, population stratification, surveillance response, resource allocation or policy planning. Studies were excluded when predictions were intended solely to guide diagnosis, prognosis, triage, treatment selection or bedside management of the predicted individual. Studies using a single routine dataset remained eligible when this system-level decision link was present. We also excluded studies lacking a predictive or forecasting component, as well as reviews, protocols, editorials, commentaries, conference abstracts, theses and methodological papers without applied predictive analysis. Search outputs were exported to EndNote 2025 [31] for de-duplication and screening in Rayyan.

To examine whether implementation-stage evidence was underrepresented in the peer-reviewed evidence base, we conducted an additional grey-literature assessment (Supplementary Appendix D). Overton was searched in full text on 27 July 2026 for English-language documents published between 1 January 2014 and 30 November 2025. WHO IRIS and PAHO IRIS were assessed on 5 August 2026 using fixed targeted searches; PAHO IRIS was searched in English, Spanish and Portuguese. Searches combined terms for predictive methods, routine or population health data, implementation and system-level decision support, and the same substantive eligibility criteria were applied throughout. Supplementary findings were analysed separately from the prespecified peer-reviewed evidence base. Full search strategies, record-level screening decisions and exclusion reasons are provided in Supplementary Appendix D. XX verified screening decisions, and YY and ZZ independently reviewed the three included reports and agreed on their eligibility and implementation-stage classifications.

2.3. Study selection

Selection of the peer-reviewed evidence was conducted in two stages. Titles and abstracts were screened independently by two reviewers (XX and YY). Where screening decisions differed, disagreements were resolved through discussion, with a third reviewer (ZZ) providing a final decision when agreement could not be reached. Before screening, 30 studies were piloted to ensure consistent interpretation of eligibility criteria. Full-text screening was undertaken for all potentially eligible studies, using the same approach to resolve disagreements. Reasons for full-text exclusion are summarised in Fig. 1.

2.4. Data charting process

A structured data charting tool was created in Microsoft Excel by one reviewer (XX) and refined iteratively. Ten studies were piloted to refine extraction fields and coding rules. One reviewer (XX) extracted data from all included studies, and a second reviewer (YY) checked each record for completeness and consistency. Disagreements were resolved through discussion with a third reviewer (ZZ).

3. Data items

The extraction captured bibliographic details, population and setting, data sources, integration methods, modelling approaches, predicted outcomes and stated decision-support purpose. Additional fields captured implementation maturity, governance, interoperability, linkage quality, uncertainty, actionability and workflow integration. Implementation maturity was assigned to the highest documented stage: development only, defined as model construction or analytical evaluation without qualifying independent validation or live use; validation, defined as temporal, prospective, external or otherwise independent performance assessment without live decision use; pilot implementation, defined as limited or time-bounded real-world use; or operational deployment, defined as recurring use within a live workflow, official programme or organisational decision process.

Author-stated intended application was coded separately from a documented decision pathway. Intended application described the use claimed by the authors. In contrast, a documented pathway required an identifiable predictive output, decision actor, specific action and route by which the output reached or informed that action.

For this review, we use the term “claim-to-action reporting gap” to describe the discrepancy between author-stated decision relevance and documented evidence that a predictive output reached an identifiable decision actor and informed a specific action through a described delivery route. The term refers to a gap in documented decision integration and should not be interpreted as evidence that decision use was absent in practice.

Pathways were classified as explicit when all four elements were reported and implied when the actor and action were identifiable but one linkage step was incomplete. Workflow integration required evidence that outputs were delivered through a routine tool or process; a statement that a model could support planning was insufficient. Unit of prediction, analytical context and level of decision supported were coded as separate variables. Prediction horizon was assigned using the longest explicitly reported forward horizon: same-day or real-time; short-term, defined as more than same-day but less than 12 months; seasonal; annual or longer, defined as 12 months or more; or not stated or not applicable. Linkage-quality reporting was assessed only where record-level, spatial or contextual linkage was applicable; general model-performance measures were not counted as linkage-quality metrics unless they explicitly evaluated the linkage process. For reporting-sensitive domains, we distinguished documented present, explicitly absent and not reported. When evidence did not satisfy a higher category, the lower category was assigned. Thus, an article describing a model as “for service planning” without identifying an actor, action or delivery route was coded as an intended application, but not as a documented decision pathway or workflow integration. Full definitions, decision rules and worked examples are provided in Supplementary Appendix C.

We conducted a focused extraction of barriers and enablers to implementation. Only factors explicitly reported by study authors were coded. Codes were consolidated into seven domains: data quality and interoperability; validation and transportability; workflow and usability; governance, accountability and regulation; workforce and stakeholder capacity; financing, procurement and organisational readiness; and monitoring, maintenance and model updating. Articles could contribute to multiple domains; domain totals could therefore exceed 161. Failure to report a barrier or enabler was coded as not reported.

3.1. Synthesis of results

Evidence was synthesised using descriptive and interpretive approaches. Studies were grouped by the review’s core analytical dimensions, including decision-support level (for example, facility, district, and national), data sources and integration patterns, modelling family and validation strategy, and indicators of implementation maturity. We also synthesised reporting on governance and interoperability considerations, linkage quality, uncertainty, and actionability, because these factors determine whether predictive outputs can be operationalised in routine decision-making.

R version 4.5.2 [32] and Python version 3.13.3 [33] were used to generate counts and percentages describing the distribution of included studies across these dimensions. Comparative summaries were produced using 2025 World Bank income classifications [34]. Findings were stratified by decision-support level and by country income group to examine whether implementation maturity, governance reporting, and uncertainty communication differed across settings. All figures and tables were generated directly from the extracted dataset to ensure reproducibility.

We then applied narrative synthesis to interpret and contextualise the quantitative patterns. Interpretation focused on how datasets were integrated, how predictive outputs were positioned within workflows and which barriers and enablers to implementation were explicitly reported. Barrier and enabler domains were summarised by implementation stage and synthesised narratively. This synthesis was guided by the review’s conceptual framing of predictive analytics as an end-to-end decision-support system rather than a standalone modelling exercise. Consistent with scoping review methodology, the review mapped existing literature rather than evaluating effectiveness. Therefore, a formal critical appraisal or risk-of-bias assessment of included studies was not undertaken.

