Predicting Thirty-Day Hospital Readmissions Among Adults With Multiple Chronic Conditions: A Critical Methodological Review

Oroma Chukuigwe[1*], Florence Eribenne[2], Esther Sydney[3]

[1] Innovista Medical Center, Houston, TX, USA
oromachukuigwe@gmail.com

[2] Bay Cove Human Services, Boston, MA, USA
feribenne1980@gmail.com

[3] Independent Researcher, Ontario, Canada
esydney19@gmail.com

[*]Corresponding author: Oroma Chukuigwe, oromachukuigwe@gmail.com

Abstract

Adults with multiple chronic conditions account for a disproportionate share of thirty-day hospital readmissions, yet most readmission risk models were developed and judged in general or single-disease populations. This critical methodological review examines what is required for such a model to be trustworthy in multimorbid adults. Literature was identified purposively and by citation chaining, anchored on published systematic reviews of readmission models and on the reporting and appraisal standards for prognostic research, and was appraised against CHARMS, PROBAST, TRIPOD+AI and published guidance on evaluating prediction model performance. The review does not pool performance estimates and reports no new quantitative synthesis. Prior syntheses consistently describe modest discrimination, sparse calibration reporting, and little external validation, and direct evidence in older adults shows widely used scores discriminating poorly. The review argues that the ceiling on performance lies more in what routine data measure than in how models are estimated: functional status, frailty, falls, medication access and social circumstances are repeatedly associated with readmission but seldom captured in administrative data. It sets out an appraisal framework, a predictor-domain map, a transportability checklist and a worked illustration of why modest discrimination yields low positive predictive value. The Multimorbidity Readmission Prediction Readiness framework links infrastructure, measurement, model governance and care-coordination capacity, and generates four testable propositions. Local recalibration, calibration reporting, decision-analytic evaluation and equity auditing should precede deployment.

Keywords: multimorbidity; hospital readmission; prognostic models; clinical prediction rules; calibration; PROBAST; transportability; transitional care

DOI: https://doi.org/10.21590/ijtmh20241004416

1. Introduction

1.1 The clinical and economic burden of readmission in multimorbidity

Thirty-day readmission remains a high-volume, high-cost outcome, and adults with multiple chronic conditions carry much of that burden. According to the Healthcare Cost and Utilization Project, in 2020 there were 14.0 readmissions within thirty days per 100 adult index admissions in the United States, at an average cost of $17,700 per readmission; the rate was 17.0 per 100 among Medicare stays, reached 28.2 for chronic kidney disease and 22.3 for heart failure, and five conditions regarded as potentially preventable (heart failure, diabetes with complications, pneumonia, chronic obstructive pulmonary disease, and urinary tract infection) together accounted for 18.7% of all adult readmissions [1]. A companion statistical brief reported that the all-cause thirty-day rate remained at 13.9 per 100 index admissions across 2016 to 2020 [2]. The persistence of the problem is itself informative: if readmission were chiefly a consequence of discrete, correctable failures at discharge, the financial penalties and transitional care programs introduced over the preceding decade might have been expected to move the national rate further than they have.

Comorbidity burden is one of the most consistent correlates of that persistence. In a Swiss cohort that linked patient-level hospital administrative data to the population census, unplanned thirty-day readmission among patients with chronic health conditions followed a social gradient [3]. Among Medicare beneficiaries in the Health and Retirement Study, an ICD-coded multimorbidity-weighted index that embeds physical functioning was associated with both thirty-day readmission and post-discharge mortality [4]. Chronic conditions rarely occur in isolation; they cluster, for example along a cardiometabolic continuum in which a single metabolic marker combined with echocardiographic parameters has been proposed to improve cardiovascular risk stratification [5]. Readmission in this population is therefore better understood as an emergent property of several interacting conditions moving through a fragmented care system than as a property of a single index diagnosis, and a model that sees only the index diagnosis code sees little of the process it is asked to forecast.

1.2 Why multimorbidity changes the prediction problem

Multimorbidity describes the coexistence of two or more chronic diseases in one person. It rises steeply with age and drives admissions, specialist referral, and resource use; in the Medicare

analysis cited above, beneficiaries in the highest quartile of the multimorbidity-weighted index had substantially higher odds of thirty-day readmission than those in the lowest quartile, and the index was proposed as a way of bringing physical functioning into claims-based risk models [4]. Older adults who combine multimorbidity with disability and social complexity are those whose individual characteristics claims-based risk adjustment captures least well. Three features of such populations make prediction structurally harder than in single-condition cohorts.

First, the causal pathway to readmission runs through care coordination as much as through disease severity. Care-coordination frameworks designed to reduce readmissions among chronic disease patients identify handoff quality, medication reconciliation, and follow-up scheduling as the operative levers [6], and continuity-of-care frameworks for cardiometabolic patients reach the same conclusion from the policy side [7]. These levers act after the index discharge, when most prediction models have already produced their output.

Second, many of the determinants that matter are unmeasured in routine data. In a cohort of 1,509 older adults admitted to a geriatric ward, comprehensive geriatric assessment found that moderately impaired gait ability, together with male sex, was independently associated with thirty-day readmission, information that the scores routinely applied to such patients do not contain [8]; among patients aged 80 years and over, multimorbidity and prior falls have been reported to correlate with readmission risk in a prospective cohort [9]. Medication access is a comparable blind spot. Equitable access to chronic therapies in hospital pharmacy settings has been framed as a continuity-of-care problem [10], supply chain inefficiency translates into affordability failure in resource-constrained systems [11], last-mile distribution barriers determine whether a discharge prescription is filled [12], and unscreened drug interactions can turn the post-discharge regimen itself into a hazard, as a framework for systematic interaction screening in antenatal care illustrates in a different clinical population [13].

Third, the infrastructure needed to capture these variables is unevenly distributed. Analyses of health system transformation place a governance backbone for interoperability, data quality, and analytics at the base of any analytic capability [14], and argue that scientific innovation fails to improve outcomes without a functioning delivery infrastructure [15]. Diagnostic capacity belongs to that backbone: integrated laboratory information systems improve detection accuracy

[16], and quality systems determine whether laboratory results are trustworthy enough to be used as model inputs [17].

1.3 The state of the prediction literature and the gap in prior reviews

Two decades of model development have produced a crowded field with limited cumulative progress. A landmark systematic review identified 30 studies of 26 unique readmission models; among the 21 studies reporting a C-statistic, values ranged from 0.55 to 0.83, only six exceeded 0.70, and the large United States population-based models designed for hospital comparison generally discriminated poorly [18]. An update covering 60 studies and 73 unique models reported C-statistics spanning 0.21 to 0.88, with consistently moderate discrimination (C-statistic of 0.70 or above) in 11 of the 13 models built for medical condition-related readmissions, and the authors emphasized inconsistent performance across the field [19]. A review of 41 studies that used electronic medical record data found that such models perform better than models built on administrative data, but only modestly; it reported no statistically significant difference between the average C-statistics of regression and machine learning models, and it noted that most studies lacked socioeconomic features, failed to calibrate their models, and did not discuss clinical impact [20]. In hospitalized older adults, a group in which multimorbidity is the norm, both the LACE index and the HOSPITAL score discriminated at a C-statistic of 0.59 [8].

Three gaps follow from this record. No prior synthesis has examined model evaluation in a population defined by multimorbidity rather than by index diagnosis or by general admission; the review of thirty-day readmission models in paediatrics, for instance, addressed applicability in a quite different population [21]. The earliest general reviews preceded PROBAST and so did not apply a structured, domain-level appraisal of bias, although appraisals in adjacent prognostic fields have since shown how pervasive analysis-domain flaws are, as in models predicting impairment after critical illness [22] and early warning scores for ward deterioration [23]. Finally, no review has examined calibration, reconstructability, and transportability together, even though these, rather than discrimination, determine whether a published model can be adopted elsewhere. Reviews of machine learning early warning systems describe the same gap between the number of models developed and the number evaluated in new settings [24].

The applied analytics literature, meanwhile, has matured quickly outside clinical prediction. AI-based decision support for healthcare operations and resource planning [25], performance

intelligence for public service outcomes [26], data-driven lifecycle performance management [27], and predictive detection of financial risk and fraud in public systems [28] show a toolkit that is no longer technically scarce. The same methods have been applied to humanitarian logistics [29], predictive maintenance of commercial platforms [30], student retention and institutional risk [31], and budgeting from operational signals [32]. Capacity planning for multi-program portfolios [33], cost optimization through machine learning [34], crop resilience forecasting [35], and predictive strategy in large service organizations [36] complete a picture in which estimator availability is not the constraint. At population scale, AI-based risk stratification for emergency preparedness [37], big-data outbreak anticipation [38], and climate-integrated mortality prediction [39] apply the toolkit to public health. The gap is therefore not methodological capacity but its alignment with the clinical problem, with the data actually available, and with the care system that must act on the output.

1.4 Conceptual grounding

Two bodies of theory frame this review. The first is the chronic care model, which holds that outcomes for people with chronic illness depend on productive interactions between an informed, activated patient and a prepared, proactive practice team, supported by delivery system design, decision support, clinical information systems, and community resources [40]. Read through that model, a readmission risk score is a decision-support and clinical-information component; its value is conditional on the delivery-system and self-management components that must act on it. This framing predicts that discrimination and clinical value will diverge whenever the delivery system lacks capacity, which is a prediction that the literature reviewed here can illuminate but not yet test.

The second is the methodological framework for prognostic model research. The CHARMS checklist specifies what must be extracted to judge a prediction model study [41], PROBAST structures judgments of risk of bias and applicability across participants, predictors, outcome, and analysis [42], TRIPOD+AI sets reporting requirements for regression and machine learning models alike [43], and methodological guidance on reviewing prediction model performance specifies how C-statistics and calibration measures should be interpreted, transformed, and, where appropriate, combined [44]. Within this framework, discrimination, calibration, and clinical utility are distinct properties, and transportability to a new setting is an empirical

question rather than an assumption. A structure-process-outcome reading in the Donabedian tradition, and the behavioral model of health service use that separates predisposing, enabling, and need factors, also informed the grouping of predictor domains.

1.5 Objectives and contributions

This critical methodological review aims to set out what a reader, a health system, or a journal should require before trusting a model that predicts thirty-day all-cause unplanned readmission in adults with multiple chronic conditions. Its objectives are to summarize what published systematic reviews have established about the performance of readmission models [18-20]; to translate the appraisal and reporting standards for prognostic research into concrete questions for readmission models [42,43]; to examine, from the cited primary and review literature, which predictor domains are likely to matter most in multimorbid adults and why they are usually absent; and to situate model performance within the infrastructure and care-coordination conditions that determine whether a model can be implemented.

The review makes five contributions. (1) It provides an appraisal framework that maps each named methodological standard to the questions a reader should ask of a readmission model, so that published models can be judged consistently. (2) It sets out the predictor-domain argument, that functional, frailty, medication, and social information matter more than algorithmic sophistication, as a reasoned case built from cited evidence and stated as a hypothesis to be tested rather than as a measured effect. (3) It explains, with a clearly labeled worked illustration, why modest discrimination at realistic readmission rates yields low positive predictive value, and why calibration and decision-analytic evaluation are therefore indispensable. (4) It proposes the Multimorbidity Readmission Prediction Readiness (MRPR) framework, which treats performance as a joint property of a model and its operating environment and yields four testable propositions. (5) It sets out a prioritized research agenda that favors external validation, recalibration, and impact evaluation over further model development, consistent with the view of the chronic care model that information systems add value only when coupled to delivery [40].

2. Methods

2.1 Review design and rationale

This article is a critical methodological review rather than a systematic review with meta-analysis. The choice follows from the question. The question is not how well the average published model discriminates, which earlier systematic reviews have addressed for general and condition-specific populations [18-20], but what a model must demonstrate before it can be trusted in multimorbid adults and which features of the evidence base prevent that demonstration. A question of this kind is answered by appraisal against explicit standards and by reasoned argument from the best available evidence, not by pooling. The review therefore does not claim the exhaustive search, duplicate screening, or quantitative synthesis that the PRISMA 2020 statement describes for systematic reviews [45], and readers should not read it as one. It borrows from that tradition the obligation to state how sources were found, how they were judged, and what the review can and cannot conclude. The scope of the review is summarized in

Table 1. Scope of the review framed with PICOTS.

Element	Scope
Population	Hospitalized adults with two or more chronic conditions, or populations in which multimorbidity is the explicit modeling target
Index model	Any multivariable model estimating individual risk, whether regression, tree-based, or neural
Comparator	Other models, clinician judgment, or simple rules, where the cited literature reports them
Outcome	All-cause unplanned readmission to any acute hospital within 30 days of index discharge
Timing	Prediction at admission or at discharge, with attention to whether predictors are available at the stated time
Setting	Acute inpatient care, any country, any payer
Evidence considered	Published systematic reviews, primary cohort studies in older and multimorbid adults, reporting and appraisal standards, and impact studies from adjacent prediction tasks

Table 1 using the PICOTS elements recommended for questions about prognostic models [44], which separate the population, the index model, any comparator, the outcome, the timing of prediction, and the setting.

2.2 How literature was identified

The literature was identified purposively and by citation chaining rather than by a registered database search. The review was anchored on four groups of sources that are listed in full in the reference list. The first group comprises published systematic reviews of readmission prediction models, namely the general review that established the field [18], its update [19], the review of models built on electronic medical record data [20], and the review of paediatric thirty-day readmission models [21]; their included studies and conclusions define what the field is already known to show. The second group comprises the reporting and appraisal standards for prognostic research: CHARMS [41], PROBAST [42], TRIPOD+AI [43], and guidance on systematic review and meta-analysis of prediction model performance [44]. The third group comprises primary studies of readmission determinants in older and multimorbid adults [3,4,8,9] and national surveillance statistics [1,2]. The fourth group comprises evaluations of prediction models deployed in adjacent acute care tasks, where external validation and impact studies are more advanced [23,46,47]. From these anchors, backward citation chaining and forward reading of related work were used to locate literature on predictor domains, transportability, implementation, and the infrastructure and governance conditions discussed in Section 2.6. Selection was guided by relevance to a methodological question and by the need to represent each standard and each predictor domain, not by an attempt at completeness.

2.3 How the evidence was appraised

Each source was read for what it could support. Systematic reviews were used for what their abstracts and main text report about the number and performance of models, and their figures are attributed to them rather than restated as findings of the present review [18-20]. Primary studies were used for what their design permits: a single cohort can show that a variable is associated with readmission in that setting [8], but it cannot show that adding the variable to a model will improve performance elsewhere. The appraisal standards were then applied in a structured way. CHARMS was used to identify the information a reader needs to judge a readmission model study, including data source, participants, outcome definition, candidate predictors, sample size,

missing data, model development, performance, and model evaluation [41]. PROBAST was used to organize concerns about bias and applicability across participants, predictors, outcome, and analysis [42]. TRIPOD+AI was used to define what must be reported for a model to be reconstructed, validated, and implemented, including calibration and full model specification [43]. Guidance on reviewing prediction model performance was used to frame the interpretation of C-statistics, calibration measures, heterogeneity, and prediction intervals [44]. Table 2 translates these standards into questions a reader should ask of any readmission model intended for multimorbid adults. For studies evaluating the effect of using a model on patient outcomes, ROBINS-I for non-randomized studies [48] and RoB 2 for randomized trials [49] are the appropriate instruments, and the distinction matters because a well-calibrated model can still be evaluated in a badly biased impact study.

Table 2. Appraisal framework: methodological standards mapped to the questions a reader should ask of a thirty-day readmission model in multimorbid adults.

Standard or concept	What it addresses	Questions to ask of a readmission model
CHARMS [41]	Information needed to judge a prediction model study	Is the data source, setting, and period stated? How is multimorbidity defined? How many index admissions and readmission events contribute, and how many candidate predictors were considered?
PROBAST, participants domain [42]	Whether the study population matches the intended use	Were exclusions (deaths, transfers, planned admissions) handled so that the modeled population resembles the one in which the model will be used?
PROBAST, predictors domain [42]	Whether predictors are defined and available at the time of use	Is every predictor available at the stated prediction time? Would a model intended for use at admission depend on length of stay or discharge information?
PROBAST, outcome domain [42]	Whether readmission is ascertained completely and consistently	Can readmissions to other hospitals be observed? Are planned readmissions excluded consistently? How is death before readmission handled?
PROBAST, analysis domain [42]	Whether the analysis is likely to produce optimistic or biased estimates	Was sample size adequate for the number of candidate predictors? Were missing data imputed rather than excluded? Were continuous predictors kept continuous? Was univariable screening avoided? Was optimism corrected?

