A PRESCRIPTIVE ANALYTICS FRAMEWORK FOR RISK-INTEGRATED MATERNAL HEALTHCARE RESOURCE ALLOCATION IN ZIMBABWE

Ruramai Judith Yotamu

MTech Data Science and Analytics, School of Information Sciences and Technology, Harare Institute of Technology, Zimbabwe

ABSTRACT

This paper presents a Prescriptive Analytics Framework for Maternal Health (PAMH) for risk-integrated maternal healthcare resource allocation in Zimbabwe. The framework combines machine-learning risk prediction with a mixed-integer linear programming optimisation model to allocate midwives, delivery kits and ambulances across 84 facilities under six budget scenarios. Logistic Regression, Random Forest and Gradient Boosting models were trained on 5,001 patient encounters, with Gradient Boosting achieving the strongest predictive performance (test ROC-AUC 0.913). Facility-level risk scores were embedded as priority weights in the optimisation objective, enabling risk-sensitive allocation under budget and equity constraints. Baseline optimisation achieved 97.9% budget utilisation, while the austerity scenario showed a 157% rise in weighted unmet demand. A six-page decision support dashboard translates the framework into actionable intelligence for district health officers.

KEYWORDS

Prescriptive analytics, MILP, Maternal health, Machine learning, Resource allocation

1. INTRODUCTION

Zimbabwe's Maternal Mortality Ratio (MMR) stood at approximately 462 per 100,000 live births in the most recent national estimates (WHO, 2023; ZimStat, 2022), placing the country well above the Sustainable Development Goal (SDG) 3.1 target of fewer than 70 maternal deaths per 100,000 live births by 2030 (WHO, 2023). Globally, an estimated 287,000 women die annually from preventable pregnancy-related causes, with Sub-Saharan Africa accounting for more than 66% of those deaths (WHO, 2023; UNFPA, 2023). Zimbabwe's persistent burden reflects structural rather than purely clinical failure: rural clinics and district hospitals in Manicaland, Mashonaland West, and Midlands provinces consistently operate below minimum staffing and equipment thresholds, while certain urban-adjacent facilities register surpluses against their demand requirements (MoHCC, 2021; ZimStat, 2022). This allocation inefficiency is compounded by the Three Delays framework delays in deciding to seek care, in reaching a facility, and in receiving adequate treatment each of which is independently fatal when the institutional resource environment is poorly matched to local need (Thaddeus & Maine, 1994).

Current resource allocation in Zimbabwe's Ministry of Health and Child Care (MoHCC) relies on historical headcount-based formulae and district administrative discretion (MoHCC, 2023; Haddara & Elragal, 2015). Neither approach incorporates probabilistic foresight on facility-specific complication risk, seasonal access burden, or demand fluctuation. Enterprise Resource Planning (ERP) systems deployed in the sector record completed transactions retrospectively but

DOI:10.5121/ijaia.2026.17304

offer no prospective allocation intelligence (Haddara & Elragal, 2015). The resulting decision intelligence gap the inability to proactively match resource deployment to risk-weighted facility demand before commitment drives the persistent maternal mortality disparity documented across Zimbabwe's provincial health data (MoHCC, 2022; ZimStat, 2022).

Operations research and data science literature demonstrate that Mixed-Integer Linear Programming (MILP) models can produce near-optimal facility-to-resource assignments under budget and equity constraints (Crown et al., 2018; Leung et al., 2023), and that ensemble machine learning classifiers can generate reliable risk probabilities from structured health records (Grinsztajn et al., 2022; Obsie et al., 2020). However, these two capabilities have not previously been combined within a single, integrated, health-system-adapted decision-intelligence artefact for maternal care in Southern Africa. This paper addresses that gap directly.