3.2. Protocol

A protocol was developed before study initiation but was not prospectively registered or publicly posted. It is available from the corresponding author on reasonable request. The review was conducted and reported in accordance with the PRISMA-ScR reporting guideline [28].

3.3. Ethics considerations

This review analysed publicly available, peer-reviewed literature and did not involve primary data collection from human participants; therefore, ethical approval was not required. The review followed established best practices in research reporting and transparency.

3.4. Patient and public involvement

Patients and members of the public were not involved in the design, conduct or reporting of this scoping review.

4. Results

4.1. Study selection

Database searching identified 2,894 records across five sources. After removing 271 duplicates, 2,623 records were screened, and 2,209 were excluded. Full texts were sought for 414 reports, of which 16 could not be retrieved. Of 398 full-text reports assessed, 237 were excluded, leaving 161 studies for inclusion (Fig. 1). A full list of included studies with citation details is provided in Supplementary Table 1 to support

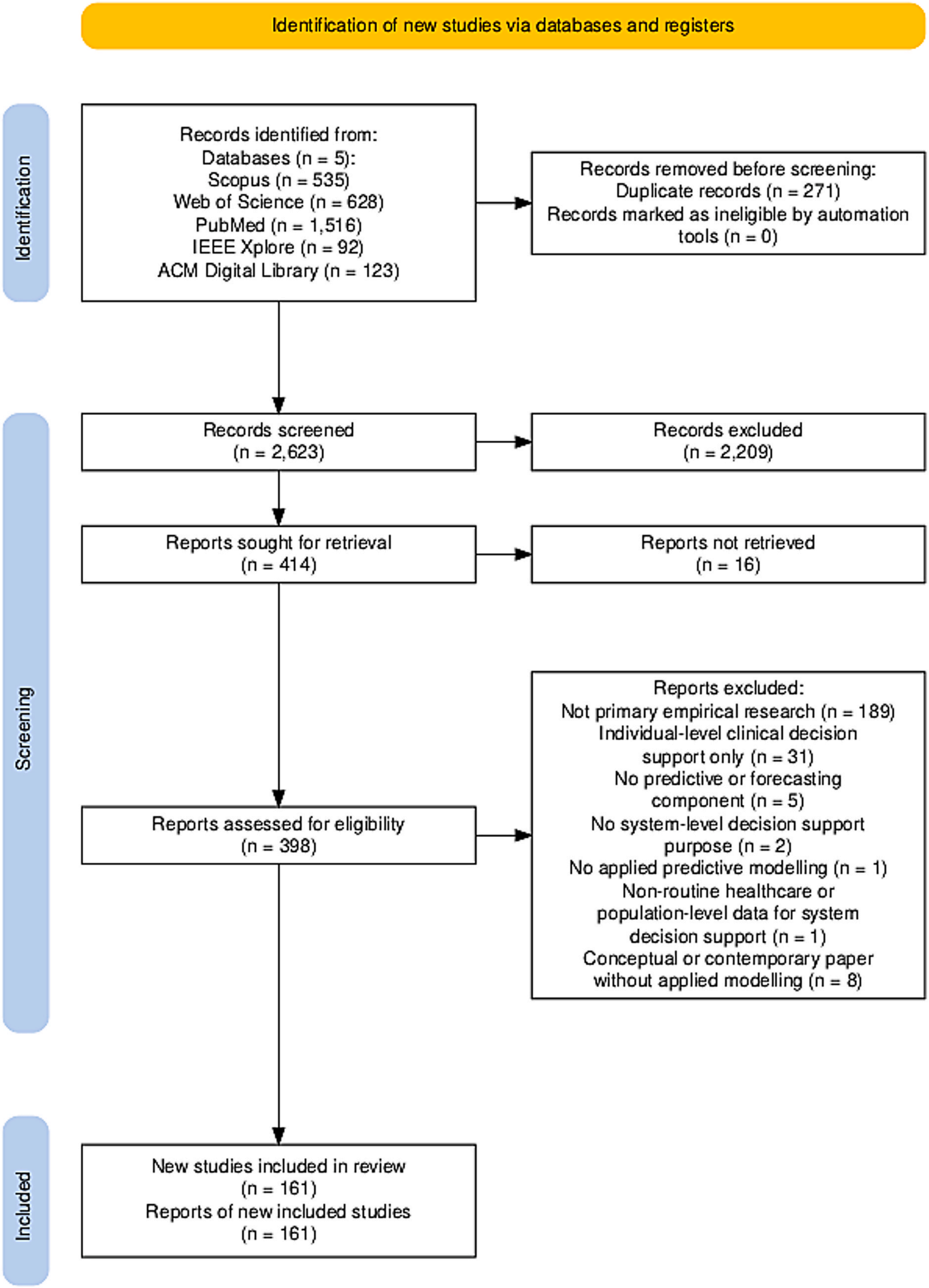

Fig. 1. PRISMA-ScR flow diagram of study selection.

transparency and reproducibility.

4.2. Temporal and geographic distribution of studies

Three features defined the evidence base: rapid growth after 2020, marked concentration in high-income settings and dominance of studies from the United States. Publication volume increased over time (Table 1; Supplementary Fig. 1). Only 16 (9.9%) studies were published between 2014 and 2017, and 32 (19.9%) between 2018 and 2020. Output increased from 2021 onwards, with 45 (28.0%) studies published between 2021 and 2023 and the highest output observed in 2024–2025 (68; 42.2%). Overall, 113 (70.2%) studies were published from 2021 onwards, indicating rapid growth in recent years.