Standard or concept	What it addresses	Questions to ask of a readmission model
TRIPOD+AI [43]	Reporting sufficient for reconstruction, validation, and implementation	Is the full model specification (coefficients and intercept, or the trained model and preprocessing pipeline) available? Are intended use and threshold rationale stated? Is code available?
Calibration [43,44]	Whether predicted probabilities match observed risk	Is a calibration plot shown with calibration slope and intercept, or an observed-to-expected ratio? Is calibration assessed in the highest-risk groups where decisions concentrate?
Discrimination [44]	Whether the model ranks patients correctly	Is the C-statistic reported with a confidence interval, and is it an apparent, internally validated, or externally validated estimate?
Transportability and external validation [44,47]	Whether performance holds in a new setting	Has the model been validated in a different hospital, system, or time period? Were case mix, coding, and linkage in the validation setting described?
Heterogeneity and prediction intervals [44,50]	How much performance varies between settings	Where several validations exist, what range of performance should a new setting expect, rather than what is the average?
Clinical utility and decision curves [43,44]	Whether using the model does more good than harm at a relevant threshold	Is net benefit reported across a range of thresholds relevant to the intended action, compared with treating all patients and treating none?
Equity and subgroup performance [3,51]	Whether errors fall unevenly on disadvantaged groups	Are discrimination and calibration reported by race and ethnicity, payer, deprivation, sex, and age band?
Impact evaluation [48,49,52,53]	Whether using the model changes outcomes	Has the model been evaluated in an early-stage clinical study and then in a trial or well-designed non-randomized comparison?

2.4 Operationalization of multimorbidity

The definition of the population is a methodological choice with consequences for every downstream judgment. Four operationalizations of multimorbidity recur in the readmission literature and in the sources cited here: a count of two or more conditions from a named list; a Charlson or Elixhauser index above a stated threshold; a validated multimorbidity index such as the weighted index applied to Medicare data [4]; and national classifications of pluripathology. The heterogeneity of these definitions is a measurement problem with governance consequences. Quality monitoring frameworks in long-term care have shown how differently constructed

indicators produce different accountability conclusions from the same underlying population [54], and geriatric quality measure sets face the same construct-validity challenge when a single label covers patients with very different needs [55]. A Charlson threshold weights conditions by their association with mortality, whereas a condition count treats a controlled hypertension and a decompensated heart failure alike; the two definitions therefore select overlapping but distinct populations. For appraisal, this means that a model developed under one definition cannot be assumed to transport to a population selected under another, and the definition should be treated as part of the model's intended-use statement rather than as a detail of cohort assembly.

2.5 Outcome definition and the thirty-day window

The review focuses on all-cause unplanned readmission within thirty days of index discharge. The thirty-day window is the horizon used in payment policy and in the national statistics that define the burden [1], and mixing windows would conflate models answering different questions. The updated general review, for example, deliberately included both 28-day and 30-day models [19], which is appropriate for a general survey but means that pooled or tabulated figures from that review are not direct estimates for a thirty-day question. Three further features of the outcome matter for appraisal. Planned readmissions should be excluded consistently, because a planned return that goes well is not a failure of transitional care. Readmissions to other hospitals must be observable, which depends on record linkage rather than on the model. And death after discharge competes with readmission: post-discharge mortality is itself strongly associated with multimorbidity [4], so a model that predicts death well could appear to predict readmission well if the two were not separated, while a model evaluated with deaths censored may look better than it is in the sickest patients. The review therefore treats composite outcomes that cannot be disaggregated, and outcome ascertainment restricted to a single hospital, as threats to applicability that a reader should check before relying on a reported performance estimate [42].

2.6 Interpretive synthesis of infrastructure and governance literature

Conventional prognostic reviews treat performance as a property of the model. This review additionally treats it as a property of the model and its operating environment jointly, on the reasoning that where richer data depend on infrastructure, infrastructure becomes a determinant of performance. To examine this, a structured interpretive synthesis was conducted of literature on data infrastructure, diagnostic capacity, medication access, workforce, governance, and

security, much of it drawn from health systems research and from cross-sector analogues. Analyses of health system transformation locate analytic capability downstream of interoperability and data quality [14], work on real-world evidence frames predictive analytics as one component of health system performance [56], and studies of laboratory networks treat infrastructure as a public health intervention in its own right [57] whose returns require explicit measurement [58]. This literature is largely conceptual. It was therefore used to generate and interpret hypotheses about the conditions under which models perform or fail, never to supply effect sizes.

The interpretive synthesis followed three rules. Each source was cited only for what its title, venue, and scope indicate it examined, so that, for example, a framework for laboratory spatial planning was used to support the claim that facility design affects diagnostic throughput and not any claim about readmission. Cross-sector sources, such as work on predictive analytics in finance, logistics, education, or energy, were used as analogues that illustrate a mechanism, never as evidence that the mechanism operates in hospital care. And every interpretive claim was tied to a pattern documented in the clinical prediction literature, such as the modest gain from electronic record data [20] or the performance loss on external validation observed in adjacent tasks [47], so that the contextual synthesis could be checked against the clinical evidence rather than floating free of it. Documentation discipline for the review drew on data integrity principles from regulated life sciences operations, in which attributable and verifiable records are treated as preconditions for any downstream decision [59], and on the logic of corrective and preventive action, in which recurring errors lead to a change in procedure rather than case-by-case correction [60].

2.7 What this review does and does not estimate

The limits of the review should be stated plainly because they determine how its conclusions may be used. The review does not pool C-statistics, calibration measures, or classification metrics, and it reports no pooled estimate, confidence interval, prediction interval, heterogeneity statistic, or risk-of-bias tally of its own. Where numbers appear, they are either attributed to the published source that reported them, such as national readmission statistics [1,2] and the counts and ranges reported by earlier systematic reviews [18-21], or they are part of a worked illustration that is labeled as such and whose assumptions are stated. The review also does not

claim that the sources it cites are a complete or random sample of the literature. Its conclusions are of three kinds. The first are statements about standards, which rest on the published standards themselves [41-43]. The second are statements about what prior reviews found, which rest on those reviews. The third are reasoned arguments and hypotheses, such as the predictor-domain argument and the MRPR propositions, which are offered for testing and are labeled as such. Readers seeking an estimate of average model performance in multimorbid adults should treat that as an open question that an individual participant data meta-analysis, or a systematic review designed around the appraisal framework in Table 2, would be needed to answer [44].

3. Methodological Appraisal of the Evidence

3.1 What prior syntheses have established

Four published systematic reviews define the baseline against which models for multimorbid adults should be judged. The earliest general review, which identified 30 studies of 26 unique models, found C-statistics from 0.55 to 0.83 among the 21 studies that reported one, with only six above 0.70, and concluded that most models discriminated modestly; the population-based models designed for hospital comparison performed worst [18]. The update found a wider range, from 0.21 to 0.88 across 56 of 60 studies, with more consistent moderate discrimination among models for specific medical conditions, and stressed inconsistency across its 73 models [19]. The review of models built on electronic medical record data found that 17 of its 41 studies reported a C-statistic of 0.75 or more, that electronic record data improved on administrative data only modestly, that the difference between regression and machine learning was not statistically significant, and that relatively few models were reported to use calibration techniques [20]. The paediatric review, finally, identified 37 models from 28 studies, found comorbidity and length of stay to be the most common significant predictors, and reported median adherence to TRIPOD of 59% [21].

Three messages are consistent across these reviews and are directly relevant to multimorbid adults. Discrimination is usually modest, and higher values tend to come from narrower populations or richer data. Calibration is reported far less often than discrimination, so the correspondence between predicted and observed risk is usually unknown. And reporting is

incomplete enough that many models cannot be reconstructed or independently validated. None of these reviews defined its population by multimorbidity, so none answers directly how models perform when multimorbidity is the defining feature. The closest direct evidence comes from older inpatients, in whom two widely used scores discriminated at 0.59 [8], a value near the lower end of the ranges reported for general populations. That observation is consistent with, though it does not prove, the expectation that multimorbid populations are harder to predict than the general inpatient populations on which most models were built.

3.2 Discrimination: what the C-statistic can and cannot tell

The C-statistic is the most reported and most misread property of a readmission model. It is the probability that a randomly chosen patient who is readmitted receives a higher predicted risk than a randomly chosen patient who is not. A value of 0.50 represents ranking no better than chance, and a value of 0.70 is often described as acceptable, but the conventional thresholds have no clinical meaning in themselves [44]. Three properties of the statistic matter for appraisal. First, it is rank-based, so it is silent on whether the predicted probabilities are correct; a model can discriminate well and still assign risks that are systematically too high or too low. Second, it depends on the spread of risk in the population. The same model will appear to discriminate better in a heterogeneous cohort, such as a general admission population that includes both elective surgical patients and frail medical patients, than in a homogeneous one, such as a cohort restricted to older adults with several chronic conditions. This is one reason why a score validated in general populations can look weaker when applied in geriatric wards [8], and why a C-statistic from one population is a poor guide to performance in another. Third, the statistic reported in a development sample is optimistic by construction unless it has been corrected for overfitting through bootstrapping or cross-validation, and even a corrected internal estimate says little about performance in a different hospital [42].

For a reader appraising a readmission model, the practical questions are therefore not only how high the C-statistic is, but which kind of estimate it is (apparent, internally validated, or externally validated), in which population it was obtained, and how that population compares with the intended one. A modest C-statistic from an external validation in a similar multimorbid population is more informative than a higher value from a development sample in a general population. Comparison with other fields can also mislead: early warning scores for in-hospital

deterioration often report higher discrimination [61], but they predict an outcome hours away from physiological signals that are measured repeatedly, whereas readmission models predict an outcome weeks away from data that omit most of what happens after discharge.

3.3 Calibration: whether predicted risks mean what they say

Calibration describes the agreement between predicted probabilities and observed event rates. It is typically assessed with a calibration plot, a calibration slope, and a calibration intercept or calibration-in-the-large, and in reviews it is often summarized as an observed-to-expected ratio [44]. The calibration slope has a direct interpretation. A slope below 1 on validation indicates that predictions are too extreme, with high risks too high and low risks too low, which is the characteristic signature of overfitting; a slope above 1 indicates predictions that are too conservative. The intercept, or an observed-to-expected ratio different from 1, indicates that the model systematically over- or under-predicts overall, which commonly happens when a model is moved to a setting with a different baseline readmission rate. The Hosmer-Lemeshow test, historically the most common calibration statistic in this literature, is the least informative: it is sensitive to sample size, indicates neither the direction nor the location of miscalibration, and has been superseded in reporting guidance, which asks for calibration plots with slope and intercept [43].

Calibration matters more than discrimination for most decisions that readmission models are asked to support. Enrollment in a transitional care program, allocation of a pharmacist visit, or referral for home-based follow-up is usually made by applying a threshold to predicted risk. If the model underpredicts risk in the highest-risk patients, fewer of them cross the threshold than should, even if the ranking of patients is correct. The review of electronic record models reported that relatively few of its included models were described as using calibration techniques [20], and the earlier reviews focused on discrimination because that is what their included studies reported [18,19]. For multimorbid adults, whose baseline risk varies widely between settings, calibration should be assumed not to transport and should be checked and, where necessary, updated in every new setting before deployment [44].

3.4 Heterogeneity and prediction intervals

When several studies report the performance of the same or similar models, their estimates usually differ by more than chance would explain. Guidance on meta-analysis of prediction

model performance therefore recommends random-effects methods, analysis of the C-statistic on the logit scale, and, most importantly for a reader considering adoption, a prediction interval [44]. The distinction between a confidence interval and a prediction interval is central. A confidence interval around an average C-statistic describes how precisely that average has been estimated; with many large studies it can be very narrow. A prediction interval describes the range within which the performance in a new, comparable setting is expected to fall, and it widens with the between-study variance. The classical moment estimator of between-study variance [50] is still widely used, but guidance for prediction model reviews notes that alternative estimators and interval methods can better reflect uncertainty when heterogeneity is large and the number of studies is small [44].

The conceptual consequence for readmission models is that an average is not a forecast. Even if a well-conducted meta-analysis were to find a respectable average C-statistic for a family of models, a wide prediction interval would mean that a health system adopting one of those models could obtain performance barely distinguishable from chance or performance that supports useful risk ranking, with no way of knowing in advance which to expect. The ranges reported by the general reviews, from 0.55 to 0.83 in one [18] and from 0.21 to 0.88 in another [19], are not prediction intervals, because they mix different models, populations, and validation types, but they show why a reader should expect wide variation. The appraisal question is therefore whether the model has been validated in settings like the reader's own, and if so what range of performance those validations show, rather than what its best reported C-statistic was.

3.5 Predictor domains: what routine data make easy to measure

The predictors used in readmission models are best read as a map of what routine data systems make easy to measure rather than of what clinicians believe drives readmission. The general reviews found that most models rely on comorbidity, prior utilization, length of stay, and demographic variables available in administrative data [18,19], and the paediatric review found comorbidity and length of stay to be the most common significant risk factors [21]. The review of electronic record models found that most studies lacked socioeconomic features [20]. By contrast, the direct evidence in older and multimorbid adults points toward domains that administrative data record poorly. Impaired gait ability was independently associated with readmission in older inpatients undergoing comprehensive geriatric assessment [8]; prior falls

Table 3. Predictor domains for thirty-day readmission in multimorbid adults, described qualitatively with the cited basis for each.

Predictor domain	Typical data source	Basis in the cited literature	Principal appraisal concern
Comorbidity burden	Claims or electronic record diagnosis codes	Consistently used and associated with readmission; a weighted multimorbidity index is associated with readmission and mortality [4,19]	Coding intensity differs between systems, so the same index captures different burden in different settings
Prior utilization and length of stay	Claims or electronic record	Among the most common predictors in general and paediatric reviews [18,21]	Length of stay is not available at admission; prior utilization reflects access as well as need
Laboratory values	Electronic record	Part of the richer data that give electronic record models a modest advantage [20]	Availability depends on diagnostic infrastructure and test-ordering practice [16,17]
Functional status and gait	Geriatric assessment or structured nursing fields	Impaired gait independently associated with readmission in older inpatients [8]; function embedded in a multimorbidity index [4]	Rarely recorded in structured form; measured differently across settings
Frailty and falls	Prospective assessment or structured fields	Prior falls correlated with readmission in patients aged 80 and over [9]	Frailty instruments vary; falls history is often incomplete
Medication access and regimen complexity	Pharmacy claims, dispensing records, or interview	Continuity and affordability of chronic therapies shape post-discharge regimens [10-12]; adherence measurable from pharmacy data [62,63]	Usually reduced to a drug count; fill and adherence data require linkage
Social risk	Linked area data, census linkage, or screening	Social gradient in readmission among patients with chronic conditions [3]; socioeconomic features usually absent [20]	Area-level proxies misclassify individuals; inclusion raises questions for payment adjustment
Caregiver capacity	Interview or care-planning records	Caregiver burden linked to patient impairment after critical illness [64,65]	Almost never recorded in routine data
Care-process and follow-up variables	Scheduling, pharmacy, and community care records	Handoff, reconciliation, and follow-up are the operative levers in care-coordination frameworks [6,7]	Occur after the prediction is made, so cannot be used as predictors at discharge without redefining timing

correlated with readmission in patients aged 80 and over [9]; a multimorbidity index that embeds physical functioning was associated with readmission among Medicare beneficiaries [4]; and unplanned readmission among patients with chronic conditions showed a social gradient when hospital data were linked to census information [3].

Table 3 summarizes the main predictor domains qualitatively, with the typical data source, the cited basis for their relevance, and the principal appraisal concern. The table deliberately contains no frequencies, because the review did not count predictors across a defined set of models. Its purpose is to show the inversion that runs through the literature: the domains that direct geriatric and social evidence identifies as informative are those least available in the data most models use. In the terms of the chronic care model, the literature has modeled the variables held by clinical information systems and largely omitted those held by self-management support and community resources [40].

3.6 Data content versus algorithmic complexity

A recurring question in readmission modeling is whether better performance should be sought through more flexible estimators or through better data. The cited evidence favors data. The review of electronic record models found that the difference between regression and machine learning was small and not statistically significant, and that electronic record data improved on administrative data only modestly [20]. Comparative work in ward deterioration reached a conclusion that is instructive precisely because its inputs were unusually rich: in a multicenter comparison, machine learning methods outperformed conventional regression when both were given the same repeated vital sign and laboratory data [66], and that comparison built on an earlier multicenter risk stratification tool whose performance rested on combining vital signs, laboratory values, and demographic data [67]. In both cases the estimator worked on dense, time-varying inputs. Readmission models rarely have that advantage, because the information that separates a patient who returns from one who does not is mostly functional, social, and post-discharge, and is mostly absent from the data.