The Prescriptive Analytics Framework for Maternal Health (PAMH) integrates three components: (1) a risk prediction module employing Logistic Regression, Random Forest, and Gradient Boosting classifiers trained on 5,001 patient records to estimate facility-level complication probability; (2) a MILP optimisation engine that embeds these probabilities as priority weights in a demand-minimisation objective, allocating midwives, delivery kits, and ambulances across 84 facilities in 21 districts under six budget scenarios; and (3) a six-page interactive DSS dashboard that delivers all outputs to district health officers through a browser-accessible interface requiring no statistical expertise.

The specific contributions of this paper are: (i) the first risk-integrated MILP formulation for maternal health resource allocation in Sub-Saharan Africa; (ii) a documented prescriptive analytics pipeline from raw health records to actionable allocation plans; (iii) a quantified equity-efficiency frontier for six Zimbabwean health budget scenarios; and (iv) an open-architecture DSS designed for deployment within Zimbabwe's existing DHIS2-compatible health information infrastructure. To the best of the author's knowledge, no prior study has embedded ML-derived facility risk scores as continuous objective weights within a MILP formulation for maternal healthcare resource allocation in Sub-Saharan Africa. PAMH addresses this gap directly, moving beyond prediction to prescription within a single integrated artefact.

The remainder of this paper is structured as follows. Section 2 reviews the related literature across three domains. Section 3 describes the PAMH framework architecture and mathematical formulation. Section 4 presents the dataset and experimental setup. Section 5 reports results across predictive, optimisation, and comparison dimensions. Section 6 discusses theoretical contributions, deployment pathway, and limitations. Section 7 concludes with policy implications and future work directions.

2. Related Work

2.1. Maternal Health Resource Allocation in LMICs

Resource allocation in low- and middle-income country (LMIC) maternal health systems has been studied through both equity and efficiency lenses. Ozawa et al. (2019) established that LMIC health resource allocation frameworks systematically under-account for facility-level demand heterogeneity, producing avoidable mismatches between supply and need; their analysis found that population-proportional formulae fail to reflect within-district variation in complication risk, leading to over-allocation to low-risk urban facilities at the expense of high-risk rural ones. Leung et al. (2023) demonstrated through constrained optimisation modelling that realistic capacity constraints and equity floors could simultaneously be satisfied in LMIC health settings when the allocation problem was formulated mathematically rather than handled through administrative discretion, achieving a 22% improvement in equity-adjusted service coverage over the administrative baseline. Boutilier and Chan (2020) applied integer programming to ambulance deployment in LMICs and found that optimised positioning reduced mean emergency response times by 18–24%, with the largest gains concentrated in underserved rural zones.

Critically, however, these studies share three methodological limitations that PAMH directly addresses. First, they are predominantly single-resource in scope allocating either staff or equipment but not both simultaneously meaning that complementarities between midwife availability and ambulance access are never modelled jointly (Crown et al., 2018). Second, allocation priorities are set using historical mortality statistics or population headcounts rather than forward-looking probabilistic risk estimates, making them blind to facilities where predicted risk diverges from past mortality (Ozawa et al., 2019). Third, none of the reviewed studies produce an interactive decision support interface deployable without specialist statistical expertise, limiting their practical uptake in district-level planning contexts (Leung et al., 2023). PAMH addresses all three gaps through joint multi-resource optimisation, ML-embedded risk weighting, and a browser-based DSS.

2.2. Machine Learning for Maternal Complication Prediction

Ensemble learning has emerged as the dominant paradigm for obstetric risk classification in structured clinical datasets. Obsie et al. (2020) developed a predictive model for severe acute maternal morbidity in low-resource settings using Random Forest and Gradient Boosting, achieving AUC values of 0.81–0.87 and demonstrating that ensemble classifiers significantly outperform logistic regression on imbalanced complication datasets. Grinsztajn et al. (2022) provided theoretical and empirical grounding across 45 tabular benchmarks, showing that tree-based ensembles consistently outperform deep learning architectures on structured data the class of data encountered in health management information systems such as DHIS2 and OpenLMIS. Friedman (2001) established the residual-fitting mechanism underpinning Gradient Boosting's superiority on heterogeneous feature spaces with non-linear interaction effects.