Studies were conducted across all WHO regions but were unevenly distributed geographically (Table 1). Most studies were undertaken in the Americas (82; 50.9%), followed by the European Region (30; 18.6%) and the Western Pacific Region (28; 17.4%). Fewer studies were conducted in the Eastern Mediterranean Region (10; 6.2%), while the African Region (3; 1.9%) and the South-East Asia Region (2; 1.2%) together accounted for less than 4% of included studies. Six (3.7%) studies were global or multi-country in scope.

At the country level, the United States accounted for nearly half of all studies (74; 46.0%), followed by China (13; 8.1%) and the United Kingdom (11; 6.8%), with most other countries contributing fewer than five studies each (Supplementary Fig. 2). The global distribution of studies is shown in Supplementary Fig. 3. These figures describe the geographic distribution of eligible publications and should not be interpreted as a direct measure of where operational predictive systems exist.

Table 1
Characteristics of included studies and study contexts (N = 161).

Characteristic	Category	n (%)
Publication year		
	2014–2017	16 (9.9)
	2018–2020	32 (19.9)
	2021–2023	45 (28.0)
	2024–2025	68 (42.2)
WHO region †		
	African Region	3 (1.9)
	Americas	82 (50.9)
	Eastern Mediterranean Region	10 (6.2)
	European Region	30 (18.6)
	South-East Asia Region	2 (1.2)
	Western Pacific Region	28 (17.4)
	Global or multi-country	6 (3.7)
Health domain		
	Communicable diseases	50 (31.1)
	Non-communicable diseases	37 (23.0)
	Maternal, neonatal, and child health	3 (1.9)
	Health systems and multidomain analytics	71 (44.1)
Analytical context of the study		
	National	61 (37.9)
	Regional or district	12 (7.5)
	Facility or service level	36 (22.4)
	Facility network *	8 (5.0)
	Population or community level	44 (27.3)
World Bank income setting		
	High-income	128 (79.5)
	Upper-middle-income	20 (12.4)
	Lower-middle-income	5 (3.1)
	Low-income	1 (0.6)
	Mixed or multi-country	7 (4.3)

† Studies classified as Global or multi-country or Mixed (multi-country) spanned more than one country or World Bank income category. Each study was assigned to a single contextual category based on the highest level of aggregation reported. * A coordinated group of health facilities that share data or operational decision processes.

4.3. Data sources and linkage approaches

The data architecture was mixed rather than uniformly integrated: routine health-system sources predominated, 40.4% of articles used a single source, and reporting of how multi-source data were linked was often incomplete. Electronic health records or clinical systems were the most common primary source (66; 41.0%), followed by public health surveillance systems (42; 26.1%) and administrative or claims data (39; 24.2%). Nearly two-fifths of articles used a single dataset (65; 40.4%), whereas 45 (28.0%) integrated two sources and 51 (31.7%) integrated three or more sources (Supplementary Table 2).

When stratified by World Bank income group, integration depth did not show a consistent gradient across settings (Supplementary Fig. 4). In high-income countries, studies were distributed relatively evenly across single-source (49/128; 38.3%), two-source (37/128; 28.9%), and three-or-more-source (42/128; 32.8%) designs. Upper-middle-income settings showed a similar distribution. Lower-middle-income settings more frequently relied on single-source analyses (3/5; 60%), although denominators were small. The single low-income study integrated two data sources. Mixed multi-country studies were predominantly single source (6/7; 85.7%).

Among studies reporting integration characteristics, record linkage was the dominant approach (64; 39.8%). Feature-level fusion (3; 1.9%) and model-level integration (1; 0.6%) were rarely reported. Integration patterns were frequently under-reported, with 28 (17.4%) studies not describing the integration approach (Supplementary Table 2).

Linkage was most often performed at the individual level (86; 53.4%), with fewer articles linking at the district or regional level (27; 16.8%). Reporting of linkage identifiers and methods was often incomplete. Direct identifiers were reported in 21 (13.0%) studies and pseudonymous or encrypted identifiers in 10 (6.2%), whereas 57 (35.4%) studies did not report the identifiers used. Deterministic linkage was reported in 43 (26.7%) studies; privacy-preserving record linkage was uncommon (4; 2.5%); and 52 (32.3%) studies did not report a linkage method (Supplementary Table 2).

Facility-level decision support was most aligned with single-source designs and record linkage. In contrast, district and national decision contexts were frequently associated with surveillance-based and multi-source architectures (Fig. 2).

4.4. Predictive and machine-learning approaches

The key contrast was not between modelling families, but between frequent author-stated decision relevance and limited documentation that outputs were connected to action. Regression-based models were the most frequently reported modelling family (88; 54.7%), followed by tree-based ensemble models (54; 33.5%). Neural or deep-learning models were reported in 32 (19.9%) studies. Time-series models (23; 14.3%) and Bayesian models (22; 13.7%) were reported less frequently. Mechanistic or hybrid models were uncommon (9; 5.6%).

Authors most often framed their applications as relevant to resource allocation and capacity planning (118/161, 73.3%) and policy or strategic planning (99/161, 61.5%), followed by surveillance and early warning (75/161, 46.6%) and care pathway or clinical operations (33/161, 20.5%). These values describe stated purposes and do not demonstrate that outputs were delivered to or used by decision-makers. The intended decision level was facility management in 86 articles (53.4%), district or regional planning in 58 (36.0%) and national policy in 17 (10.6%).

Reporting relevant to model trustworthiness was limited. Interpretability was explicitly reported in 9 (5.6%) studies. External validation was reported in 35 (21.7%) studies (Table 2).