The argument of this review, stated as a hypothesis, is therefore that the ceiling on discrimination in multimorbid adults lies more in what is measured than in how it is modeled. The hypothesis is consistent with the direct geriatric evidence [8,9], with the modest data-source gain reported for electronic records [20], and with the chronic care model's emphasis on self-

management and community information [40]. It is not established by a within-cohort comparison, and it should be tested by adding functional, medication, and social predictors to a fixed modeling method in the same cohort and reporting the change in discrimination, calibration, and net benefit (Section 4.6).

3.7 Transportability and external validation

Transportability is the central unsolved problem for readmission models. A model developed in one system encodes that system's case mix, coding practices, linkage completeness, and care pathways; when any of these differ, both discrimination and calibration shift. The general reviews found external validation to be uncommon [18,19], and the pattern is not unique to readmission. External validation of a widely implemented proprietary sepsis model in an independent health system found performance substantially below that reported by its developer [47], and systematic reviews of AI for predicting in-hospital deterioration report that few models are externally validated and fewer still are tested prospectively [24,68]. Transportability also depends on the stability of the inputs. Scores built on routinely recorded vital signs, such as the National Early Warning Score [61], travel comparatively well because their inputs are measured similarly everywhere, whereas scores that draw more widely on the electronic record, such as an alert score for inpatient deterioration [69] and a later system for detecting deteriorating ward patients [70], depend more heavily on the local data model and warrant revalidation in each new setting. Most readmission models are of the second kind.

Table 4 sets out a transportability checklist that a health system can apply before adopting a published readmission model. Each item corresponds to a way in which a development setting can differ from an adoption setting, and each is grounded in an appraisal standard or in the cited evidence.

3.8 Classification performance and clinical utility: a worked illustration

Readmission models are often summarized by sensitivity and specificity at an author-selected threshold, but the quantity that matters to a program manager is the positive predictive value, the proportion of flagged patients who are actually readmitted. Positive predictive value depends heavily on the readmission rate. The following is a worked illustration, not a result of this review.

Table 4. Transportability checklist for adopting a published readmission model in a new setting.

Item	Question for the adopting setting	Why it matters
Population definition	Is multimorbidity defined the same way, and is the age and frailty profile similar?	Different definitions select different populations [4,55]
Baseline readmission rate	Is the local thirty-day rate similar to that in the development setting?	A different baseline shifts calibration-in-the-large even if ranking is preserved [44]
Outcome capture	Can readmissions to other hospitals be observed locally, as they were (or were not) in development?	Incomplete linkage biases observed rates and apparent performance [42]
Predictor definitions and timing	Is each predictor recorded in the same way and available at the same moment?	Coding and documentation practices vary; timing errors inflate performance [42]
Data source	Does the model depend on laboratory, functional, or social data that are incomplete locally?	Input quality depends on infrastructure [14,16]
Model specification	Is the full specification available so that the model can be run and recalibrated?	Without it, independent validation is impossible [43]
Local validation	Has discrimination and calibration been measured locally, ideally during silent operation?	Performance in a new system can fall well below developer reports [47]
Recalibration plan	Is there a plan to update the intercept and slope, and to monitor drift after deployment?	Calibration is setting-specific and changes with case mix [44]
Subgroup performance	Have discrimination and calibration been checked by race and ethnicity, payer, and deprivation?	Social gradients in readmission can translate into unequal errors [3]
Action capacity	Is there a defined, resourced action for patients above the threshold?	A score without a workflow changes little [46,71]

It assumes a thirty-day readmission rate of 17 per 100, the national figure for Medicare stays reported by the Healthcare Cost and Utilization Project [1], and four hypothetical operating points chosen to span the kinds of trade-off that a model of modest discrimination offers. Positive predictive value is calculated by Bayes' theorem as sensitivity multiplied by prevalence, divided by the sum of that quantity and the product of one minus specificity and one minus

prevalence. Net benefit is calculated at an illustrative risk threshold of 20%, as the proportion of true positives minus the proportion of false positives weighted by the threshold odds (0.20 divided by 0.80), following the logic of decision curve analysis [44]; the default strategy of treating everyone has a net benefit of 0.17 minus 0.83 multiplied by 0.25, which is negative (about minus 3.8 per 100), and treating no one has a net benefit of zero.

Table 5. Worked illustration (hypothetical operating points; not results of this review). Assumptions: readmission rate 17 per 100 [1]; 1,000 patients; net benefit at a 20% risk threshold, expressed as net true positives per 100 patients.

Hypothetical sensitivity	Hypothetical specificity	Patients flagged per 1,000	Readmitted among flagged	Not readmitted among flagged	Positive predictive value	Negative predictive value	Net benefit per 100
0.80	0.50	551	136	415	0.25	0.92	3.2
0.70	0.70	368	119	249	0.32	0.92	5.7
0.50	0.85	210	85	125	0.41	0.89	5.4
0.30	0.95	93	51	42	0.55	0.87	4.1

Three points follow from the illustration. First, at a readmission rate in the teens, most flagged patients are not readmitted unless specificity is high; at the second operating point, roughly two of every three flagged patients would not return within thirty days. At the all-payer adult rate of 14.0 per 100 [1], the same operating point would give a positive predictive value of about 0.28. Second, raising the positive predictive value requires sacrificing sensitivity, so that a high-specificity model misses most readmissions. Third, net benefit depends on the threshold, which in turn should reflect the cost and burden of the action triggered by a flag relative to the harm of a missed readmission. Each hypothetical operating point in the illustration beats both default strategies at a 20% threshold, but the ordering would change at other thresholds. This is why TRIPOD+AI asks authors to state intended use and threshold rationale [43], and why decision curve analysis across a range of thresholds is more informative than a single operating point. Whether a given positive predictive value is acceptable depends entirely on what the flag triggers: a low-cost, broadly beneficial action such as medication reconciliation tolerates many false positives, whereas a scarce and expensive program does not.

3.9 Risk of bias: recurring patterns and how to recognize them

PROBAST organizes risk of bias into four domains, and the recurring concerns in readmission modeling can be located within them [42]. In the participants domain, the main concern is that exclusions made for analytic convenience, such as removing patients who died, were transferred, or had missing data, produce a population that differs from the one in which the model will be used. In the predictors domain, the main concern is timing: a model described as usable at admission may rely on length of stay, discharge disposition, or other information that exists only at discharge. In the outcome domain, the main concern is ascertainment, particularly the inability of single-hospital data to observe readmissions elsewhere, a completeness failure analogous to the data-quality gaps that evaluations of routine disease surveillance systems document [72]. In the analysis domain, the familiar concerns are insufficient events relative to candidate predictors, complete-case analysis without justification, univariable screening of predictors, dichotomization of continuous variables, and failure to correct for optimism.

These analysis-domain flaws are not peculiar to readmission research. Systematic appraisal of early warning scores for adult inpatients found widespread methodological weakness and poor reporting [23], and appraisal of models predicting impairment after critical illness reached similar conclusions [22]. In each literature the dominant flaws are methodological choices that are well understood and avoidable, which suggests that the bottleneck is less statistical knowledge than the absence of enforced reporting and review standards. For machine learning models, additional questions arise that PROBAST was not originally designed to capture: whether release gating and automated evaluation were described, including awareness of bias in automated judging [73], and whether data curation decisions made during model fitting were documented [74]. A reader appraising a readmission model should check each domain directly rather than relying on the authors' own description of the model as validated.

3.10 Reporting completeness and reconstructability

Reporting completeness determines whether a model can be used by anyone other than its developers. Without the intercept and coefficients of a regression model, or the trained object and preprocessing pipeline of a machine learning model, a second team cannot externally validate the model, cannot recalibrate it, and cannot compare it head to head with an alternative in its own population. TRIPOD+AI requires that the full model be reported, together with its

intended use, the handling of missing data, calibration, and, for machine learning models, hyperparameters and code availability [43]. The paediatric readmission review found median adherence to TRIPOD of 59%, ranging from 33% to 81%, and called for more complete reporting of the items relevant to implementation [21], and the review of electronic record models noted that most studies failed to calibrate their models or discuss clinical impact [20].

Reporting deficits matter especially for flexible models, whose performance is more sensitive to preprocessing, class-imbalance handling, and hyperparameter choice than that of a prespecified regression. A model whose specification is not published is closed to the independent evaluation that PROBAST and TRIPOD+AI assume will follow publication. A literature in which many models cannot be reproduced accumulates claims rather than evidence, and each new development study adds a model without adding knowledge the next study can build on. For appraisal, the implication is simple: a model that cannot be reconstructed from its publication or accompanying materials should not be considered for adoption until its developers make the specification available and it has been validated locally [44].

3.11 Subgroup performance and equity

Equity is a property of how a model distributes its errors, not only of its average accuracy. National statistics show that readmission rates differ by payer and by race and ethnicity; the Healthcare Cost and Utilization Project reported the highest rates among non-Hispanic Black patients and the lowest among non-Hispanic Asian and Pacific Islander patients (16.0 versus 11.7 per 100 index admissions) over 2016 to 2020, and particularly high rates among Medicare beneficiaries aged 21 to 64 years and Medicaid patients aged 45 to 64 years [2]. Among patients with chronic conditions, readmission follows a social gradient [3]. Two kinds of error can arise when a model is applied across such groups. Lower discrimination within a subgroup means that the model ranks patients in that group less well, which reduces the efficiency of targeting. Underprediction means that the group as a whole is assigned risks below its true risk, which reduces its share of referrals at any fixed threshold. The second error is the more harmful for equity, and it is also the one that group-specific recalibration can correct.

The mechanism by which such errors arise is not mysterious: the predictors most associated with social disadvantage, such as function, social support, and medication access, are those least often measured, so their influence is absorbed into residual error that may concentrate in the

groups where they matter most. Outside health care, an analysis of deep learning load forecasting has argued that a forecast can be accurate on average and still distribute its errors inequitably across districts [75], an analogue of the pattern that readmission models may reproduce. Comparative governance literature has examined executive oversight, regulatory structures, and accountability for AI-driven healthcare management across jurisdictions [51], and systematic work on digital equity and access frameworks in health education addresses the same concern from another direction [76]. For appraisal, the question is whether subgroup discrimination and calibration were reported at all, and whether any subgroup analysis was prespecified.

3.12 Optimism, small studies, and selective reporting

Small development samples overfit, overfitting inflates apparent discrimination, and small studies with inflated performance may be more likely to be published. In prediction model research, small-study effects and optimism are therefore partly the same phenomenon, and funnel-plot asymmetry in a meta-analysis of C-statistics has a specific interpretation that differs from its interpretation in intervention reviews [44]. For a reader appraising an individual model, the relevant questions are whether the sample size was planned for the number of candidate predictors rather than assumed adequate because the cohort is large, whether optimism was corrected by resampling, and whether the performance reported is the performance of the model as it would be used, rather than of the best of several models tried. Selective reporting of the best-performing specification, or of a favorable subgroup, produces the same inflation as overfitting and is invisible without a prespecified analysis plan [42].

These concerns suggest that the ranges reported in the general reviews [18,19] should, if anything, be read as optimistic for the performance a health system is likely to achieve with a model developed elsewhere. They also explain why guidance on reviewing prediction model performance recommends separate treatment of development, internal validation, and external validation estimates rather than combining them [44]. A review or a health system that pools apparent and externally validated performance without distinction will overstate what can be expected in practice.

4. The Multimorbidity Readmission Prediction Readiness Framework

4.1 Rationale for a layered framework

The appraisal supports a claim that no single statistic captures: predictive performance in multimorbid populations is bounded from below by the data environment and from above by the care system's capacity to act. A model is one component in a chain, and the weakest link in the chain determines what the model can contribute to outcomes. Existing frameworks represent the links separately. Prognostic model reporting standards describe the model in isolation [43]; health system frameworks describe the delivery environment, including its data backbone, without reference to prediction [14]; and the chronic care model places decision support and clinical information systems alongside delivery system design without specifying how the quality of the former depends on the latter [40]. The Multimorbidity Readmission Prediction Readiness (MRPR) framework joins them. It specifies four nested layers and the dependencies between them (Figure 1), and it is intended both as an interpretive device for the evidence reviewed here and as a source of propositions that future studies can test.

4.2 Layer 1: Infrastructure foundation

The base layer comprises diagnostic and laboratory capacity, the interoperability and data quality backbone, cross-site record linkage and identity, and security, privacy, and availability. Diagnostic capacity determines which laboratory values exist to be modeled, and it is itself a product of sustainable laboratory infrastructure models [77], infrastructure-driven expansion of diagnostic access into underserved regions [78], laboratory spatial planning that affects accuracy and throughput [79], and regulatory-compliant design of molecular and pathology environments [80]. Lifecycle evaluation links that design to long-term service delivery [81], and quality systems determine whether the resulting values are reliable [17]. Interoperability governs whether values held in separate systems can be combined [14]. Record linkage determines whether the outcome can be observed at all, since a model developed in data that cannot see readmissions to other hospitals is trained on an incomplete outcome; blockchain-based cross-border data exchange [82] and privacy-focused digital identity verification [83] illustrate architectural routes past that constraint. Security and availability determine whether the system functions under stress, and systematic evidence associates cybersecurity investment with healthcare financial performance [84]. The clinical signature of this layer is the modest but

consistent advantage of models built on richer electronic record data over those built on administrative data [20].

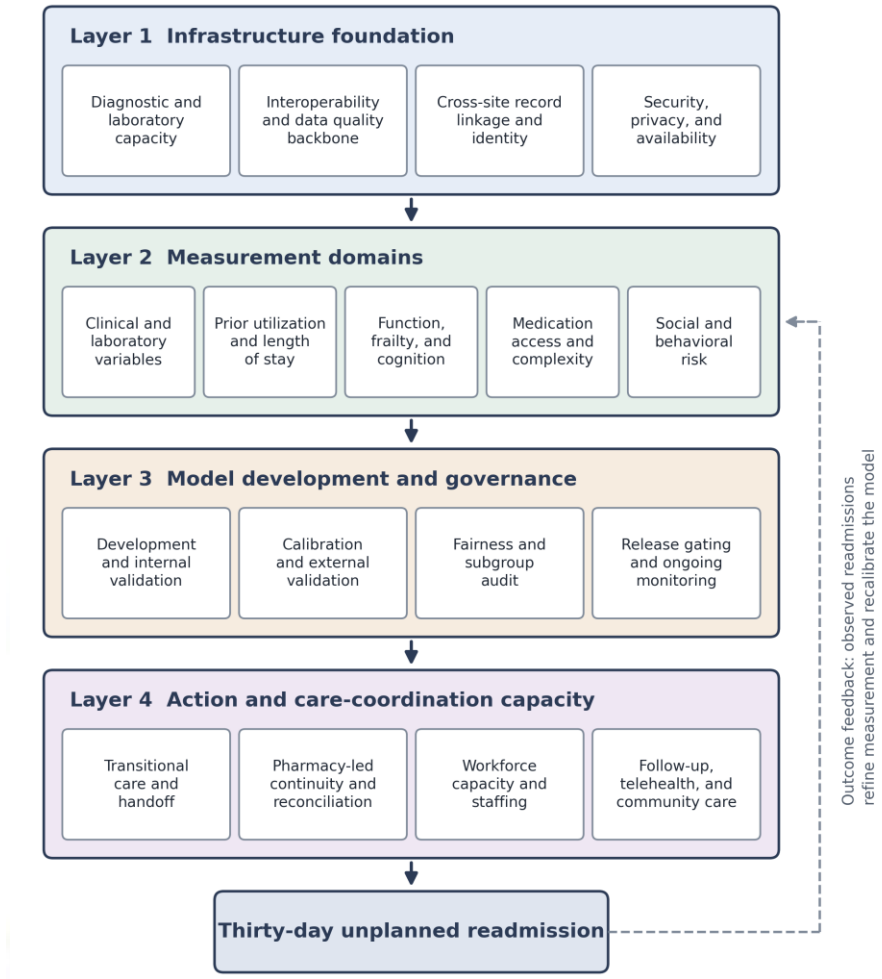

Figure 1. The Multimorbidity Readmission Prediction Readiness (MRPR) framework. Each layer bounds the one below it: predictive performance cannot exceed what the measurement layer captures, measurement cannot exceed what the infrastructure layer supports, and a prediction changes outcomes only through the action layer. The dashed path shows outcome feedback used to refine measurement and recalibrate the model.