A persistent and critical limitation of published maternal health ML studies is that they operate at the individual patient level and stop at prediction: they produce a complication probability for a given patient encounter but do not aggregate those probabilities into facility-level operational signals, and they do not connect those signals to resource allocation decisions (Ngoma et al., 2022; Obsie et al., 2020). This means a clinician receives a risk score for an individual patient but a district health officer receives no guidance on where to deploy additional midwives or ambulances. This prediction-to-prescription translation gap is the core architectural contribution of PAMH: Gradient Boosting outputs are aggregated into facility-level risk weights that are embedded directly in the MILP objective function, so the optimisation engine is guided by predicted risk rather than past mortality statistics.

2.3. Decision Support Systems for Public Health

Holsapple et al. (2014) established empirically that DSS adoption produces measurable organisational performance improvements when decision-makers interact directly with model outputs in real time rather than receiving static analytical reports. Kogler and Rauch (2022) documented significant efficiency gains in operations management contexts where optimisation-linked DSS platforms replaced manual planning, providing a transferable design precedent for healthcare operations. Hevner et al. (2004) formalised the Design Science Research (DSR) paradigm, establishing the theoretical basis for evaluating information systems artefacts on their utility, rigour, and design coherence the evaluation framework applied in this paper.

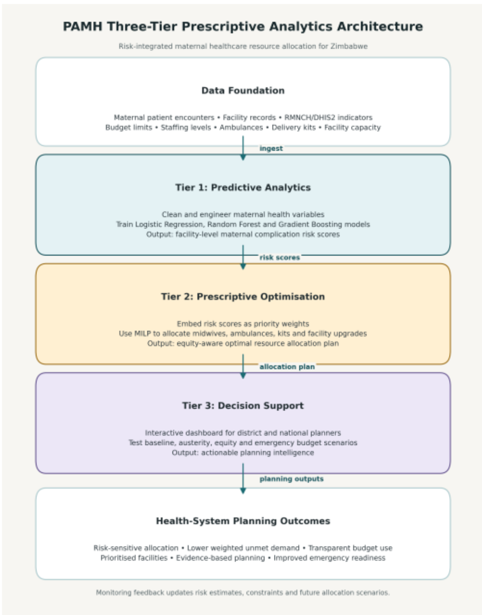
Existing public health DSS deployments in Sub-Saharan Africa, including those built on the DHIS2 platform, provide retrospective dashboards that visualise historical data but do not generate forward-looking allocation recommendations (MoHCC, 2023). They report what happened MMR last year, ANC attendance last quarter but do not advise health officers on where to deploy the next available midwife or ambulance given a constrained budget. This gap between descriptive reporting and prescriptive intelligence is precisely what PAMH fills, and explains why a technically novel DSS artefact is warranted rather than an extension of existing reporting platforms.

3. FRAMEWORK ARCHITECTURE AND PROBLEM FORMULATION

3.1. Conceptual Architecture

PAMH operationalises a three-tier analytics pipeline. The Data Tier ingests pre-allocation operational variables from five structured dataset sheets spanning patient-level encounters, facility metadata, facility-year summaries, resource allocation parameters, and province-year aggregate statistics. The Predictive Risk Tier applies trained machine learning classifiers to produce calibrated facility-level complication probability estimates, which are normalised and embedded as priority weights. The Prescriptive Optimisation Tier executes the MILP model, allocating midwives, delivery kits, and ambulances under budget and equity constraints, with results rendered through an interactive DSS dashboard.

Figure 1. PAMH three-tier prescriptive analytics architecture.



The architectural distinction from prior work is the direct coupling between the predictive and optimisation tiers. Rather than using risk as a binary filter or post-optimisation diagnostic, PAMH embeds risk as a multiplicative weight on demand urgency within the MILP objective function, ensuring the solver actively prioritises allocations that reduce unmet demand at high-risk facilities.