4.5. Outcomes and health system performance indicators

Individual-level prediction was common, but the unit of prediction

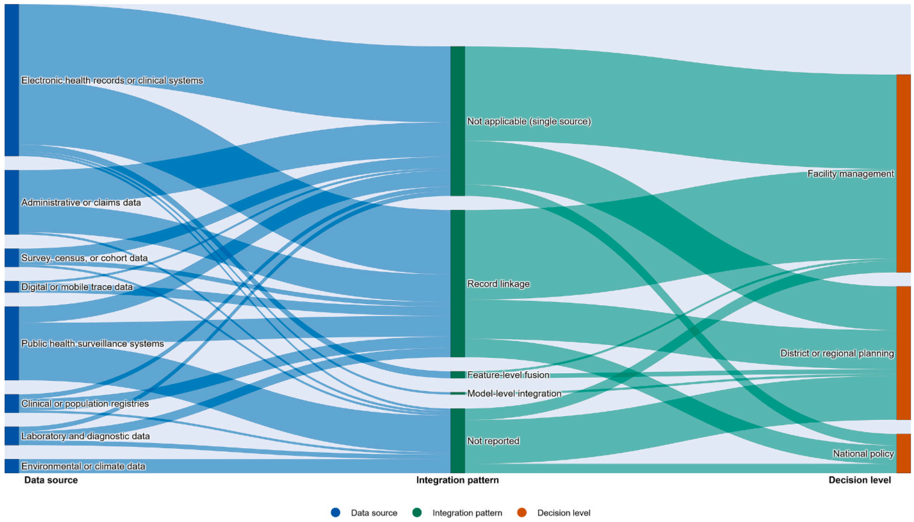

Fig. 2. Relationships between data sources, integration patterns, and levels of decision support. Sankey diagram showing relationships among primary data sources, the reported integration approach, and the decision-support level. Each link represents the number of studies reporting a given combination. Categories include both routine health system data and population-based data sources.

and the level of decision supported were not equivalent. The most frequently modelled outcome category was individual-level risk or case status (59; 36.6%), followed by incidence or case counts (40; 24.8%) and mortality (23; 14.3%) (Supplementary Table 3). Outcomes relating to hospital admissions or service utilisation were reported in 13 (8.1%) studies. Coverage, uptake, or retention outcomes were reported in 12 (7.5%) studies. Nine (5.6%) studies reported heterogeneous outcomes that could not be mapped consistently to predefined outcome families.

Prediction horizons were same-day or real-time in 18 articles (11.2%), short-term in 63 (39.1%), seasonal in 6 (3.7%) and annual or longer in 40 (24.8%); 34 articles (21.1%) did not state an extractable forward horizon or used an analysis for which a forward horizon was not applicable. Units of prediction were most frequently at the individual level (80; 49.7%) or district or regional level (63; 39.1%), with fewer articles modelling at the facility level (7; 4.3%) or national level (11; 6.8%) (Supplementary Table 3).

In contrast to the frequent statements of intended use, only 9/161 articles (5.6%) described an output-to-decision pathway. Three articles (1.9%) explicitly identified the predictive output, decision actor, action and linkage mechanism, while six (3.7%) described an identifiable but incomplete pathway. The remaining 152 articles (94.4%) did not report how a predictive output reached a specific decision. The 67.7 percentage-point difference between author-stated resource-allocation relevance (73.3%) and documented output-to-decision pathways (5.6%) illustrates the claim-to-action reporting gap defined in Section 2.4, indicating a marked discrepancy between stated decision relevance and documented decision integration (Supplementary Table 3).

Facility-level decision support most frequently focused on individual-level risk or case status (n = 34) and mortality (n = 16), whereas district or regional planning more often modelled incidence or case counts (n = 24). National policy applications were less common overall and were most frequently associated with individual-level risk or case status (n = 9) and mortality (n = 3) (Supplementary Fig. 5).

4.6. *Governance, linkage quality, and uncertainty in decision support*

Across governance, linkage quality and uncertainty, incomplete reporting limited assessment of operational trustworthiness. Ethical or regulatory approval was reported in 51 articles (31.7%), but 106 (65.8%) reported no governance mechanism beyond ethics approval. Privacy protection was not reported in 95 articles (59.0%), and bias or fairness assessment was reported in only 6 (3.7%) (Table 3; Supplementary Table 4). These figures describe article-level reporting and do not establish that the underlying practices were absent.

Linkage-quality assessment was applicable in 62 articles. Only 3/62 (4.8%) reported a linkage accuracy or completeness metric, while 59/62 (95.2%) did not; general model-performance statistics were not treated as linkage-quality measures. Missing-data handling was not reported in 100 articles (62.1%). Uncertainty quantification was reported in 90 articles (55.9%), most commonly through confidence or prediction intervals (82; 50.9%), but only 24 (14.9%) described how uncertainty informed decisions; calibration was reported in 32 (19.9%) (Table 3; Supplementary Table 5).

4.7. *Operational use and implementation characteristics*

Operational use remained uncommon in the peer-reviewed evidence base. The highest documented stage was development in 139/161 articles (86.3%); 10 (6.2%) reached validation, 3 (1.9%) pilot implementation and 9 (5.6%) operational deployment (Table 3; Supplementary Table 6).

Routine workflow integration was documented in 14 articles (8.7%), explicitly reported as absent in 37 (23.0%) and not reported in 110 (68.3%). Among the 14 integrated applications, 8 used planning or decision-support platforms, 3 alerting or triggered workflows, 2 dashboards or reporting interfaces and 1 another embedded workflow tool. Post-deployment monitoring was reported in 6 articles (3.7%) (Table 3; Supplementary Table 6).

Table 1 describes the analytical context represented in each article, whereas Fig. 3 groups articles by the organisational level of decision intended to be supported; their cross-tabulation is provided in Supplementary Table 8. Development-stage articles predominated across decision levels and most income settings, although the single low-income article was classified as operational deployment and lower-income categories had small denominators (Fig. 3; Supplementary Fig. 6). Workflow reporting by income setting is shown in Supplementary Fig. 7. These proportions describe eligible journal articles and should not be interpreted as estimates of all predictive systems operating in practice.

4.8. *Reported barriers and enablers to implementation*

At least one implementation barrier was explicitly reported in 150/161 articles (93.2%), and at least one enabler in 108 (67.1%). Overall,

Table 2
Predictive and machine-learning approaches and author-stated intended decision-support applications.