4.3 Layer 2: Measurement domains

The second layer specifies what is measured: clinical and laboratory variables; prior utilization and length of stay; function, frailty, and cognition; medication access and complexity; and social

and behavioral risk. The appraisal in Sections 3.5 and 3.6 suggests that the largest recoverable gain sits in this layer, because the domains most clearly associated with readmission in older and multimorbid adults, such as gait, falls, function, and social circumstances [3,4,8,9], are the least available in routine data. The framework treats medication access as a first-order domain rather than a covariate, because affordability, last-mile distribution, and continuity of supply each determine whether a discharge regimen is actually taken [10-12], and because shortages are themselves predictable from supply signals, as drug shortage early-warning models show [85]. Program evidence from medication therapy management in chronically ill Medicare populations [86], evaluation of the eligibility criteria used to target such services [87], and systematic evidence on medication synchronization and utilization [88] all indicate that medication-related information is collectible at scale when a program requires it. The design of a cluster-randomized tele-pharmacy trial for chronic disease adherence shows how such measures can be embedded prospectively [89].

4.4 Layer 3: Model development and governance

The third layer covers development and internal validation, calibration and external validation, fairness and subgroup audit, and release gating with ongoing monitoring. It is the layer that the readmission literature addresses most directly, and the published reviews show that it does so incompletely: calibration is reported less often than discrimination [20], external validation is uncommon [18,19], and reporting is often insufficient for reconstruction [21]. The framework positions governance as part of development rather than as post-hoc compliance, consistent with work on release gating and automated evaluation of models before and after deployment [73], with frameworks for documenting data curation during model fitting [74], and with the TRIPOD+AI requirement that model specifications be reported in full [43]. For models that will influence care, the DECIDE-AI guideline for early-stage clinical evaluation of AI-driven decision support provides the bridge between technical validation and prospective use [53]. Threat models developed for agentic AI workflows, covering tool poisoning and manipulation of memory and inputs, are a reminder that a deployed model's inputs are an attack surface as well as a data pipeline [90].

4.5 Layer 4: Action and care-coordination capacity, and the feedback path

The fourth layer covers transitional care and handoff, pharmacy-led continuity and reconciliation, workforce capacity and staffing, and follow-up through telehealth or community care. It explains why discrimination alone is a poor guide to value: a score changes outcomes only through the actions it triggers, and those actions require capacity that varies independently of model quality. Care-coordination frameworks specify handoff mechanics [6,7], pharmacy-led access models supply continuity [91], workforce governance determines whether protocols are executed [92], and telehealth adoption frameworks for underserved populations extend reach [93]. Reviews of telehealth effectiveness across rural and urban settings indicate how far such follow-up can realistically compensate for distance [94], and workflow automation in small practices shows how patient engagement can be scaled when staff time is the constraint [95].

One feedback path runs from observed readmissions back to the measurement layer, recalibrating the model against local case mix and revealing which unmeasured determinants drive residual error. The published reviews describe a literature in which models are typically developed, reported, and left static, with little evidence of scheduled updating [18,19]. Engineering domains offer templates for continuous recalibration, including sensor-based predictive energy management with ongoing model updating [96] and tiered digital twins that bring predictive diagnostics to resource-constrained sites [97].

4.6 Testable propositions

The framework yields four propositions (Table 6). None can be tested with the literature as it stands, because the required infrastructure and capacity variables are rarely reported; reporting them is itself a recommendation of this review. The propositions are deliberately falsifiable. If, for example, external validation performance proved unrelated to linkage completeness and laboratory availability, the infrastructure layer would lose its claimed role as a determinant of transportability, and the framework would need to locate the loss of performance on external validation elsewhere, for instance in case-mix differences alone. External validation of a widely implemented proprietary sepsis model in a new health system [47], and the deployment of an automated deterioration alert across a large integrated system [46], illustrate in adjacent fields the kind of natural experiment in which such propositions can be examined.

Table 6. Testable propositions derived from the MRPR framework.

Proposition	Layer	Statement	Basis in the cited literature	Design needed to test
P1	Measurement	Within a fixed modeling method, discrimination increases with the number of measurement domains represented, and the increment from function, medication, and social domains exceeds that from more flexible estimators	Direct associations of gait, falls, function, and social gradient with readmission [3,4,8,9]; no significant regression versus machine learning difference in a prior review [20]	Individual participant data meta-analysis or single-cohort study with nested model comparison
P2	Infrastructure	External validation performance degrades in proportion to the infrastructure gap between development and validation settings	Performance loss on independent validation in an adjacent task [47]; input dependence of record-based scores [69,70]	Multi-site external validation recording linkage completeness and laboratory availability
P3	Action	The clinical value of a given C-statistic is conditional on care-coordination capacity	Positive and negative impact studies of alerts with and without structured responses [46,71]	Decision curve analysis and cluster trials across systems with differing transitional care capacity
P4	Governance and feedback	Models embedded in a recalibration loop maintain calibration over time, whereas static models drift with case-mix change	Calibration is setting-specific and rarely reported [20,44]	Temporal validation with scheduled recalibration versus a static comparator

4.7 Relation to existing frameworks

The MRPR framework is intended to complement, not replace, established models. Relative to the chronic care model [40], it narrows the focus to one decision-support component and makes

explicit the dependencies that the chronic care model leaves implicit: a decision-support tool is only as informative as the clinical information system that feeds it, and only as useful as the delivery system design that acts on it. Relative to PROBAST [42], which asks whether a model's performance estimate is at risk of bias in the study that produced it, MRPR asks whether the conditions that produced that estimate exist in the setting where the model will be used. The two questions are distinct: a model can be at low risk of bias in its development study and still fail in a new setting whose linkage, laboratory, and staffing conditions differ. Relative to conceptual models explaining why scientific innovation fails to improve patient outcomes without delivery infrastructure [15], MRPR specifies the particular infrastructure elements on which prognostic modeling depends and links each to an observable property of the data or the care process. The framework is therefore narrower than these antecedents but more directly testable, which is its intended contribution.

5. Discussion

5.1 Principal arguments

This review advances four arguments. First, the evidence summarized by published systematic reviews shows that readmission models discriminate modestly, that calibration is reported far less often than discrimination, and that external validation is uncommon [18-20]; direct evidence in older inpatients, among whom multimorbidity is the norm, shows widely used scores discriminating poorly [8]. Second, the ceiling on performance in multimorbid adults plausibly lies more in what routine data measure than in how models are estimated, because the domains most clearly associated with readmission in these patients are those least often recorded. Third, calibration and decision-analytic evaluation matter more than discrimination for the decisions readmission models are asked to support, and the worked illustration in Section 3.8 shows why modest discrimination at realistic readmission rates yields low positive predictive value. Fourth, reconstructability and transportability, rather than reported discrimination, determine whether a published model can be adopted, and both are constrained by reporting practice and by the infrastructure of the adopting setting.

Together these describe a literature that is productive but not cumulative. Each new study adds a model; few add knowledge the next study can build on, because the information required to validate, recalibrate, or extend a published model is usually missing [43]. The response proposed here is not another model but a change in what is measured, what is reported, and what is evaluated.

5.2 Comparison with prior reviews

The earliest general review judged the evidence too heterogeneous to pool and concluded that most models performed poorly [18]. The update reported moderate discrimination mainly among condition-specific models [19], and the review of electronic record models reported a modest advantage for richer data and no significant advantage for machine learning [20]. The paediatric review addressed a population too different to serve as a direct comparator, but its finding that reporting of implementation-relevant items was incomplete applies equally to adult models [21]. The ordering implied by these reviews, with condition-specific models tending to discriminate better than general ones, is expected. Restricting to a single index condition removes between-patient heterogeneity that a model does not then have to explain; multimorbid populations reinstate it, and they do so in the dimensions (function, frailty, medication, social context) that routine data capture least well.

This review adds three elements that the prior syntheses did not provide: a focus on populations defined by multimorbidity, a structured translation of PROBAST, TRIPOD+AI, and related standards into appraisal questions specific to readmission models, and an explicit analysis of the infrastructure and care-coordination conditions on which performance depends. The methodological weaknesses discussed here are also not unique to readmission research: systematic appraisal of early warning scores found widespread analysis-domain flaws and poor reporting [23], and appraisal of models predicting impairment after critical illness reached similar conclusions [22]. Prognostic modeling in acute and post-acute care shares a common methodological deficit that reporting standards have not yet corrected.

5.3 Why performance remains modest

The ceiling appears substantive rather than technical. Readmission in multimorbid patients reflects an interaction of baseline vulnerability, acute decompensation, and post-discharge care processes, and two of those three are determined after the prediction is made. A model built at

discharge cannot observe whether a follow-up appointment was kept, whether a prescription was filled, or whether a caregiver was available in the second week at home. Moving prediction from admission to discharge adds information about the hospital episode but still omits the period in which most of the missing signal arises.

The literature on post-discharge and inpatient care processes makes the point concrete. Multidisciplinary collaboration changes outcomes in emergency care [98], postoperative pain management depends on integrated nursing communication pathways [99], patient safety in high-acuity units depends on integrated clinical policies [100], and wound care in acute settings follows evidence-based protocols whose adherence varies across units [101]. Process variables of this kind are rarely available to readmission models, yet each plausibly separates a patient who returns from one who does not. The chronic care model anticipates this: outcomes emerge from interactions between prepared teams and activated patients, and a model trained on the characteristics of the patient alone omits the team [40].

5.4 Measurement as the binding constraint

Adjacent domains offer analogues for the argument that measurement, not method, is the binding constraint. Business intelligence applications have been proposed to improve the allocation of mental health and substance use resources by making previously unrecorded program data visible [102,103]; informatics frameworks for protecting vulnerable populations depend on population-level social data [104]; and public health governance models built on process metrics make the same argument for accountability [105]. The social gradient in readmission among patients with chronic conditions is real and measurable [3], and an integrative maternal health model [106] and a trauma-informed framework for adolescents in underserved communities [107] describe how social and psychosocial information can be built into program design from the outset. What readmission research lacks is less a method of measurement than a decision to measure.

That decision has costs. Structured functional assessment, frailty screening, medication reconciliation that records access and adherence, and social risk screening each require staff time and workflow change. The argument of this review is not that every hospital should collect every domain, but that the value of collecting them should be tested directly, within cohorts, against the value of more complex modeling of the data already held [8,20].

5.5 Transportability as the central problem

A model that performs well where it was built and poorly elsewhere is of limited use to the health systems that would adopt it. The appraisal in Section 3.7 and the checklist in Table 4 frame transportability as a property that must be demonstrated rather than assumed. Evidence from adjacent prediction tasks shows how large performance shifts can be when a model moves between systems [47], and reviews of deterioration prediction show how rarely such shifts are measured [24,68]. Abnormal vital signs often precede in-hospital cardiac arrest [108], and delayed activation of rapid response teams is associated with worse outcomes [109], which shows that even a well-discriminating signal is only as useful as the response system attached to it.

For readmission, this argues for models whose inputs are standardized across sites, for mandatory local recalibration whenever they are not, and for a period of silent operation in which predictions are recorded but not acted upon, so that local discrimination and calibration can be measured before any patient is affected. Guidance on the evaluation of prediction models treats updating of the intercept and slope as a routine step when a model is moved, not as a sign of failure [44]. The same guidance implies that a health system should prefer validating and recalibrating an existing model to developing a new one when its data are limited.

5.6 From discrimination to impact

Little of the readmission prediction literature measures whether using a model changes patient outcomes. The closest evidence comes from inpatient deterioration and sepsis, where impact studies are more advanced and instructive in both directions. A randomized trial of a real-time deterioration alert on general medical wards did not demonstrate improvement in hard outcomes [71], whereas deployment of an automated early warning program across a large integrated system was associated with lower mortality when the alert was coupled to a structured clinical response [46]. A multicenter clinical intervention study of a machine learning early warning score reported reduced hospital mortality [110], a prospective multi-site study associated timely clinician response to a sepsis alert with better outcomes [111], and a deep learning sepsis model was associated with improved quality of care and survival after deployment [112]. Implementation of a machine learning early warning system in an Australian hospital describes the operational work required to realize such benefits [113].

The common feature of the positive studies is that the score triggered a defined action by a team with capacity to act; the negative trial is a reminder that an alert without a workflow changes little. A longitudinal review of implementing AI to predict deterioration in adult hospitals concludes that implementation outcomes remain the least studied part of the pathway [114]. For readmission prediction, the implication is that impact evaluations should be designed and reported to the standards of CONSORT-AI for trials of AI-enabled interventions [52], appraised with RoB 2 [49] or ROBINS-I [48] according to design, and preceded by early-stage clinical evaluation of the kind DECIDE-AI describes [53].

5.7 Medication determinants as an under-modeled domain

Medication determinants are the clearest example of a first-order cause that prediction models omit. Stewardship frameworks govern prescribing quality at system level [115], medicine quality failures carry documented clinical consequences [116], traceability and chain-of-custody integrity determine whether a dispensed product is authentic [117], and field-deployable methods now allow substandard and falsified products to be detected outside reference laboratories [118]. Value-based pricing [119] and biosimilar substitution [120] shape affordability, while quality assurance in manufacturing is increasingly understood as a determinant of supply reliability [121]. At the regulatory level, risk-based decision-making in manufacturing environments [122], risk-based interdiction of illicit medical products [123], harmonization of medicine quality assurance [124], and supplier risk assessment in international procurement [125] each influence whether the medicine a discharged patient receives is available, affordable, and of assured quality.

The patient-level evidence is equally relevant. Comprehensive medication review under Medicare Part D has been evaluated for its effects on adherence among patients with Alzheimer's disease [63] and on racial and ethnic disparities in statin nonadherence in the same population [126], and adherence among Medicare-enrolled older adults with asthma and chronic obstructive pulmonary disease has been compared before and during the COVID-19 pandemic [62]. These are multimorbid populations in which medication behavior is measurable from pharmacy data, yet readmission models typically reduce medication information to a count of drugs. A model that omits whether the patient can obtain and tolerate their medication omits a plausible cause, not merely a correlate.

5.8 Infrastructure as a precondition for prediction

Model performance plausibly tracks the maturity of the data environment, which aligns with a literature that treats infrastructure as a binding constraint on health system performance. Infrastructure has been framed as a public health intervention in its own right, drawing on evidence from large laboratory networks [57], and converting infrastructure investment into population health outcomes requires explicit measurement frameworks [58]. Capital project delivery for high-risk facilities [127], program management for coordinated multi-site expansion [128], design of high-performance laboratories for pandemic preparedness [129], and accountability in public-private partnerships [130] describe the delivery mechanics. Risk-managed construction of complex facilities [131], sustainable materials and energy strategies [132], operational standards for blood collection networks [133], and compliance frameworks for healthcare infrastructure projects [134] shape the environment in which clinical data are generated.

The consequence for prediction is direct. A model trained where laboratory turnaround is unreliable, diagnostic coverage incomplete, or record linkage fragmented inherits those defects as noise. Integrated laboratory information systems [16] are therefore relevant to prediction not as analogies but as upstream determinants of input quality, and data-driven preparedness frameworks for urban hospitals [135] show how facility readiness can be assessed systematically before a new tool is introduced. Few published readmission models describe the diagnostic environment that generated their laboratory inputs, so the literature cannot yet test whether input quality explains part of the variation in performance between settings. A practical corollary is that health systems considering a published model should compare their own laboratory coverage, coding practices, and linkage completeness with those of the development setting before adoption. Program management approaches for multi-site infrastructure [128] offer a transferable template for staging such a rollout across facilities with different levels of readiness.

5.9 Governance, privacy, security, and the workforce conditions for deployment

Deployment introduces a second set of constraints. Privacy-preserving governance models comparing blockchain and cryptographic strategies address how predictive systems can use sensitive data lawfully across jurisdictions [136], and legal and ethical risk modeling for enterprise data protection supplies a complementary governance lens [137]. Security is not

peripheral: cyber risk pricing for health technology platforms [138], forensic techniques for detecting cyber-enabled fraud in telemedicine [139], and structured frameworks for clinical record protection [140] describe the residual risk landscape. Continuous monitoring of cloud misconfiguration [141] and evidence on information-security risk from personal devices in administrative offices [142] show how quickly an unmonitored environment accumulates exposure, and quantitative models of compliance risk in emerging economies [143] show how that exposure can be scored. A deployed readmission model inherits all of these risks through its data pipeline.