3.2. Predictive Risk Model Formulation

Let each patient be represented by a 16-dimensional pre-admission feature vector aligned to the Three Delays framework. Delay 1 features include ANC visits, gestational age at first visit, and urgency score. Delay 2 features include distance to facility, referral delay hours, transport access, and ambulance availability. Delay 3 features include skilled birth attendance, haemoglobin level (g/dL), facility beds available, and power outage days per month. Clinical risk indicators HIV status, hypertension, diabetes, body mass index, and history of postpartum haemorrhage span all three delay domains.

A Gradient Boosting classifier is trained to estimate the probability of severe maternal complication p_ij for each patient i at facility j. Facility-level risk weights are computed as the mean predicted probability across all patients served by each facility:

$$R_j = \frac{1}{n_j} \sum_{i=1}^{n_j} p_{ij}$$

(Equation 1)

where n_j is the number of patient encounters at facility j. These weights are normalised to the unit interval [0,1] before entry into the MILP objective.

3.3. MILP Optimisation Formulation

Let $J = \{1, \dots, 84\}$ be the set of health facilities and $K = \{midwives, kits, ambulances\}$ the set of resource types. Decision variables: x_{kj}(number of resource k allocated to facility j, integer ≥ 0); u_j (binary upgrade indicator); s_j(non-negative slack variable for unmet demand, continuous ≥ 0).

The composite demand score for facility j is:

$$R_j = \frac{1}{n_j} \sum_{i=1}^{n_j} p_{ij} \tag{9}$$

$$D_j = 0.30 \left(\frac{\mathrm{Vol}_j}{\max \mathrm{Vol}} \right) + 0.25 \left(\frac{\mathrm{CR}_j}{\max \mathrm{CR}} \right) + 0.30 \left(R_j \right) + 0.15 \left(1 - \frac{\mathrm{SBA}_j}{\max \mathrm{SBA}} \right) \tag{10}$$

(Equation 2)

where vol_j= annualised delivery volume, CR_J = complication rate, R_J = ML risk weight from Equation 1, and SB_J= skilled birth attendance rate. Weights reflect the relative primacy of delivery volume and ML risk consistent with the Three Delays evidence base (Thaddeus & Maine, 1994).

The objective function minimises total risk-weighted unmet demand plus an access burden penalty (α = 0.5):

$$\min \sum_{j \in J} R_j \cdot s_j + 0.5 \sum_{j \in J} A_j \cdot s_j$$

(Equation 3)

Where A_J is the access burden index (normalised mean distance to facility weighted by delivery volume).

Resource-class penalty coefficients P_k differentiate the cost of under-allocation across resource types.

Table 1 reports the explicit penalty structure:

Resource Class	Unit Cost (USD)	Penalty Weight P_k	Justification
Midwives	7,200 / yr	1.00 (base)	Primary skilled care provider; baseline reference
Delivery Kits	150 / kit	0.25	Lower unit cost; scaled proportionally to cost ratio
Ambulances	25,000 / yr	3.47	Highest mortality impact per Boutilier & Chan (2020); scaled to unit cost ratio

Table 1. Resource-class penalty coefficients P_k in the MILP objective function. An ambulance shortfall at a high-risk facility is penalised 3.47× more heavily than a midwife shortfall and 13.9× more heavily than a delivery kit shortfall.

Constraints:

1. **Budget constraint:**

$$\sum_{j \in J} \sum_{k \in K} c_k x_{kj} + \sum_{j \in J} u_j \cdot 50{,}000 \leq B \tag{1}$$

2. **Minimum staffing:**

$$x_{kj} \geq m_{kj}^{\min} \qquad \forall\, j \in J,\; k \in K \tag{2}$$

3. **Resource pool limit:**

$$\sum_{j \in J} x_{kj} \leq P_k \qquad \forall\, k \in K \tag{3}$$

4. **Unmet demand:**

$$s_j \geq D_j - \sum_{k \in K} x_{kj} \qquad \forall\, j \in J \tag{4}$$

5. **Upgrade logic:**

$$u_j \leq \frac{x_{\text{midwives},j}}{m^{\text{upgrade}}} \qquad \forall\, j \in J \tag{5}$$