Modelling characteristic	Category	n (%)
Modelling approach family †		
	Regression-based models	88 (54.7)
	Tree-based ensemble models	54 (33.5)
	Bayesian models	22 (13.7)
	Time-series models	23 (14.3)
	Neural and deep-learning models	32 (19.9)
	Mechanistic or hybrid models	9 (5.6)
Analytical task †		
	Forecasting	47 (29.2)
	Risk prediction or stratification	77 (47.8)
	Detection or early warning	23 (14.3)
	Resource optimisation	40 (24.8)
Author-stated intended decision-support application †		
	Surveillance and early warning	75 (46.6)
	Resource allocation and capacity planning	118 (73.3)
	Care pathway or clinical operations	33 (20.5)
	Policy and strategic planning	99 (61.5)
Level of decision supported		
	Facility management	86 (53.4)
	District or regional planning	58 (36.0)
	National policy	17 (10.6)
Model interpretability reported		
	Yes	9 (5.6)
	No	152 (94.4)
External validation reported		
	Yes	35 (21.7)
	No	126 (78.3)

† Studies reported more than one modelling approach, analytical task, and author-stated intended decision-support application; therefore, category totals exceed n = 161. Intended-application categories describe how authors framed the potential use of the model. They do not establish that outputs were delivered through a routine workflow, linked to a decision rule or used to make an enacted decision.

the most frequently reported barriers concerned data quality and interoperability (112; 69.6%) and validation and transportability (104; 64.6%); the most frequently reported enablers were data and interoperability arrangements (55; 34.2%), validation and transportability measures (38; 23.6%) and workforce or stakeholder involvement (26; 16.1%). Development-stage articles most often reported data quality and interoperability barriers (99/139; 71.2%) and validation and transportability barriers (92/139; 66.2%). In contrast, operational deployments more often described workflow integration and usability (8/9; 88.9%), data and interoperability arrangements (5/9; 55.6%), organisational readiness or partnership (4/9; 44.4%) and monitoring or updating (3/9; 33.3%) as enablers. Because articles could contribute to multiple domains and only 3 pilot and 9 operational articles were identified, these comparisons are descriptive and do not establish causal determinants of deployment (Supplementary Table 9; Supplementary Fig. 9).

Table 3
Key governance, uncertainty, decision-integration and implementation indicators (N = 161).

Domain	Indicator	n (%)
Governance	Governance mechanisms not reported	106 (65.8)
Privacy	Privacy approach not reported	95 (59.0)
Bias/fairness	Bias or fairness assessment reported	6 (3.7)
Linkage quality	Linkage quality or completeness metric reported among articles with applicable linkage	3/62 (4.8)
Missing data	Missing-data handling not reported	100 (62.1)
Uncertainty	Uncertainty quantification reported	90 (55.9)
Decision integration	Uncertainty propagated to decisions	24 (14.9)
Workflow integration	Documented routine workflow integration	14 (8.7)
Implementation maturity	Development only	139 (86.3)
Operational use	Operational deployment	9 (5.6)
Monitoring	Post-deployment monitoring described	6 (3.7)

Values describe article-level reporting. “Not reported” does not establish absence. The linkage-quality denominator is 62 articles with applicable linkage; all other percentages use N = 161. Implementation stage represents the highest documented stage. Workflow integration required evidence that predictive outputs were delivered through a routine tool or organisational process.

4.9. Supplementary grey-literature implementation evidence

Three independent reports met the supplementary eligibility criteria: two from Overton and one from WHO IRIS; no PAHO IRIS report met the criteria (Supplementary Fig. 8; Supplementary Table 7). In Virginia, an operational statewide analytics portal combined hospital electronic health records, discharge and readmission data and geographic stroke indicators to support hospital quality improvement, emergency medical services planning and statewide reporting [35]. In Rhode Island, a pilot neighbourhood-level model and interactive maps identified areas vulnerable to opioid overdose and associated HIV and hepatitis C outbreaks to target prevention resources [36]. WHO’s 2024 PhilHealth report described operational use of the Machine Learning Identification, Detection and Analysis System (MIDAS), which used claims anomaly detection to identify unusual provider claiming patterns for management review and investigation [37]. Across the three reports, the common practical function was anticipatory risk identification and prioritisation: routinely generated data were converted into actionable signals that helped organisations identify where attention was needed, target resources and preventive action, or trigger further review and investigation. Although applied to different health-system problems, the cases illustrate how predictive health informatics can support preparedness and organisational response when analytical outputs are embedded in decision processes. The reports were analysed separately and were not added to the primary denominator of 161 peer-reviewed articles.

5. Discussion

5.1. Principal finding: The claim-to-action reporting gap

This review shifts attention from predictive model performance to the conditions required for operational health-system decision support. The central finding was a marked claim-to-action reporting gap. Although 118 of 161 articles (73.3%) framed predictive applications as relevant to resource allocation or capacity planning, only 9 (5.6%) documented a specific output-to-decision pathway and 14 (8.7%)