A risk score changes nothing unless someone acts on it. Workforce planning governed by explicit nurse staffing frameworks establishes the capacity precondition [92]. Work outside health care supplies analogues for managing that capacity: frameworks for staffing volatility during rapid expansion of supply chain operations [144], enterprise risk models integrating human resource strategy with operational analytics [145], and structured talent development across distributed operations [146]. Workforce safety training models show how operational risk falls when training is systematic rather than ad hoc [147]. Resilient health system frameworks argue that organizational strategy, operational excellence, and workforce capability matter more than technology adoption alone [148], which is the MRPR Layer 4 argument stated from the health system side.

5.10 The choice of outcome: readmission and its alternatives

Thirty-day readmission is a convenient outcome for payers but an imperfect one for patients. It counts a planned return that went well the same as an avoidable crisis, it is censored by death, and it misses time spent in post-acute facilities. Patient-centered alternatives have matured. Days alive and at home has been validated from routinely collected data after hip fracture [149] and associated with long-term survival and functional status after mechanical ventilation [150]. For multimorbid patients whose index admission involved critical illness, a group in which chronic critical illness is a substantial burden [151], the relevant post-discharge trajectory includes physical, cognitive, and mental health impairments that a stakeholders' conference identified as the core of long-term recovery [152] and that an international consensus conference sought to predict [153].

Evidence on interventions in this population is mixed, which bears on what a risk score could trigger. Increased hospital-based rehabilitation and information after intensive care discharge did not improve the primary outcome in the RECOVER trial [154], a meta-analysis questioned whether enhanced physical rehabilitation after ICU discharge improves outcomes [155], and patients describe services they found helpful and others they found missing [156]. The ABCDEF bundle illustrates how an organized care process, rather than a score, carries the effect [157]. Caregiver burden, driven by post-intensive care syndrome components in the patient [64], with its own risk factors in family members [65], frequency and determinants in sepsis survivors [158], and responsiveness to follow-up psychological intervention [159], is a further determinant of whether a discharged multimorbid patient can be supported at home. Readmission models that ignore caregiver capacity omit one of the social variables most directly tied to the outcome.

5.11 Convergent evidence from adjacent domains

Several arguments of this review echo patterns in adjacent fields, which suggests they are structural rather than artifacts of the readmission literature. Surveillance faces the same data-quality constraint: evaluation of integrated disease surveillance showed that indicator quality, not analytic method, limits what surveillance detects [72], and real-time surveillance dashboards [160] and epidemic preparedness systems [161] rest on the same interoperability precondition. Detection technologies follow the same logic: AI-enabled biosensors [162], One Health framing of environmental vectors [163], and automated detection in mammography [164] extend what can be measured rather than refining inference on existing data. Diagnostic integration across microbiology, virology, and serology workflows [165] and droplet digital PCR for viral reservoir quantification [166] show the measurement frontier moving in laboratory medicine.

Education and regulatory science supply organizational analogues. Ethical data practice in health education technology [167], real-time analytics dashboards [168], data-informed learning platforms [169], and evidence-based frameworks for preserving critical thinking [170] and assessment integrity [171] in the era of generative AI all depend on governance of what is measured and how outputs are used. Readiness maturity models for regulatory entry [172], risk-based accelerated approval [173], cross-border regulatory harmonization [174], and supply chain governance [175] ask the same underlying question as this review: how much confidence is warranted in a decision system given the evidence feeding it. Primary care mental health

screening [176], digital operations models for medical information flow [177], the psychosocial effects of compressed workweeks on workers [178], real-time informatics that reduce decision delays [179], and analytics-driven workflow performance models [180] complete a set of system-level variables that a model operating in isolation cannot see. Work on the comparative bioavailability of botanical compounds [181] and the pharmacological characterization of traditional botanicals [182] points to a further blind spot, the use of herbal and traditional remedies, which pharmacy-based predictors do not record.

5.12 Communicating model performance to decision makers

A C-statistic is frequently misread as the accuracy of a model. For a decision maker, four points matter. First, the statistic describes ranking, not the correctness of predicted probabilities, which is the role of calibration, and not whether acting on the predictions does more good than harm, which is the role of net benefit [44]. Second, a C-statistic estimated in a development sample overstates what the same model will achieve elsewhere, and the gap can be large when the model moves to a different system [47]. Third, an average across settings, even if one were available, should not be expected in any particular setting, because between-setting variation is typically wide and is described by a prediction interval rather than a confidence interval. Fourth, the appropriate benchmark for a readmission model is not an abstract threshold of acceptability but the comparison with current practice in the setting where it will be used, including clinician judgment and simple rules.

Communicating these points requires translating statistics into operational terms. The worked illustration in Table 5 shows how a model of modest discrimination translates, at a realistic readmission rate [1], into the proportion of flagged patients who will and will not be readmitted. Presenting model performance to program managers in this form, with the assumptions stated, makes the trade-offs visible and ties the choice of threshold to the cost and burden of the action it triggers, as TRIPOD+AI asks developers to do [43].

5.13 Scope and contribution of the review

The review has features that bear on how its arguments should be weighted. It anchors every quantitative statement either in a named published source or in a labeled worked illustration, so that readers can check each number against its origin. It applies the named appraisal and reporting standards explicitly and translates them into questions specific to readmission models

(Table 2), so that its appraisal can be reproduced by others on any model they are considering [41-43]. It distinguishes throughout between what the cited evidence shows, what prior reviews found, and what is argued or hypothesized, and it presents its central interpretive claims as testable propositions rather than conclusions. It also separates the clinical prediction evidence from the interpretive synthesis of infrastructure and governance literature, so that readers can accept the appraisal framework without accepting the MRPR framework, or the reverse [14]. The review does not, however, provide what a systematic review with meta-analysis would provide, and its contribution should be judged as a structured argument and a set of tools for appraisal rather than as an estimate of model performance.

6. Implications for Practice and Policy

6.1 Implications for clinical practice

Because models of modest discrimination at realistic readmission rates flag many patients who will not be readmitted (Section 3.8), the case for using a score rests on what the flag triggers. For low-cost, broadly beneficial actions the arithmetic is favorable even at modest accuracy: medication reconciliation, structured follow-up, fall prevention embedded in nursing policy [183], infection prevention compliance [184], evidence-based wound care in acute settings [101], and mobile monitoring with adherence support [185]. For scarce, expensive programs such as intensive case management, the same arithmetic is unfavorable, and a score should not be the sole gatekeeper.

Three recommendations follow. Use risk scores to rank patients for attention rather than to gate access to services. Pair any score with clinical judgment, since clinicians observe function, cognition, caregiver availability, and medication behavior that models built on routine data capture poorly [8]. Recalibrate locally before deployment and at scheduled intervals thereafter, and check calibration specifically in the highest-risk groups, where enrollment decisions concentrate. These recommendations are consistent with TRIPOD+AI, which asks developers to report the intended use and the threshold rationale explicitly [43].

6.2 Implications for health system investment and operations

The MRPR framework implies a sequencing rule for investment. A system lacking reliable cross-site linkage, dependable laboratory turnaround, and a functioning interoperability layer should build those before commissioning a predictive model, because a model built on weak inputs will underperform and its apparent failure will discredit an approach that was never given usable data. Resilient health system frameworks make the same argument from the delivery side [148]. Workforce capacity deserves specific attention: nurse staffing and workforce planning determine whether transitional care protocols are carried out [92], and without capacity to act, even an accurate list of high-risk patients produces no change in outcomes.

Operational leaders should also treat a deployed model as a controlled system with monitoring, audit, and corrective action. Release gating and automated evaluation [73], corrective and preventive action methodology for remediation [60], and data integrity controls from regulated industries [59] supply a governance template that is more mature than anything typically described in readmission model publications. The additional cost of these controls is small relative to the cost of a miscalibrated model directing transitional care away from the highest-risk patients for months before anyone notices.

6.3 Implications for policy, payment, and reporting standards

The policy context shapes what a readmission model is for. Health insurance financing and provider payment reform determine the balance between coverage expansion and financial protection [186], and real-world evidence and predictive analytics are increasingly the means by which payment programs are evaluated [56]. This connects to an unresolved tension in risk adjustment. Multimorbidity measured with a weighted index predicts both readmission and post-discharge mortality among Medicare beneficiaries [4], so hospitals serving high-need populations may incur penalties that reflect case mix rather than performance when such burden, together with disability and social complexity, is incompletely captured; yet including social determinants in payment formulas risks normalizing lower standards for disadvantaged populations. The arguments of this review support a separation of purposes. Social and functional variables are likely to improve the clinical targeting of transitional care. Whether they belong in a payment formula is a distinct normative question that prediction accuracy cannot settle, and conflating the two has produced much of the current confusion.

Three reporting changes would materially improve the cumulative value of the literature at negligible cost: publication of calibration plots, slopes, and intercepts as a minimum; publication of full model specifications sufficient for independent implementation; and description of the infrastructure characteristics of the development setting, including linkage completeness and data provenance. The first two are already required by TRIPOD+AI [43]; journals and funders could make compliance a condition of publication and funding.

6.4 Implications for model developers, reviewers, and journals

Developers can act on these arguments immediately. Before building a new model, they should search for existing models and consider external validation and recalibration of the most promising candidate, which will often add more than another development study [44]. Where development is justified, sample size should be planned for the number of candidate predictors rather than assumed adequate because the cohort is large, missing data should be handled by multiple imputation rather than complete-case analysis, continuous predictors should be kept continuous, and predictor selection should not rely on univariable screening; these are the analysis-domain practices that PROBAST identifies as sources of bias [42]. Data curation decisions made during model fitting should be documented, as frameworks for context-aware curation propose [74], and code and full specifications should be released. Peer reviewers and editors can require calibration plots, a statement of intended use and threshold rationale, and subgroup performance by at least payer and deprivation, and can use Table 2 as a checklist. None of these requirements is novel; all are in TRIPOD+AI [43]. The gap is enforcement.

7. Equity Considerations

Equity considerations follow directly from the appraisal in Section 3.11. National statistics document differences in readmission rates by race and ethnicity and by payer [2], and among patients with chronic conditions readmission follows a social gradient [3]. A model that underpredicts in disadvantaged groups will direct transitional care away from the patients whose absolute risk is highest, and deploying a model without subgroup calibration risks exactly that. The predictors most associated with social disadvantage, namely function, social support, and medication access, are those least often measured, so their influence is absorbed into residual

error. Analyses in other sectors show that a forecast can be accurate overall and inequitable across groups [75], and within health care, evaluation of Medicare comprehensive medication review has examined whether a single service narrows racial and ethnic disparities in statin nonadherence among patients with Alzheimer's disease [126].

Prespecified equity analysis should therefore be a release requirement rather than an optional supplementary analysis, consistent with governance frameworks for accountability in AI-driven healthcare management [51] and with systematic work on digital equity and access frameworks in clinical training [76]. Minimum content should include discrimination and calibration by race and ethnicity, payer, deprivation, sex, and age band; a documented decision on whether group-specific recalibration is warranted; and monitoring of these metrics after deployment. Equity also has a geographic dimension. The readmission prediction literature summarized by the general reviews comes predominantly from high-income health systems [18,19], so models are being developed where linkage and laboratory infrastructure are most mature and are least applicable where readmission definitions, follow-up capture, and care pathways differ most. Informatics frameworks for vulnerable populations [104] and integrative maternal health models addressing rural socioeconomic inequity [106] indicate how measurement of social risk can be built into programs serving the populations currently absent from the evidence.

8. Limitations

8.1 Limitations of the evidence

The evidence on which this review draws has limitations that constrain every argument it makes. None of the published systematic reviews of readmission models defined its population by multimorbidity [18-20], so their findings are indirect for the question posed here, and the direct evidence in older and multimorbid adults comes from a small number of cohorts in particular settings [4,8,9]. The prior reviews themselves combined models with different outcome windows, populations, and validation types, so the ranges they report describe the variety of the literature rather than the performance to be expected of any one model [19]. Calibration and clinical utility are rarely reported in the primary literature, so statements about them rest on methodological reasoning more than on observed data [44]. Evidence on the impact of using a

prediction model on patient outcomes comes largely from inpatient deterioration and sepsis rather than from readmission [46,110], and its transfer to readmission is an analogy, not a demonstration. Finally, predictor timing is a recurring concern: PROBAST requires that each predictor be available at the stated time of use [42], and a model intended for use at admission that relies on a variable such as length of stay would overstate what is achievable at that moment, so readers adopting an admission-time model should verify predictor timing directly.

8.2 Limitations of this review

This is a critical methodological review, and its design has consequences. The literature was identified purposively and by citation chaining, not by a registered and reproducible database search, and screening and appraisal were not performed in duplicate against prespecified eligibility criteria as PRISMA 2020 describes for systematic reviews [45]. Relevant studies may therefore have been missed, and the selection of sources may reflect the authors' judgment about what illustrates each point. The review does not pool performance estimates and does not tally risk-of-bias judgments across a defined set of studies, so it cannot say how well the average model performs in multimorbid adults or how common any particular flaw is. The predictor-domain argument is a reasoned case from cited evidence, not a measured difference between models, and it could be overturned by within-cohort comparisons. The worked illustration in Table 5 uses hypothetical operating points and a single national readmission rate, and different assumptions would yield different values. The analysis of infrastructure, governance, and workforce conditions is interpretive; the supporting literature is largely conceptual and framework-based, much of it from cross-sector analogues [14], and the MRPR framework should be read as a set of hypotheses rather than as a validated model.

9. Future Research Agenda

The most consequential gap is not predictive. Little of the readmission prediction literature measures whether using a model changes patient outcomes. Stepped-wedge or cluster-randomized designs comparing model-guided transitional care with usual targeting would answer the question the field has largely left unasked, and they should report net benefit at operationally

realistic thresholds rather than discrimination alone, following CONSORT-AI [52] and appraised with RoB 2 [49]. Where randomization is infeasible, well-designed non-randomized evaluations appraised with ROBINS-I [48] are the next best option.

Table 7. Prioritized research agenda.

Priority	Question	Recommended design	Primary outcome and reporting standard
1	Does model-guided transitional care reduce readmission and improve patient-centered outcomes compared with usual targeting?	Cluster-randomized or stepped-wedge trial	Readmission and days alive and at home; net benefit; CONSORT-AI, RoB 2
2	How much do functional, medication, and social predictors add within the same cohort and method?	Individual participant data meta-analysis; prospective cohorts	Change in C-statistic, calibration, and net benefit; TRIPOD+AI
3	How well do existing models perform in multimorbidity-defined populations, and how much does performance vary between settings?	Systematic review with meta-analysis restricted to multimorbid populations	Pooled C-statistic with prediction interval; calibration summaries; PROBAST
4	Do existing models transport, and does local recalibration restore calibration?	Multi-site external and temporal validation	Calibration slope and intercept before and after recalibration; PROBAST
5	Are models fair across race and ethnicity, payer, and deprivation?	Prespecified subgroup validation with parity metrics	Group-specific discrimination and calibration
6	Does infrastructure explain transportability (MRPR proposition 2)?	Validation studies reporting linkage completeness and laboratory availability	Performance change regressed on infrastructure gap
7	Is earlier prediction in primary care more effective than discharge prediction?	Comparative effectiveness study across prediction timing	Admissions prevented per patient flagged
8	Do deployed models drift, and does monitoring detect it?	Post-deployment surveillance with scheduled recalibration	Calibration drift over time; release-gating documentation

Measurement is the second priority. Prospective cohorts collecting frailty, cognition, functional status, medication access, and social risk at admission in multimorbid populations would

establish the practical ceiling on discrimination when measurement is not the constraint; frameworks for systematic psychosocial screening in primary care [176] and adolescent services [107] indicate how such data collection can be organized. External validation with documented local recalibration now adds more than new development, and journals should treat it accordingly. A systematic review with individual participant data, designed around the appraisal framework in Table 2 and restricted to multimorbidity-defined populations, would provide the pooled estimate and prediction interval that this review deliberately does not attempt [44]. Comparative studies across prediction timing are also needed: a model applied in primary care months before an admission, even at lower discrimination, may prevent more admissions than a better model applied at discharge, and AI-driven early detection systems for non-communicable diseases [187] and pharmacy-led access models [91] provide plausible delivery vehicles. Table 7 sets out the agenda in order of priority.