6. **Provincial equity floor** (Equity-Focused scenario only):

$$\sum_{j \in \mathcal{P}_p} x_{\text{midwives},j} \;\geq\; 0.60 \times \frac{\text{pop}_p}{\text{pop}_{\text{total}}} \times \sum_{j \in J} x_{\text{midwives},j} \qquad \forall\, p \tag{6}$$

7. **Non-negativity and integrality:**

$$x_{kj} \in \mathbb{Z}^+, \qquad u_j \in \{0,1\}, \qquad s_j \geq 0 \qquad \forall\, j \in J,\; k \in K \tag{7}$$

To guarantee solver convergence under all six budget scenarios including the Austerity scenario (USD 1.5M) where strict minimum staffing constraints risk infeasibility the model uses non-negative slack variables S_j(Equation 7). These elastic slack variables absorb constraint violations as quantified unmet demand rather than causing solver failure, converting hard demand constraints into soft penalties. This ensures the model always returns a feasible allocation plan, with the penalty magnitude signalling the severity of resource shortfall.

3.4. Operational Time Step and Risk Score Currency

The PAMH pipeline operates on an annual deployment cycle, aligned with MoHCC's fiscal planning calendar. At the start of each budget year, the Gradient Boosting model is re-trained on the most recent 12-month patient encounter data extracted from DHIS2, producing updated facility-level risk weights R_j that reflect current complication rates, staffing levels, and access conditions. The MILP model is then re-solved against these updated weights to generate the recommended allocation plan for the forthcoming year. Risk weights therefore have a maximum age of 12 months at the point of allocation, ensuring they remain operationally current within the annual planning window.

In settings where sub-annual seasonal outbreaks for example, malaria-associated anaemia spikes during the rainy season materially alter facility risk profiles within the year, quarterly re-scoring is recommended. This requires quarterly data extraction from DHIS2, which is feasible under Zimbabwe's HMIS infrastructure but not currently part of standard district reporting workflows (MoHCC, 2023). The annual cycle represents the operationally achievable baseline; quarterly re-scoring is identified as a near-term enhancement priority for deployment in high-seasonality provinces.

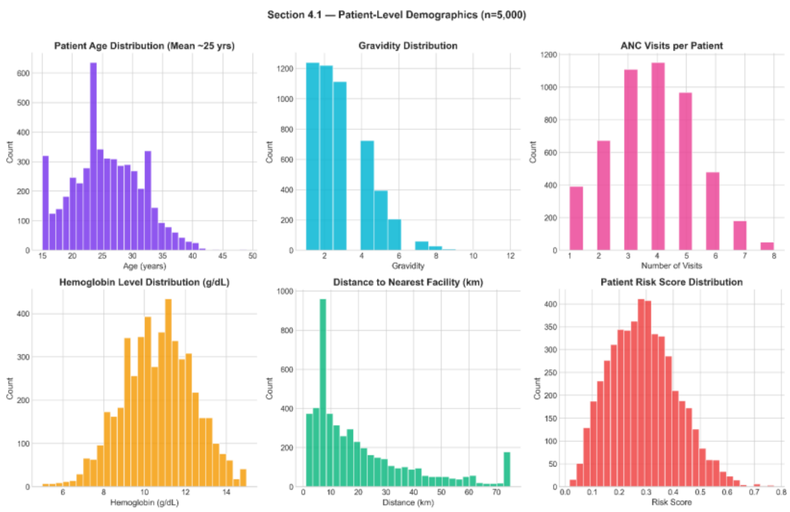
4. Dataset and Experimental Setup

4.1. Dataset Description

The study employs a multi-sheet longitudinal maternal health dataset covering 84 health facilities across three Zimbabwean provinces Manicaland, Mashonaland West, and Midlands over the period 2018 to 2023. The Patient-Level sheet contains 5,001 records with 26 variables per record, including clinical indicators, service access measures, and binary outcome flags for severe maternal complication and maternal mortality. The Facility Lookup sheet covers 84 facilities across 21 districts with structural metadata including GPS coordinates, staffing levels, bed counts, and infrastructure readiness indicators. The Facility-Year Summary sheet provides 497 facility-year aggregate observations. The Resource Allocation sheet records current versus required staffing and ambulance levels per facility. The Province-Year Summary sheet provides 18 province-year aggregate rows used for temporal trend analysis.