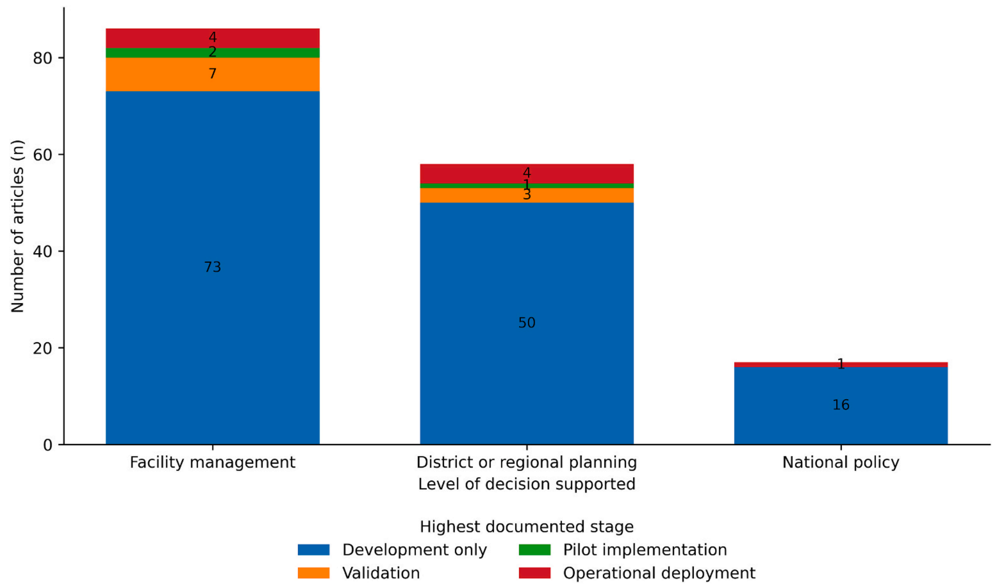

Fig. 3. Implementation maturity by level of decision supported. Bars show the number of articles at the highest documented stage: development only, validation, pilot implementation or operational deployment. Decision level refers to the organisational level at which the output was intended to inform action and is distinct from the analytical-context categories reported in Table 1. Segment labels show article counts; the cross-tabulation of analytical context and decision level is provided in Supplementary Table 8.

reported routine workflow integration. As defined in Section 2.4, this gap represents the discrepancy between author-stated decision relevance and documented decision integration; it should therefore be interpreted as a reporting gap rather than evidence that decision use was absent in practice.

This discrepancy matters because predictive performance alone does not establish operational value. For a model to function as health-system decision support, its output must reach an identifiable decision actor at an appropriate point in the workflow and be linked to an actionable choice. When these connections are not documented, it is difficult to determine whether predictive information influenced practice, how uncertainty was interpreted, who was accountable for the resulting decision, or whether the implementation could be reproduced, evaluated or transferred to another setting. The reporting gap therefore limits both assessment of real-world impact and cumulative learning about the conditions required for successful deployment.

This pattern is consistent with sociotechnical accounts of health-information technology, which emphasise alignment among technical systems, users, organisational routines and governance [21–23]. The review adds empirical evidence that stated planning relevance was common, whereas the mechanisms connecting predictive outputs to organisational decisions were rarely documented. Decision pathways should therefore be evaluated and reported as an integral component of predictive decision support rather than inferred from model performance or stated purpose alone.

5.2. Progression from development to operational use

Peer-reviewed evidence remained dominated by development-stage articles (86.3%), with 10 (6.2%) reaching validation, 3 (1.9%) pilot implementation and 9 (5.6%) operational deployment. Development-stage articles predominated across most income settings; however, the single low-income article was classified as operational deployment, and small denominators limit interpretation of the lower-income categories.

The focused synthesis indicates that progression depends on more than predictive performance. Data quality, interoperability, validation, and transportability were the most frequently reported barrier domains, whereas operational deployments more often described workflow delivery, organisational readiness and lifecycle monitoring as enablers. Because these factors were author-reported and unevenly documented, they should be interpreted as implementation conditions rather than proven causes of deployment or non-deployment.

Implementation maturity differed descriptively by decision level. Facility-level applications may have a shorter route between output and operational actor, whereas district and national decisions involve more organisations, mandates and accountability relationships. This interpretation is consistent with implementation literature but remains an explanatory hypothesis [6,18].

The three supplementary implementations [35–37], although drawn from different domains, shared a practical function: converting routinely generated data into signals that helped organisations identify elevated or unusual risk and prioritise action. Virginia used statewide stroke analytics to support quality improvement and emergency medical services planning; Rhode Island used neighbourhood vulnerability maps to target overdose and infectious-disease prevention; and PhilHealth used claims anomaly detection to focus management review and investigation. Their value therefore extends beyond showing that implementation evidence exists outside journals. They illustrate how predictive informatics can support anticipatory risk management, targeted resource use and organisational preparedness when outputs are embedded in operational workflows. The rarity of such examples is also consistent with the barriers identified in the peer-reviewed literature: operationalisation requires not only model performance, but interoperable data, validation, workflow integration, organisational readiness and ongoing monitoring. These three cases do not establish effectiveness or resilience, but they demonstrate the practical scope for informatics to strengthen system responsiveness.

5.3. Data integration and interoperability

Data integration remained a persistent constraint. Many articles relied on fragmented datasets and incompletely described linkage between routine health information, population and administrative sources. Integration depth did not increase consistently with country income, indicating that fragmentation was not confined to resource-constrained settings (Supplementary Fig. 4) [17,18,20].

5.4. Governance, workflow integration and uncertainty

Governance, workflow integration and uncertainty were often reported insufficiently to assess operational trustworthiness. In 110 articles, workflow integration was not reported, and only 24 described how uncertainty informed decisions. Lack of reporting does not establish that these activities were absent; however, it prevents assessment of decision authority, data stewardship, model ownership, maintenance, human review and accountability. Inadequate reporting is therefore itself an informatics concern [14,19].

5.5. Geographic concentration and equity

The strong concentration in high-income settings, particularly the United States, likely reflects several overlapping factors: greater availability of linked routine data and computational infrastructure; stronger health-system research and publication capacity; the English-language eligibility criterion; and bibliographic-database coverage that favours indexed journal literature. Relevant work in under-represented settings may also be published locally, in other languages or through institutional and programme reports. The observed distribution therefore reflects both research activity and the visibility of that activity within the review's information sources.

The concentration of evidence in data-rich settings also raises equity concerns. Predictive systems may reinforce existing imbalances in decision-making power when technical actors control model design and interpretation without transparent governance or inclusive decision processes [17,20] Studies should therefore report whose priorities are represented, who can challenge or override outputs and how infrastructure, data quality and governance conditions affect transferability to under-represented settings.

5.6. Implications for research, policy and practice

These findings support a minimum decision-integration reporting chain. Studies should identify: the target decision and accountable actor; the data-refresh and model-update process; the predictive output and associated uncertainty; the delivery mechanism and timing; the action threshold and available alternatives; the role of human review and escalation; governance, equity and accountability safeguards; and monitoring of use, model drift, decisions and downstream consequences. Implementation maturity should be reported as an outcome, with development, validation, pilot implementation, operational deployment and post-deployment monitoring clearly distinguished. Policy-facing evaluations should state who receives the output, which decision it changes and how use and consequences are monitored after deployment.