The field also needs subgroup calibration reported as standard, fairness audits using established parity metrics, and evaluation of whether deployed models drift as case mix changes. Continuous monitoring methods already exist in adjacent domains, including automated model governance with release gating [73], and require adaptation rather than invention. Each item in Table 7 corresponds to a layer of the MRPR framework, so progress on the agenda would also provide the data needed to test the framework itself.

Outcome definition deserves its own line of work. Readmission should be complemented by patient-centered measures such as days alive and at home, which can be derived from routinely collected data [149] and which track long-term survival and function after critical illness [150]. A model that reduces readmissions by shifting patients into post-acute facilities, or that performs well only because deaths are censored, would look successful on the conventional outcome while failing patients. Joint modeling of readmission, death, and time at home in multimorbid cohorts would resolve this ambiguity and would allow the competing-risk structure of post-discharge outcomes to be represented explicitly rather than excluded. Studies in post-intensive care populations, where prediction of long-term impairment has itself been systematically appraised [22], offer a template for the outcome sets such work should use.

10. Conclusion

Models predicting thirty-day readmission in adults with multiple chronic conditions should be judged by more than their discrimination. The published evidence describes a field in which discrimination is usually modest, calibration is seldom reported, external validation is uncommon, and many models cannot be reconstructed from their publications, and in which widely used scores discriminate poorly in older inpatients. This review argues that the ceiling on performance lies mainly in what routine data measure: functional status, frailty, falls, medication access, social circumstances, and caregiver capacity are repeatedly linked to readmission but rarely recorded, and better estimators applied to the same data are unlikely to close the gap. A readmission risk score for a multimorbid adult should be read as one input among several: useful for ranking, insufficient for deciding, and trustworthy only after local validation, recalibration, and subgroup audit in the setting where it will be used. Until the measurement and infrastructure layers improve, and until the care system has the capacity to act on a flag, better predictions will leave outcomes largely unchanged. The appraisal framework, transportability checklist, and testable propositions offered here are intended to make those conditions explicit and to direct the next generation of research toward validation, measurement, and impact rather than toward further model development.

Ethics approval and consent to participate

Not applicable. This study is a review of published literature and involved no human participants.

Consent for publication

Not applicable.

Availability of data and materials

No new data were generated or analyzed in this review. All sources discussed are published and are listed in the reference list; the worked illustration can be reproduced from the assumptions stated in the text.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

OC conceptualized the study and led the literature synthesis and drafting; FE and ES contributed to the analysis, interpretation, and critical revision of the manuscript. All authors read and approved of the final manuscript.

Acknowledgments

Not applicable.

References

1. Jiang HJ, Barrett ML. Clinical conditions with frequent, costly hospital readmissions by payer, 2020. Agency for Healthcare Research and Quality; 2024. Available from: https://hcup-us.ahrq.gov/reports/statbriefs/sb307-readmissions-2020.jsp
2. Jiang HJ, Hensche MK. Characteristics of 30-day all-cause hospital readmissions, 2016-2020. Agency for Healthcare Research and Quality; 2023. Available from: https://hcup-us.ahrq.gov/reports/statbriefs/sb304-readmissions-2016-2020.pdf
3. Zumbrunn A, Bachmann N, Bayer-Oglesby L, Joerg R, SIHOS Team. Social disparities in unplanned 30-day readmission rates after hospital discharge in patients with chronic health conditions: a retrospective cohort study using patient level hospital administrative data linked to the population census in Switzerland. PLOS ONE. 2022;17(9):e0273342. doi:10.1371/journal.pone.0273342
4. Wei MY. Multimorbidity, 30-day readmissions, and postdischarge mortality among Medicare beneficiaries using a new ICD-coded multimorbidity-weighted index. The Journals of Gerontology: Series A. 2023;78(4):727-734. doi:10.1093/gerona/glac242
5. Nnabuko OM. Integrating triglyceride-glucose index and echocardiographic parameters for improved cardiovascular risk stratification in Sub-Saharan Africa. International Journal of Cardiology. 2022;1(2):1-16. doi:10.34218/IJC_01_02_001
6. Akinlolu VS, Fapohunda M, Omaghomi TT, Atima ME, Igweonu C. A proposed care-coordination framework for reducing readmissions among chronic disease patients. International Journal of Multidisciplinary Research and Growth Evaluation. 2023;4(5):1187-1195. doi:10.54660/.IJMRGE.2023.4.5.1187-1195
7. Igweonu C, Omaghomi TT, Fapohunda M, Akinlolu VS, Atima ME. A strategic policy framework for enhancing continuity of care in patients with cardiometabolic conditions. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):2779-2790. doi:10.62225/2583049X.2024.4.6.5406

8. Sun CH, Chou YY, Lee YS, Weng SC, Lin CF, Kuo FH, et al. Prediction of 30-day readmission in hospitalized older adults using comprehensive geriatric assessment and LACE index and HOSPITAL score. International Journal of Environmental Research and Public Health. 2023;20(1):348. doi:10.3390/ijerph20010348
9. Tsai YC, Chen YM, Wen CJ, Wu MC, Chou YC, Chen JH, et al. Multimorbidity and prior falls correlate with risk of 30-day hospital readmission in aged 80+: a prospective cohort study. Journal of the Formosan Medical Association. 2023;122(11):1111-1116. doi:10.1016/j.jfma.2023.03.009
10. Asiedu CS, Asiedu AA. Equitable access to chronic therapies in hospital pharmacy settings: A conceptual framework for continuity of care. International Journal of Medical Evaluation and Physical Report. 2024;8(6):274-298. doi:10.56201/ijmepr.v8.no6.2024.pg274.298
11. Asiedu CS, Asiedu AA. Pharmaceutical supply chain inefficiencies and medication affordability in resource-constrained health systems: A narrative review. International Journal of Health and Pharmaceutical Research. 2023;8(4):192-222. doi:10.56201/ijhpr.v8.no4.2023.pg192.222
12. Asiedu CS, Asiedu AA. Last-mile drug distribution to underserved communities: A review of barriers and intervention models. Gyanshauryam, International Scientific Refereed Research Journal. 2022;5(6):439-466.
13. Asiedu CS. Drug interaction risk in antenatal care: A conceptual framework for systematic interaction screening. Shodhshauryam, International Scientific Refereed Research Journal. 2022;5(3):465-484.
14. Ilodigwe L, Adesemoye AC. The data backbone of health system transformation: A governance and architecture framework for interoperability, data quality, and advanced analytics at national scale. International Journal of Health and Pharmaceutical Research. 2021;6(2):52-76. doi:10.56201/ijhpr.vol.6.no2.2021.pg52.76
15. Ilodigwe L, Adesemoye AC. From molecules to health systems: A conceptual model explaining why scientific innovation fails to improve patient outcomes without effective healthcare delivery infrastructure. International Journal of Health and Pharmaceutical Research. 2019;5(2):56-79. doi:10.56201/ijhpr.vol.5.no2.2019.pg56.79
16. Abdulrazaq K, Yusuf HD, Osinem HE, Eribenne F. A model for improving infection detection accuracy through integrated laboratory information systems. Iconic Research and Engineering Journals. 2019;3(6):650-675. doi:10.64388/IREV3I6-1723089
17. Yusuf HD, Abdulrazaq K. A conceptual framework for strengthening diagnostic quality systems in resource-limited clinical laboratories. Iconic Research and Engineering Journals. 2018;1(11):145-179. doi:10.64388/IREV1I11-1723090
18. Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, et al. Risk prediction models for hospital readmission: a systematic review. JAMA. 2011;306(15):1688-1698. doi:10.1001/jama.2011.1515
19. Zhou H, Della PR, Roberts P, Goh L, Dhaliwal SS. Utility of models to predict 28-day or 30-day unplanned hospital readmissions: an updated systematic review. BMJ Open. 2016;6(6):e011060. doi:10.1136/bmjopen-2016-011060

20. Mahmoudi E, Kamdar N, Kim N, Gonzales G, Singh K, Waljee AK. Use of electronic medical records in development and validation of risk prediction models of hospital readmission: systematic review. BMJ. 2020;369:m958. doi:10.1136/bmj.m958
21. Niehaus IM, Kansy N, Stock S, Dötsch J, Müller D. Applicability of predictive models for 30-day unplanned hospital readmission risk in paediatrics: a systematic review. BMJ Open. 2022;12(3):e055956. doi:10.1136/bmjopen-2021-055956
22. Haines KJ, Hibbert E, McPeake J, Anderson BJ, Bienvenu OJ, Andrews A, et al. Prediction models for physical, cognitive, and mental health impairments after critical illness: a systematic review and critical appraisal. Critical Care Medicine. 2020;48(12):1871-1880. doi:10.1097/CCM.0000000000004659
23. Gerry S, Bonnici T, Birks J, Kirtley S, Virdee PS, Watkinson PJ, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. BMJ. 2020;369:m1501. doi:10.1136/bmj.m1501
24. Muralitharan S, Nelson W, Di S, McGillion M, Devereaux PJ, Barr NG, et al. Machine learning-based early warning systems for clinical deterioration: systematic scoping review. Journal of Medical Internet Research. 2021;23(2):e25187. doi:10.2196/25187
25. Adelanwa A, Basnet A, Anene UN. Advanced AI based decision support systems for healthcare operations and resource planning. Gyanshauryam, International Scientific Refereed Research Journal. 2024;7(1):244-276.
26. Adelanwa A, Basnet A, Anene UN. Performance intelligence models for optimization and outcome measurement in large scale public services. Shodhshauryam, International Scientific Refereed Research Journal. 2024;6(1).
27. Adelanwa A, Basnet A, Anene UN. Data driven digital transformation models for lifecycle performance management in infrastructure delivery. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2646-2662. doi:10.62225/2583049X.2023.3.6.5968
28. Adelanwa A, Basnet A, Anene UN. Predictive analytics models for financial risk detection and fraud prevention in public systems. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2663-2679. doi:10.62225/2583049X.2023.3.6.5969
29. Anene UN, Clement T. A resilient logistics framework for humanitarian supply chains: Integrating predictive analytics, IoT, and localized distribution to strengthen emergency response systems. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022;8(5):398-424.
30. Mayo W, Ogbole JI, Okoruwa PO, Babatope OM. Designing an AI-predictive maintenance model for e-commerce systems using machine learning and cloud analytics. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2454-2468. doi:10.62225/2583049X.2023.3.6.5418
31. Aliliele C, Mbonu IS, Iwuanyanwu U. Advances in predictive analytics models for student retention and institutional risk management systems. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2692-2711. doi:10.62225/2583049X.2023.3.6.5990

32. Adesuyi MO, Walawalkar G, Kalu A. Predictive budgeting models using operational and market signals. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022;8(5):818-837.
33. Edivri J, Oteri O. Predictive capacity planning and resource utilization forecasting models for multi-program and multi-stakeholder IT portfolios. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022;8(4):826-845.
34. Tonoyan A, Dada O, Ayivi-Donkor SS. Conceptual model for predictive cost optimization: machine learning applications in business transformation analysis. IIARD International Journal of Economics and Business Management. 2022;8(6):68-102. doi:10.56201/ijebm.v8.no6.2022.pg68.102
35. Michael ON, Ogunsola OE. Development of predictive data models for enhancing crop resilience and climate adaptation in agrarian economies. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(1):1510-1527. doi:10.62225/2583049X.2023.3.1.5143
36. Nnabueze SB, Filani OM, Sakyi JK, Okojie JS, Okereke M. Developing sustainable business models for large-scale service organizations using predictive and strategic insights. International Journal of Multidisciplinary Evolutionary Research. 2023;4(2):76-88. doi:10.54660/IJMER.2023.4.2.76-88
37. Oparah OS, Ezeh FE, Olatunji GI, Ajayi OO. AI-based risk stratification framework for large-scale public health emergency preparedness and response planning. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2021;7(1):332-366.
38. Oparah OS, Ezeh FE, Olatunji GI, Ajayi OO. Big data-enabled predictive models for anticipating infectious disease outbreaks at population and regional levels. International Journal of Advanced Multidisciplinary Research and Studies. 2022.
39. Oparah OS, Ezeh FE, Olatunji GI, Ajayi OO. Framework for integrating climate data and health outcomes to improve mortality risk prediction systems. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2024;10(2):1128-1150. doi:10.32628/CSEIT24102151
40. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. Milbank Quarterly. 1996;74(4):511-544. doi:10.2307/3350391
41. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Medicine. 2014;11(10):e1001744. doi:10.1371/journal.pmed.1001744
42. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Annals of Internal Medicine. 2019;170(1):51-58. doi:10.7326/M18-1376
43. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. doi:10.1136/bmj-2023-078378

44. Debray TPA, Damen JAAG, Snell KIE, Ensor J, Hooft L, Reitsma JB, et al. A guide to systematic review and meta-analysis of prediction model performance. BMJ. 2017;356:i6460. doi:10.1136/bmj.i6460
45. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71
46. Escobar GJ, Liu VX, Schuler A, Lawson B, Greene JD, Kipnis P. Automated identification of adults at risk for in-hospital clinical deterioration. New England Journal of Medicine. 2020;383(20):1951-1960. doi:10.1056/NEJMsa2001090
47. Wong A, Otles E, Donnelly JP, Krumm A, McCullough J, DeTroyer-Cooley O, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. JAMA Internal Medicine. 2021;181(8):1065-1070. doi:10.1001/jamainternmed.2021.2626
48. Sterne JAC, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919. doi:10.1136/bmj.i4919
49. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898. doi:10.1136/bmj.l4898
50. DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials. 1986;7(3):177-188. doi:10.1016/0197-2456(86)90046-2
51. Afrihyia E, Akinse SG, Ojukwu PU. Comparative governance of AI-driven healthcare management: executive oversight, regulatory structures, and accountability in the United States and developing countries. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):3087-3102. doi:10.62225/2583049X.2024.4.6.5920
52. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nature Medicine. 2020;26(9):1364-1374. doi:10.1038/s41591-020-1034-x
53. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nature Medicine. 2022;28(5):924-933. doi:10.1038/s41591-022-01772-9
54. Ekemezie UC, Ohunyon G, Sydney E. KPI-based quality monitoring in long-term care: A three-tier accountability model. Shodhshauryam, International Scientific Refereed Research Journal. 2023;6(6):429-475.
55. Ekemezie UC, Ohunyon G, Sydney E. HEDIS measures for geriatric care quality: Evidence, equity, and a governance framework. Gyanshauryam, International Scientific Refereed Research Journal. 2023;6(1):447-488.
56. Ilodigwe L, Adesemoye AC. Advances in real-world evidence, predictive analytics, and data-driven decision making for improving health system performance and patient outcomes. International Journal of Health and Pharmaceutical Research. 2024;9(5):174-197. doi:10.56201/ijhpr.v9.no5.2024.pg174.197

57. Aminu-Ibrahim AY, Ogbete JC. Healthcare infrastructure as a public health intervention using evidence from large laboratory networks. Shodhshauryam, International Scientific Refereed Research Journal. 2023;6(1):256-286. doi:10.32628/SHISRRJ23678
58. Ogbete JC, Aminu-Ibrahim AY. Translating healthcare infrastructure investment into measurable population health and diagnostic outcomes. International Journal of Scientific Research in Humanities and Social Sciences. 2024;1(2):955-985. doi:10.32628/IJSRSSH242775
59. Aneke OB, Sydney E, Adesemoye AC. Data integrity and governance in regulated life sciences operations: a review of failures, controls, and remediation strategies. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023;9:1000. doi:10.32628/CSEIT25112194
60. Aneke OB, Adesemoye AC, Sydney E. Corrective and preventive action effectiveness in pharmaceutical quality management systems: A systematic review of root-cause methodologies. International Journal of Health and Pharmaceutical Research. 2022;7(2):78-111. doi:10.56201/ijhpr.vol.7.no2.2022.pg78.111
61. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation. 2013;84(4):465-470. doi:10.1016/j.resuscitation.2012.12.016
62. Liu L, Silva Almodovar A, Nahata MC. Medication adherence in Medicare-enrolled older adults with asthma and chronic obstructive pulmonary disease before and during COVID-19 pandemic. Therapeutic Advances in Chronic Disease. 2023;14:20406223231205796. doi:10.1177/20406223231205796
63. Dong X, Tsang CCS, Zhao S, Browning JA, Wan JY, Chisholm-Burns MA, et al. Effects of the Medicare Part D comprehensive medication review on medication adherence among patients with Alzheimer's disease. Current Medical Research and Opinion. 2021;37(9):1581-1588. doi:10.1080/03007995.2021.1935224
64. Torres J, Carvalho D, Molinos E, Vales C, Ferreira A, Dias CC, et al. The impact of the patient post-intensive care syndrome components upon caregiver burden. Medicina Intensiva. 2017;41(8):454-460. doi:10.1016/j.medin.2016.12.005
65. Lobato CT, Camões J, Carvalho D, Vales C, Dias CC, Gomes E, et al. Risk factors associated with post-intensive care syndrome in family members (PICS-F): a prospective observational study. Journal of the Intensive Care Society. 2023;24(3):247-257. doi:10.1177/17511437221108904
66. Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Critical Care Medicine. 2016;44(2):368-374. doi:10.1097/CCM.0000000000001571
67. Churpek MM, Yuen TC, Winslow C, Robicsek AA, Meltzer DO, Gibbons RD, et al. Multicenter development and validation of a risk stratification tool for ward patients. American Journal of Respiratory and Critical Care Medicine. 2014;190(6):649-655. doi:10.1164/rccm.201406-1022OC

68. Veldhuis LI, Woittiez NJC, Nanayakkara PWB, Ludikhuize J. Artificial intelligence for the prediction of in-hospital clinical deterioration: a systematic review. Critical Care Explorations. 2022;4(9):e0744. doi:10.1097/CCE.0000000000000744
69. Kipnis P, Turk BJ, Wulf DA, LaGuardia JC, Liu V, Churpek MM, et al. Development and validation of an electronic medical record-based alert score for detection of inpatient deterioration outside the ICU. Journal of Biomedical Informatics. 2016;64:10-19. doi:10.1016/j.jbi.2016.09.013
70. Pimentel MAF, Redfern OC, Malycha J, Meredith P, Prytherch D, Briggs J, et al. Detecting deteriorating patients in the hospital: development and validation of a novel scoring system. American Journal of Respiratory and Critical Care Medicine. 2021;204(1):44-52. doi:10.1164/rccm.202007-2700OC
71. Bailey TC, Chen Y, Mao Y, Lu C, Hackmann G, Micek ST, et al. A trial of a real-time alert for clinical deterioration in patients hospitalized on general medical wards. Journal of Hospital Medicine. 2013;8(5):236-242. doi:10.1002/jhm.2009
72. Ekemezie UC, Ohunyon G, Sydney E. Evaluating IDSR implementation in Benue State, Nigeria: DHIS2 data quality and surveillance capacity. Iconic Research and Engineering Journals. 2018;2(6).
73. Nwakamma S, Ojukwu J, Oyesiji SO. LLM-as-a-judge for automated model governance: A review of methods, biases, and release-gating practices. World Journal of Innovation and Modern Technology. 2024;8(6):185-216. doi:10.56201/wjimt.v8.no6.2024.pg185.216
74. Oyesiji SO, Nwakamma S, Ojukwu J. A conceptual framework for context-aware data curation in fine-tuning pipelines. International Journal of Engineering and Modern Technology. 2024;10(11):197-228. doi:10.56201/ijemt.v10.no11.2024.pg197.228
75. Komi NM. When accurate forecasts are not fair forecasts: deep learning load prediction and equity across districts. Shodhshauryam, International Scientific Refereed Research Journal. 2023;6(6):476-546. doi:10.32628/SHISRRJ247102
76. Orise C, Niniola S. Digital equity in healthcare education: A systematic review of access frameworks and interventions for underrepresented learners in clinical training. International Journal of Scientific Research in Humanities and Social Sciences. 2024;1(2):1233-1246.
77. Aminu-Ibrahim AY, Ogbete JC, Ambali KB. Developing sustainable diagnostic laboratory infrastructure models for emerging and resource constrained health systems. Iconic Research and Engineering Journals. 2018;1(8):118-132. doi:10.64388/IREV1I8-1713586
78. Aminu-Ibrahim AY, Ogbete JC, Ambali KB. Infrastructure driven expansion of diagnostic access across underserved and rural healthcare regions. International Journal of Multidisciplinary Research and Growth Evaluation. 2020;1(5):691-706. doi:10.54660/.IJMRGE.2020.1.5.691-706
79. Ogbete JC, Aminu-Ibrahim AY, Ambali KB. Optimizing laboratory spatial planning strategies to improve diagnostic accuracy, safety, and clinical throughput. Iconic Research and Engineering Journals. 2018;2(1):156-177. doi:10.64388/IREV1I7-1713587
80. Ogbete JC, Aminu-Ibrahim AY, Ambali KB. Regulatory compliant design systems for molecular and pathology laboratories in highly controlled environments. Iconic Research and Engineering Journals. 2019;3(4):607-631. doi:10.64388/IREV3I4-1713589

81. Ogbete JC, Aminu-Ibrahim AY. Lifecycle performance evaluation of purpose built diagnostic laboratories supporting long term healthcare delivery. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2605-2621. doi:10.62225/2583049X.2023.3.6.5698
82. Kumuyi O, Akeju B, Uzoka E, Ozowara DE. Framework for blockchain-based cross-border data exchange and regulatory transparency. International Journal of Multidisciplinary Futuristic Development. 2023;4(1):99-113. doi:10.54660/IJMFD.2023.4.1.99-113
83. Kumuyi O, Uzoka E, Akeju B, Ozowara DE. Framework for privacy-focused digital identity verification supporting financial inclusion in Africa. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):3150-3162. doi:10.62225/2583049X.2024.4.6.6005
84. Ozowara DE, Adebayo A, Anunagba CO. A systematic review of cybersecurity investments and their impact on healthcare financial performance. Shodhshauryam, International Scientific Refereed Research Journal. 2022;5(1):404-427.
85. Eze FI, Anene UN, Akinleye OK. Creation of a drug shortage early warning and regulatory response model for essential medicines in global health systems. Research Journal of Pure Science and Technology. 2022;5(1):57-95. doi:10.56201/rjpst.v5.no1.2022.pg57.95
86. Perlroth D, Marrufo G, Montesinos A. Medication therapy management in chronically ill populations: Final report. Acumen LLC & Westat, prepared for the Centers for Medicare & Medicaid Services; 2013. Available from: https://www.cms.gov/priorities/innovation/files/x/mtm-evidencebase.pdf
87. Lee JS, Yang J, Stockl KM, Lew H, Solow BK. Evaluation of eligibility criteria used to identify patients for medication therapy management services: a retrospective cohort study in a Medicare Advantage Part D population. Journal of Managed Care & Specialty Pharmacy. 2016;22(1):22-30. doi:10.18553/jmcp.2016.22.1.22
88. Waghmare PH, Lindsey R, Reed JB, Gao S, Zillich AJ. Systematic review of the impact of medication synchronization on healthcare utilization, economic, clinical, and humanistic outcomes. JACCP: Journal of the American College of Clinical Pharmacy. 2023;6(6):597-614. doi:10.1002/jac5.1815
89. Choudhry NK, Isaac T, Lauffenburger JC, Gopalakrishnan C, Khan NF, Lee M, et al. Rationale and design of the Study of a Tele-pharmacy Intervention for Chronic diseases to Improve Treatment adherence (STIC2IT): a cluster-randomized pragmatic trial. American Heart Journal. 2016;180:90-97. doi:10.1016/j.ahj.2016.07.017
90. Nwakamma S, Ojukwu J, Oyesiji SO. Securing agentic AI enterprise workflows against prompt injection, tool poisoning, memory manipulation, and excessive agency threats. World Journal of Innovation and Modern Technology. 2024;8(6):217-248. doi:10.56201/wjimt.v8.no6.2024.pg217.248
91. Ilodigwe L, Adesemoye AC. Advances in pharmacy-led delivery and access models for strengthening primary healthcare systems and expanding medicine availability in emerging markets. International Journal of Health and Pharmaceutical Research. 2020;5(3):78-96. doi:10.56201/ijhpr.vol.5.no3.2020.pg78.96
92. Omaghomi TT, Akinlolu VS, Fapohunda M. A governance-driven framework for improving workforce planning and nurse-patient ratio optimization in public hospitals.

International Journal of Multidisciplinary Research and Growth Evaluation. 2023;4(6):1357-1368. doi:10.54660/.IJMRGE.2023.4.6.1357-1368

93. Fapohunda M, Akinlolu VS, Omaghomi TT. A conceptual framework for enhancing telehealth adoption among rural and underserved populations in Nigeria and the United States. International Journal of Multidisciplinary Research and Growth Evaluation. 2023;4(6):1348-1356. doi:10.54660/.IJMRGE.2023.4.6.1348-1356
94. Oparah SO, Gado P, Ezeh FE, Gbaraba SV, Adeleke AS. Comprehensive review of telehealth effectiveness in bridging rural-urban disparities in healthcare access. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(4):1535-1546. doi:10.62225/2583049X.2024.4.4.5116
95. Eyetsemitan RA, Oyeleye AO, Ambali KB, Fadayomi O. CRM and workflow automation in small healthcare practices: a process efficiency framework for scalable patient engagement. International Journal of Multidisciplinary Research and Growth Evaluation. 2024;5(6):1931-1949. doi:10.54660/.IJMRGE.2024.5.6.1931-1949
96. Kumuyi O, Uzoka E, Akeju B, Ozowara DE. Architecture for machine learning-enabled predictive energy management using IoT sensor networks. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):3138-3149. doi:10.62225/2583049X.2024.4.6.6004
97. Komi NM, Adeniji IO. A tiered digital twin that brings predictive diagnostics to resource-constrained renewable energy sites. International Journal of Engineering and Modern Technology. 2023;9(3):287-347. doi:10.56201/ijemt.v9.no3.2023.pg287.347
98. Ohunyon G, Ayinde M, Ojukwu PU. The impact of multidisciplinary collaboration on patient outcomes in emergency departments. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2727-2741. doi:10.62225/2583049X.2023.3.6.6029
99. Fapohunda M, Omaghomi TT, Akinlolu VS. A proposed framework for improving postoperative pain management through integrated nursing communication pathways. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):2800-2811. doi:10.62225/2583049X.2024.4.6.5408
100. Fapohunda M, Akinlolu VS, Omaghomi TT, Atima ME. A proposed framework for strengthening patient safety in high-acuity nursing units through integrated clinical policies . 2023.
101. Ohunyon G, Ayinde M, Ojukwu PU. Evidence-based wound care management in acute hospital settings. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):3186-3199. doi:10.62225/2583049X.2024.4.6.6066
102. Anioke SC, Atima ME. Business intelligence applications for mental health resource allocation and public health program accountability. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2549-2563. doi:10.62225/2583049X.2023.3.6.5492
103. Owoade OA, Moneke KC, Anioke SC. Leveraging business intelligence to optimize resource allocation in mental health and substance abuse centers. Journal of Scientific and Engineering Research. 2022;9(12):210-235.

104. Anioke SC, Atima ME. Public health informatics frameworks for protecting vulnerable populations through data driven policy enforcement. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2564-2579. doi:10.62225/2583049X.2023.3.6.5493
105. Anioke SC, Atima ME. Public health governance models using process optimization and performance metrics for regulatory oversight. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2534-2548. doi:10.62225/2583049X.2023.3.6.5491
106. Akinse SG, Ekechi NV, Ohanebo CG. Integrative maternal health model: addressing socioeconomic inequities in rural populations. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2832-2838. doi:10.62225/2583049X.2023.3.6.6177
107. Akinse SG, Ohanebo CG, Ekechi NV. A trauma-informed educational framework for adolescents in underserved communities. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023;9(10):434-443.
108. Andersen LW, Kim WY, Chase M, Berg KM, Mortensen SJ, Moskowitz A, et al. The prevalence and significance of abnormal vital signs prior to in-hospital cardiac arrest. Resuscitation. 2016;98:112-117. doi:10.1016/j.resuscitation.2015.08.016
109. Barwise A, Thongprayoon C, Gajic O, Jensen J, Herasevich V, Pickering BW. Delayed rapid response team activation is associated with increased hospital mortality, morbidity, and length of stay in a tertiary care institution. Critical Care Medicine. 2016;44(1):54-63. doi:10.1097/CCM.0000000000001346
110. Winslow CJ, Edelson DP, Churpek MM, Taneja M, Shah NS, Datta A, et al. The impact of a machine learning early warning score on hospital mortality: a multicenter clinical intervention trial. Critical Care Medicine. 2022;50(9):1339-1347. doi:10.1097/CCM.0000000000005492
111. Adams R, Henry KE, Sridharan A, Soleimani H, Zhan A, Rawat N, et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. Nature Medicine. 2022;28(7):1455-1460. doi:10.1038/s41591-022-01894-0
112. Boussina A, Shashikumar SP, Malhotra A, Owens RL, El-Kareh R, Longhurst CA, et al. Impact of a deep learning sepsis prediction model on quality of care and survival. npj Digital Medicine. 2024;7(1):14. doi:10.1038/s41746-023-00986-6
113. Bassin L, Raubenheimer J, Bell D. The implementation of a real time early warning system using machine learning in an Australian hospital to improve patient outcomes. Resuscitation. 2023;188:109821. doi:10.1016/j.resuscitation.2023.109821
114. van der Vegt AH, Campbell V, Mitchell I, Malycha J, Simpson J, Flenady T, et al. Systematic review and longitudinal analysis of implementing artificial intelligence to predict clinical deterioration in adult hospitals: what is known and what remains uncertain. Journal of the American Medical Informatics Association. 2024;31(2):509-524. doi:10.1093/jamia/ocad220
115. Fapohunda M, Omaghomi TT, Akinlolu VS, Igweonu C, Atima ME. A policy-based model for strengthening antimicrobial stewardship programs in emerging economies. International

Journal of Multidisciplinary Research and Growth Evaluation. 2023;4(6):1369-1378. doi:10.54660/.IJMRGE.2023.4.6.1369-1378

116. Aneke OB, Adesemoye AC, Sydney E. The role of medicine quality failures in the emergence and spread of antimicrobial resistance: A narrative review. International Journal of Medical Evaluation and Physical Report. 2022;6(1):34-62. doi:10.56201/ijmepr.v6.no1.2022.pg34.62
117. Aneke OB, Adesemoye AC. Toward an integrated conceptual framework for traceability and chain-of-custody integrity in pharmaceutical distribution. Iconic Research and Engineering Journals. 2019;3(4):655-672.
118. Aneke OB, Adesemoye AC. Recent advances in analytical and field-deployable methods for detecting substandard and falsified pharmaceuticals. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2024;10(4):1161-1199. doi:10.32628/CSEIT261233059
119. Awulu I, Obukadata VO, Williams I. Value-based pricing and market access strategies for innovative pharmaceuticals: a systematic review. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):3357-3377. doi:10.62225/2583049X.2024.4.6.6836
120. Williams I, Awulu I, Obukadata VO. The impact of biosimilars on pharmaceutical marketing and commercial strategy in global biologics markets. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):3378-3395. doi:10.62225/2583049X.2024.4.6.6837
121. Asiedu CS, Asiedu AA. Reconceptualizing quality assurance in pharmaceutical manufacturing as a determinant of supply reliability. International Journal of Health and Pharmaceutical Research. 2024;9(5):148-173. doi:10.56201/ijhpr.v9.no5.2024.pg148.173
122. Aneke OB, Adesemoye AC. A conceptual framework for risk-based decision-making in good manufacturing practice environments. Iconic Research and Engineering Journals. 2020;4(6):325-341.
123. Aneke OB, Adesemoye AC. Reframing trade and import compliance: a conceptual model for risk-based interdiction of illicit medical products. International Journal of Scientific Research in Science and Technology. 2021;8(6):783. doi:10.32628/IJSRST1229274
124. Aneke OB, Adesemoye AC, Sydney E. Regulatory harmonization and convergence in global medicine quality assurance: A review of frameworks, pathways, and implementation barriers. International Journal of Medical Evaluation and Physical Report. 2024;8(6):213-252. doi:10.56201/ijmepr.v8.no6.2024.pg213.252
125. Akin-Oluyomi OT, Atima ME, Akinleye OK. Regulatory compliance and supplier risk assessment frameworks in international pharmaceutical procurement. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2194-2204. doi:10.62225/2583049X.2023.3.6.5326
126. Browning JA, Tsang CCS, Dong X, Wan JY, Chisholm-Burns MA, Finch CK, et al. Effects of Medicare comprehensive medication review on racial/ethnic disparities in nonadherence to statin medications among patients with Alzheimer's disease: an observational analysis. BMC Health Services Research. 2022;22(1):159. doi:10.1186/s12913-022-07483-8