5.7. Strengths and limitations

This review provides a global synthesis of predictive analytics for health-system decision support across five bibliographic databases and applies an explicitly operationalised framework that distinguishes author-stated intended applications from documented decision pathways, workflow integration and implementation maturity. The review also separates documented absence from non-reporting, applies a four-stage implementation classification, and provides record-level coding rules and audit materials to support transparency and reproducibility. An additional grey-literature assessment of Overton, WHO IRIS and PAHO IRIS broadened coverage to implementation evidence reported through policy and organisational sources, while preserving the pre-specified peer-reviewed evidence base and denominator [38].

Several limitations should nevertheless be considered when interpreting the findings. The review mapped the breadth and characteristics of the literature rather than evaluating causal effectiveness, and no formal risk-of-bias assessment was undertaken, consistent with its scoping purpose. Coding necessarily depended on information reported in articles and supplementary materials; consequently, “not reported” should not be interpreted as evidence that a feature was absent in practice. Similarly, the barrier and enabler synthesis captures author-reported implementation conditions and cannot establish causal determinants of progression to deployment.

The supplementary grey-literature assessment was focused on implementation-stage evidence. Overton provided a complete exportable search denominator, whereas WHO IRIS and PAHO IRIS were assessed using a fixed targeted search approach because their repository interfaces did not provide a stable exportable result set comparable with Overton. These sources substantially broadened coverage of policy and organisational evidence and identified additional pilot and operational implementations not represented in the primary peer-reviewed set. Nevertheless, relevant implementation evidence may also be reported through national government, health-service, insurance or institutional sources outside the repositories examined. A prospectively designed multi-source grey-literature review could further characterise implementation maturity and potentially yield a different maturity profile by identifying additional pilot or operational systems reported outside peer-reviewed journals. The primary review was restricted to English-language peer-reviewed articles and used a pragmatic 2014 start date, which may have excluded relevant earlier or non-English evidence.

6. Conclusion

Among the peer-reviewed evidence examined, predictive analytics using routine and population health data expanded substantially but remained dominated by development-stage studies, with a marked gap between stated health-system relevance and documented integration into decision pathways and routine workflows. The three supplementary implementations demonstrate how routine data can be converted into actionable risk signals for service planning, targeted prevention and investigation, illustrating the practical value of predictive informatics for preparedness, prioritisation and organisational responsiveness. The principal informatics challenge is therefore not model development alone, but demonstrating how predictive outputs reach accountable decision-makers, inform specific actions, and are governed, integrated, monitored and maintained in practice. Future evaluations should report the complete decision-integration chain, including decision actors, predictive outputs and uncertainty, delivery mechanisms, decision rules, actions, governance, workflow integration and lifecycle monitoring. This is particularly important in under-represented settings, where infrastructure, data quality, organisational capacity and governance may determine transferability, equity and sustainability.

7. Data sharing statement

The extracted review dataset, coding framework, summary tables and reproducible analysis scripts supporting this review are provided as supplementary materials. No individual participant data were used. Further information can be obtained from the corresponding author upon reasonable request.

CRediT authorship contribution statement

William Dormechele: Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. **Pinar Guven-Uslu:** Writing – review & editing, Supervision, Project administration, Investigation, Conceptualization. **Beatriz De La Iglesia:** Writing – review & editing, Supervision, Conceptualization.

Funding

William Dormechele is supported by the South and East Network for Social Sciences (SENSS) Doctoral Training Partnership (DTP), which is funded by the Economic and Social Research Council (ESRC). The funder had no role in the study design, screening, data extraction, analysis, interpretation, manuscript writing, or the decision to submit the manuscript for publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors thank Doug Broadbent-Yale, social sciences librarian at the University of East Anglia, for reviewing the search strategy and providing guidance on the database search.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2026.106678.