127. Aminu-Ibrahim AY, Ogbete JC, Ambali KB. Capital project delivery models for high risk healthcare infrastructure in developing national health systems. Iconic Research and Engineering Journals. 2019;2(10):626-649. doi:10.64388/IREV2I10-1713588
128. Aminu-Ibrahim AY, Ogbete JC, Ambali KB. Program management models for coordinated multi site healthcare infrastructure expansion projects. International Journal of Multidisciplinary Research and Growth Evaluation. 2021;2(6):661-678. doi:10.54660/.IJMRGE.2021.2.6.661-678
129. Aminu-Ibrahim A, Ogbete JC, Ambali KB. Design and development of high-performance medical laboratories supporting pandemic preparedness and response. Shodhshauryam, International Scientific Refereed Research Journal. 2022;5(3):302-335. doi:10.32628/SHISRRJ225486
130. Aminu-Ibrahim AY, Ogbete JC. Governance and accountability models for public private partnerships in healthcare infrastructure development. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):2943-2960. doi:10.62225/2583049X.2024.4.6.5699
131. Ogbete JC, Aminu-Ibrahim AY, Ambali KB. Risk managed construction strategies for complex and highly regulated healthcare facility developments. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2021;7(5):362-392. doi:10.32628/CSEIT217561
132. Ogbete JC, Aminu-Ibrahim AY, Ambali KB. Sustainable materials selection and energy efficiency strategies for modern medical laboratory facilities. International Journal of Multidisciplinary Research and Growth Evaluation. 2020;1(5):674-690. doi:10.54660/.IJMRGE.2020.1.5.674-690
133. Ogbete JC, Aminu-Ibrahim AY, Ambali KB. Design standards and operational planning frameworks for scalable blood collection networks. Gyanshauryam, International Scientific Refereed Research Journal. 2022;5(6):267-300. doi:10.32628/GISRRJ225635
134. Amayo EB, Owulade OA, Isi LR. Risk management strategies and compliance frameworks in healthcare infrastructure projects: case studies from West Africa. International Journal of Management and Organizational Research. 2023;2(1):161-168. doi:10.54660/IJMOR.2023.2.1.161-168
135. Akinlolu VS, Omaghomi TT, Fapohunda M. A data-driven framework for improving health facility preparedness and disaster risk mitigation in urban hospitals. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):2757-2769. doi:10.62225/2583049X.2024.4.6.5404
136. Afrihyia E, Ojukwu PU, Akinse SG. Privacy-preserving health data governance models: A comparative review of blockchain and cryptographic strategies in U.S. and developing healthcare systems. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):3071-3086. doi:10.62225/2583049X.2024.4.6.5919
137. Mbonu IS, Aliliele C, Iwuanyanwu U, Oluoha OM. A conceptual framework for legal and ethical risk modeling in enterprise data protection governance systems. Iconic Research and Engineering Journals. 2018;2(2):207-226. doi:10.64388/IREV2I2-1714911

138. Adebayo A, Adepoju PA, Ozowara DE. Conceptual model for cyber risk pricing and insurance structuring in health technology platforms. Gyanshauryam, International Scientific Refereed Research Journal. 2023;6(6).
139. Ekechi NV, Anunagba CO, Ozowara DE. Advances in financial forensics techniques for detecting cyber-enabled fraud in telemedicine services. Shodhshauryam, International Scientific Refereed Research Journal. 2022;5(1):428-450.
140. Nnaji N, Akinlolu VS. Enhancing clinical record protection through structured health data security frameworks in regulated healthcare environments. International Journal of Scientific Research in Humanities and Social Sciences. 2024. Available from: https://ijsrhss.com/home/article/view/IJSRSSH243676
141. Aliliele C, Mbonu IS, Iwuanyanwu U. A conceptual framework for continuous cloud misconfiguration monitoring and enterprise risk mitigation strategies. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2023;9(10):373-394. doi:10.32628/CSEIT2361071
142. Onche VO, Adegbite MP, Ogbonna CS. BYOD adoption and information-security risk in university administrative offices: evidence from Nigerian federal universities. International Journal of Multidisciplinary Research and Growth Evaluation. 2023;4(6):1549-1563. doi:10.54660/.IJMRGE.2023.4.6.1549-1563
143. Akomolafe O, Agu MU, Bello A. A quantitative conceptual model for assessing compliance risk in emerging economies. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2287-2296. doi:10.62225/2583049X.2023.3.6.5360
144. Falemi A, Akhigbe R, Akin-Oluyomi OT. A supply chain workforce optimization framework for addressing staffing volatility during rapid expansion cycles. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(1):1745-1762. doi:10.62225/2583049X.2023.3.1.5342
145. Falemi A, Akhigbe R, Akin-Oluyomi OT. A conceptual enterprise risk management model that integrates HR strategy with operational analytics. International Journal of Multidisciplinary Research and Growth Evaluation. 2024;5(6):1730-1753. doi:10.54660/.IJMRGE.2024.5.6.1730-1753
146. Falemi A, Akhigbe R, Akin-Oluyomi OT. A conceptual supply chain talent development model for capability building across distributed operations. International Journal of Multidisciplinary Research and Growth Evaluation. 2020;1(5):439-456. doi:10.54660/.ijmrge.2020.1.5.439-456
147. Obriki OD, Obogo SF, Arumosoye OM. Advances in workforce safety training models for operational risk reduction. Gyanshauryam, International Scientific Refereed Research Journal. 2022;5(2):349-382. doi:10.32628/GISRRJ225638
148. Ilodigwe L, Adesemoye AC. Beyond technology: A strategic framework for building sustainable and resilient health systems through organizational strategy, operational excellence, and workforce capability. International Journal of Medical Evaluation and Physical Report. 2024;8(6):299-319. doi:10.56201/ijmepr.v8.no6.2024.pg299.319
149. McIsaac DI, Talarico R, Jerath A, Wijeysundera DN. Days alive and at home after hip fracture: a cross-sectional validation of a patient-centred outcome measure using routinely

collected data. BMJ Quality and Safety. 2023;32(9):546-556. doi:10.1136/bmjqs-2021-013150

150. Taran S, Coiffard B, Huszti E, Li Q, Chu L, Thomas C, et al. Association of days alive and at home at day 90 after intensive care unit admission with long-term survival and functional status among mechanically ventilated patients. JAMA Network Open. 2023;6(3):e233265. doi:10.1001/jamanetworkopen.2023.3265
151. Kahn JM, Le T, Angus DC, Cox CE, Hough CL, White DB, et al. The epidemiology of chronic critical illness in the United States. Critical Care Medicine. 2015;43(2):282-287. doi:10.1097/CCM.0000000000000710
152. Needham DM, Davidson J, Cohen H, Hopkins RO, Weinert C, Wunsch H, et al. Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. Critical Care Medicine. 2012;40(2):502-509. doi:10.1097/CCM.0b013e318232da75
153. Mikkelsen ME, Still M, Anderson BJ, Bienvenu OJ, Brodsky MB, Brummel N, et al. Society of Critical Care Medicine's international consensus conference on prediction and identification of long-term impairments after critical illness. Critical Care Medicine. 2020;48(11):1670-1679. doi:10.1097/CCM.0000000000004586
154. Walsh TS, Salisbury LG, Merriweather JL, Boyd JA, Griffith DM, Huby G, et al. Increased hospital-based physical rehabilitation and information provision after intensive care unit discharge: the RECOVER randomized clinical trial. JAMA Internal Medicine. 2015;175(6):901-910. doi:10.1001/jamainternmed.2015.0822
155. Taito S, Yamauchi K, Tsujimoto Y, Banno M, Tsujimoto H, Kataoka Y. Does enhanced physical rehabilitation following intensive care unit discharge improve outcomes in patients who received mechanical ventilation? A systematic review and meta-analysis. BMJ Open. 2019;9(6):e026075. doi:10.1136/bmjopen-2018-026075
156. O'Neill B, Green N, Blackwood B, McAuley D, Moran F, MacCormac N, et al. Recovery following discharge from intensive care: what do patients think is helpful and what services are missing?. PLOS ONE. 2024;19(3):e0297012. doi:10.1371/journal.pone.0297012
157. Ely EW. The ABCDEF bundle: science and philosophy of how ICU liberation serves patients and families. Critical Care Medicine. 2017;45(2):321-330. doi:10.1097/CCM.0000000000002175
158. Kosilek RP, Schmidt K, Baumeister SE, Gensichen J, SMOOTH Study Group. Frequency and risk factors of post-intensive care syndrome components in a multicenter randomized controlled trial of German sepsis survivors. Journal of Critical Care. 2021;65:268-273. doi:10.1016/j.jcrc.2021.07.006
159. Wang Y, et al.. Follow-up focused on psychological intervention initiated after intensive care unit in adult patients and informal caregivers: A systematic review and meta-analysis. 2023. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10259442/
160. Oparah OS, Ezeh FE, Olatunji GI, Ajayi OO. Framework for designing national real-time disease surveillance dashboards for public health stakeholders. Shodhshauryam, International Scientific Refereed Research Journal. 2023;6(1):208-227.
161. Ogunboye I, Adebayo IPS, Anioke SC, Egwuatu EC, Ajala CF, Awuah SB. Enhancing Nigeria's health surveillance system: a data-driven approach to epidemic preparedness and

response. World Journal of Advanced Research and Reviews. 2023;20(1):1352-1369. doi:10.30574/wjarr.2023.20.1.2078

162. Olorunkosebi MT, Ayinde MK, Mensah GK, Penddah AL. Integrating artificial intelligence and advanced biocensors for rapid and sensitive detection of antimicrobial-resistant foodborne pathogens. Journal of Advances in Microbiology Research. 2023;4(2):175-182. doi:10.22271/micro.2023.v4.i2b.293
163. Ayinde MK, Penddah AL, Mensah GK, Olorunkosebi MT. Environmental microplastics as vectors for avian virus transmission and their implications for food safety within a one health framework. Journal of Current Research in Food Science. 2024;5(2):210-219. doi:10.22271/foodsci.2024.v5.i2c.288
164. Ejofodomi OA, Gideon EN, Oladipo GO, Oshomah ER. Automated detection of architectural detection in mammograms using template matching. International Journal of Biomedical Science and Engineering. 2014;2(1):1-6. doi:10.11648/j.ijbse.20140201.11
165. Asiedu CS. Diagnostic integration in infectious disease management: A review of microbiology, virology, and serology workflows. International Journal of Medical Evaluation and Physical Report. 2023;7(4):173-195. doi:10.56201/ijmepr.v7.no4.2023.pg173.195
166. Asiedu CS. Droplet digital PCR for HIV viral reservoir quantification: A review of diagnostic accuracy and clinical utility. International Journal of Medical Evaluation and Physical Report. 2024;8(6):253-273. doi:10.56201/ijmepr.v8.no6.2024.pg253.273
167. Orise C, Niniola S. Privacy, safeguarding, and ethical data practices in healthcare EdTech: A conceptual model for compliance and trust. Shodhshauryam, International Scientific Refereed Research Journal. 2023;6(1):507-529.
168. Orise C, Niniola S. Real-time analytics for educational excellence: A conceptual framework for performance dashboards in health professional training. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022;8(2):802-815.
169. Orise C, Niniola S. Data-informed digital learning platforms: a conceptual framework for evidence-based educational technology in healthcare and STEM environments. International Journal of Multidisciplinary Research and Growth Evaluation. 2020;1(5):1078-1089. doi:10.54660/.IJMRGE.2020.1.5.1078-1089
170. Adams V. Preserving critical thinking in the era of generative artificial intelligence: Evidence-based frameworks for human-centered learning. International Journal of Social Sciences and Management Research. 2023;9(11):321-332. doi:10.56201/ijssmr.v9.no11.2023.pg321.332
171. Adams V. Assessment integrity in the age of artificial intelligence: Redesigning learning evaluation for authentic human competency. World Journal of Innovation and Modern Technology. 2024;8(6):185-195. doi:10.56201/wjimt.v8.no6.2024.pg185.195
172. Eze FI, Anene UN, Akinleye OK. Design of an integrated regulatory readiness maturity model for pharmaceutical companies seeking global market entry. International Journal of Health and Pharmaceutical Research. 2022;7(1):77-112. doi:10.56201/ijhpr.v7.no1.2022.pg77.112

173. Eze FI, Akinleye OK, Anene UN. Development of a risk-based accelerated approval decision model balancing speed, safety, and evidence quality in pharmaceutical regulatory science. Research Journal of Pure Science and Technology. 2024;7(6):120-161. doi:10.56201/rjpst.v7.no6.2024.pg120.161
174. Eze FI, Akinleye OK, Anene UN. Development of a cross-border regulatory harmonization model for African pharmaceutical markets: The ARCH-Model framework. International Journal of Health and Pharmaceutical Research. 2024;9(5):148-186. doi:10.56201/ijhpr.v9.no5.2024.pg148.186
175. Eze FI, Anene UN, Akinleye OK. Pharmaceutical supply chain governance, environmental risk management, and sustainable access to essential medicines in emerging economies: The PSCG-ERM framework. International Journal of Health and Pharmaceutical Research. 2023;8(4):192-241. doi:10.56201/ijhpr.v8.no4.2023.pg192.241
176. Akinlolu VS, Omaghomi TT, Fapohunda M, Atima ME. A systems-level policy framework for integrating mental health screening into primary healthcare in low-resource settings. International Journal of Multidisciplinary Research and Growth Evaluation. 2022;3(6):879-887. doi:10.54660/.IJMRGE.2022.3.6.879-887
177. Nnaji N, Akinlolu VS. Strengthening medical information flow through digital operations models for modern mental health care. Gyanshauryam, International Scientific Refereed Research Journal. 2024;7.
178. Oparah SO, Gado P, Ezeh FE, Gbaraba SV, Adeleke AS. Assessing the operational and psychosocial impact of the compressed workweek: a meta-analytic review of four-day work week trials across industries. Global Multidisciplinary Perspectives Journal. 2024;1(6):51-60. Available from: https://www.multiperspectivesjournal.com/uploads/archives/20251023114701_GMP-2025-1-024.1.pdf
179. Nnaji N, Akinlolu VS. Reducing decision delays through real-time health informatics and performance analytics. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 2022.
180. Nnaji N, Akinlolu VS. Addressing productivity and workflow inefficiencies using analytics-driven performance models in healthcare operations. Shodhshauryam, International Scientific Refereed Research Journal. 2023.
181. Asiedu CS, Asiedu AA. Comparative bioavailability of fresh versus dried botanical compounds: A review and cost-effectiveness modelling framework. Research Journal of Pure Science and Technology. 2023;6(3):226-244. doi:10.56201/rjpst.v6.no3.2023.pg226.244
182. Gbadamosi IT, Obogo SF. Chemical constituents and in vitro antimicrobial activities of five botanicals used traditionally for the treatment of neonatal jaundice in Ibadan, Nigeria. Nature and Science. 2013;11(10):130-135.
183. Omaghomi TT, Fapohunda M, Akinlolu VS, Atima ME. A conceptual framework for improving fall prevention outcomes through policy-integrated nursing workflows. International Journal of Advanced Multidisciplinary Research and Studies. 2024;4(6):2770-2778. doi:10.62225/2583049X.2024.4.6.5405

184. Omaghomi TT, Fapohunda M, Akinlolu VS. Multidimensional framework for strengthening infection prevention and control (IPC) compliance in tertiary hospitals. International Journal of Multidisciplinary Research and Growth Evaluation. 2022;3(6):888-896. doi:10.54660/.IJMRGE.2022.3.6.888-896

185. Oparah SO, Gado P, Ezeh FE, Gbaraba SV, Omotayo O, Adeleke AS. Framework for scaling mobile health solutions for chronic disease monitoring and treatment adherence improvement. International Journal of Multidisciplinary Research and Growth Evaluation. 2021;2(4):1074-1094. doi:10.54660/.IJMRGE.2021.2.4.1074-1094

186. Ilodigwe L, Adesemoye AC. A critical review of health insurance financing mechanisms and provider payment reform strategies for balancing coverage expansion and financial protection in emerging markets. International Journal of Health and Pharmaceutical Research. 2019;5(2):31-55. doi:10.56201/ijhpr.vol.5.no2.2019.pg31.55

187. Afrihyia E, Akinse SG, Ojukwu PU. Development and implementation of an AI-driven early detection system for non-communicable diseases. International Journal of Advanced Multidisciplinary Research and Studies. 2023;3(6):2680-2691. doi:10.62225/2583049X.2023.3.6.5977