References

[1] A.R. Oduro, et al., Profile of the Navrongo health and demographic surveillance system, Int. J. Epidemiol. 41 (4) (2012) 968–976.
[2] D.E. Phillips, et al., Are well functioning civil registration and vital statistics systems associated with better health outcomes? Lancet 386 (10001) (2015) 1386–1394, https://doi.org/10.1016/S0140-6736(15)60172-6.
[3] United Nations. "The Sustainable Development Goals Report 2023: Special Edition. ." https://unstats.un.org/sdgs/report/2023/ (accessed 29 September, 2025).
[4] G. Jiang, et al., Real-time risk prediction model for sepsis in patients with acute gastrointestinal bleeding: development and multicenter validation of a dynamic monitoring tool, EClinicalMedicine 90 (2025) 103574, https://doi.org/10.1016/j.eclinm.2025.103574.
[5] J. Mellor, et al., Forecasting influenza hospital admissions within English sub-regions using hierarchical generalised additive models, Commun. Med. (Lond). 3 (1) (2023) 190, https://doi.org/10.1038/s43856-023-00424-4.
[6] A. Ankolekar, et al., Using artificial intelligence and predictive modelling to enable learning healthcare systems (LHS) for pandemic preparedness, Comput. Struct. Biotechnol. J. 24 (2024) 412–419, https://doi.org/10.1016/j.csbj.2024.05.014.
[7] L.J. Frey, Data integration strategies for predictive analytics in precision medicine, Personalized Med. 15 (6) (2018) 543–551.
[8] R. Hariharan, et al., An interpretable predictive model of vaccine utilization for Tanzania, Front. Artificial Intelligence 3 (2020).
[9] A. Sepehri, S. Moshiri, W. Simpson, S. Sarma, Taking account of context: how important are household characteristics in explaining adult health-seeking behaviour? The case of Vietnam, Health Policy Plan. 23 (6) (2008) 397–407.
[10] C.E. Utazi, K. Nilsen, O. Pannell, W. Dotse-Gborgbortsi, A.J. Tatem, District-level estimation of vaccination coverage: Discrete vs continuous spatial models, Statistics in Med. 40 (2021) 2197–2211.
[11] F. Yusuph, J.E. Ntwenya, A. Kinyaga, N.S. Gibore, Routine health data use for decision making and its associated factors among primary healthcare managers in dodoma region, BMC Health Serv. Res. 24 (1) (2023) 1168, https://doi.org/10.1186/s12913-024-11658-w.
[12] O. Yaeesh, et al., Integrating machine learning models and explainable AI to predict DTaP vaccine demand in rural primary care, Neural Comput. & Applic. (2025) 1–20.
[13] N. Kuunibe, Using routine data for long term impact evaluation: Methodological reflections from a complex health system intervention in a low-income context, Cogent Public Health 10 (1) (2022), https://doi.org/10.1080/27707571.2022.2153448.
[14] T. Ciecierski-Holmes, R. Singh, M. Axt, S. Brenner, S. Barteit, Artificial intelligence for strengthening healthcare systems in low-and middle-income countries: a systematic scoping review, NPJ Digital Medicine 5 (1) (2022) 162.
[15] R.A. El Arab, et al., Artificial intelligence in hospital infection prevention: an integrative review, Front. Public Health 13 (2025), https://doi.org/10.3389/fpubh.2025.1547450.
[16] R. Haneef, et al., Innovative use of data sources: a cross-sectional study of data linkage and artificial intelligence practices across European countries, Arch. Public Health 78 (1) (2020), https://doi.org/10.1186/s13690-020-00436-9.
[17] K. Adegoke, et al., Interoperability as a catalyst for digital health and therapeutics: a scoping review of emerging technologies and standards (2015–2025), Int. J. Environ. Res. Public Health 22 (10) (2025) 1535.
[18] R. El-Yafouri, L. Klieb, "A scoping review of electronic health records interoperability levels, expectations, approaches, and problems," (in eng), Health Informatics J. 31 (4) (2025) 14604582251385986, https://doi.org/10.1177/14604582251385986.
[19] S.Y. Lyu, et al., The development and use of data warehousing in clinical settings: a scoping review, Front. Digital Health 7 (2025), https://doi.org/10.3389/fdgth.2025.1599514.
[20] M. Leventer-Roberts, R. Balicer, Data integration in health care, in: Handbook integrated care, Springer, 2017, pp. 121–129.
[21] T. Greenhalgh, et al., Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies, J. Med. Internet Res. 19 (11) (2017) e367.
[22] S. Kashyap, K.E. Morse, B. Patel, N.H. Shah, "A survey of extant organizational and computational setups for deploying predictive models in health systems," (in eng), J. Am. Med. Inform. Assoc. 28 (11) (2021) 2445–2450, https://doi.org/10.1093/jamia/ocab154.
[23] D. F. Sittig and H. Singh, "A new sociotechnical model for studying health information technology in complex adaptive healthcare systems," (in eng), *Qual Saf Health Care,* vol. 19 Suppl 3, no. Suppl 3, pp. i68–74, Oct 2010. doi: 10.1136/qshc.2010.042085.
[24] J. Wiens, et al., Do no harm: a roadmap for responsible machine learning for health care, Nat. Med. 25 (9) (2019) 1337–1340, https://doi.org/10.1038/s41591-019-0548-6.
[25] H. Arksey, L. O'Malley, Scoping studies: towards a methodological framework, Int. J. Soc. Res. Methodol. 8 (1) (2005) 19–32.
[26] D. Levac, H. Colquhoun, K.K. O'Brien, Scoping studies: advancing the methodology, Implement. Sci. 5 (1) (2010) 69.
[27] M.D.J. Peters, et al., Best practice guidance and reporting items for the development of scoping review protocols, JBI Evidence Synth. 20 (4) (2022) 953–968.
[28] A.C. Tricco, et al., PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation, Ann. Intern. Med. 169 (7) (2018) 467–473.
[29] H. Khalil, A.C. Tricco, Differentiating between mapping reviews and scoping reviews in the evidence synthesis ecosystem, J. Clin. Epidemiol. 149 (2022) 175–182.
[30] D. Pollock, et al., Undertaking a scoping review: A practical guide for nursing and midwifery students, clinicians, researchers, and academics, J. Adv. Nursing. 77 (2021) 2102–2113.
[31] *EndNote (EndNote 2025)*. (2023). Clarivate, Philadelphia, PA. [Online]. Available: https://endnote.com/.
[32] *R: A Language and Environment for Statistical Computing*. (2026). R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: https://www.R-project.org/.
[33] *Python (Version 3.13.3)*. (2026). [Online]. Available: https://www.python.org/.
[34] World Bank. *World Bank country classifications by income level for 2024–2025*. [Online]. Available: https://archive.ourworldindata.org/20251021-211801/grapher/world-bank-income-groups.html.
[35] Virginia Department of Health, "Data-Driven Action Steps and Statewide Capacity Building Pursuant to Stroke Care Quality Improvement in Virginia: 2021 Report to the Virginia General Assembly," Commonwealth of Virginia, 2022. [Online]. Available: https://www.vdh.virginia.gov/content/uploads/sites/133/2024/01/2022_RD139-Data-Driven_Action_Steps_and_Statewide_Capacity_Building_Pursuant_to_Stroke_Care_Quality_Improvement_in_Virginia.pdf.
[36] State of Rhode Island, "Response to the House of Representatives Committee on Energy and Commerce's Request for Information Regarding How Rhode Island Is Addressing the Opioid Crisis and Its Use of Federal Funding to Promote Treatment and Recovery Efforts, 2020.," House Committees, 2020. [Online]. Available: https://www.congress.gov/event/116th-congress/house-event/110367.
[37] World Health Organization, "Use of machine learning for fraud detection within the claims management process in the Philippines: insights from the national health insurance scheme PhilHealth," Geneva: World Health Organization, 2024, vol. 2026. [Online]. Available: https://www.who.int/publications/i/item/9789240101333.
[38] A. Paez, Gray literature: an important resource in systematic reviews, J. Evid. Based Med. 10 (3) (2017) 233–240.