Across the full patient cohort, the severe maternal complication rate was 28.6% and the maternal mortality rate was 3.90%. HIV-positive prevalence was 24.3%, consistent with UNAIDS (2023) national estimates. Facility-level resource statistics showed a mean midwife gap of −1.8 per facility and a mean ambulance gap of −0.6; a subset of rural clinics recorded positive gaps indicating genuine shortfalls.

Figure 2. Dataset summary and maternal health data structure.



4.2. Data Preparation and Feature Engineering

Missing value imputation applied column-wise median substitution for continuous features and mode substitution for binary clinical indicators. Outlier detection using the interquartile range method identified extreme values in age (9 cases, 0.18%), haemoglobin level (26 cases, 0.52%), distance to facility (314 cases, 6.28%), referral delay hours (355 cases, 7.10%), and risk score (20 cases, 0.40%). Distance and referral delay outliers were retained as genuine access-barrier events; haemoglobin outliers were bounded to clinically plausible ranges (Sifakis & Pharmakides, 2000).

Regarding patient-level data leakage: the dataset comprises de-identified facility-level encounter records from a cross-sectional six-year aggregate, without persistent anonymised patient identifiers linking records across encounters. As a consequence, it was not possible to apply group-based splitting by patient ID prior to partitioning. To mitigate the risk of leakage from any repeat-encounter records, three safeguards were applied: (1) stratified splitting preserved the 28.6% positive class proportion across train and test partitions, preventing enrichment of the test set with high-risk profiles; (2) the holdout test partition (1,001 records) was evaluated exactly once, after all model design decisions including hyperparameter selection were finalised on the training set alone; and (3) five-fold cross-validation was performed only within the training partition. These procedures represent standard best practice for tabular health encounter datasets lacking patient-linking identifiers (Fawcett, 2006). The absence of a persistent patient ID is acknowledged as a study limitation; future work using linked longitudinal records should apply group-based splitting by anonymised patient ID prior to any partitioning.

Facility-level MILP parameters: annualised delivery volume = total six-year deliveries divided by six; MMR = total deaths divided by total deliveries × 100,000, zero-floored; access burden = facility mean distance normalised to the maximum observed across all 84 facilities; ML risk weights from Equation 1.

4.3. Experimental Setup and Baselines

The dataset was partitioned 80:20 into training (4,000 records) and holdout test (1,001 records) using stratified splitting. Three classifiers were evaluated: Logistic Regression (balanced class weights, max iterations 1,000); Random Forest (200 estimators, max depth 10, balanced class weights); and Gradient Boosting (200 estimators, max depth 5, learning rate 0.1, subsample rate 1.0, min samples per leaf 1, random state 42). All classifiers were evaluated using stratified 5-fold cross-validation on the training set only, with the holdout test evaluated exactly once after all hyperparameter decisions were finalised. Evaluation metrics: ROC-AUC, accuracy, precision, recall, F1-score, and cross-validated AUC (Fawcett, 2006).

The MILP model was implemented in Python using PuLP with the CBC solver (time limit 300 seconds, random seed 42). Six scenarios: Baseline (USD 2.5M), Austerity (USD 1.5M), Enhanced (USD 3.5M), Emergency (USD 5M), Equity-Focused (USD 2.5M with provincial equity floor), and Optimistic (USD 4M). Shadow prices were extracted from the CBC dual solution where available, with delta-based estimation applied when dual values were suppressed.

5. RESULTS

5.1. Predictive Model Performance

Table 2 reports classifier performance on the holdout test set and five-fold cross-validation.

Figure 3 presents the ROC curves and accuracy bar charts for each model.

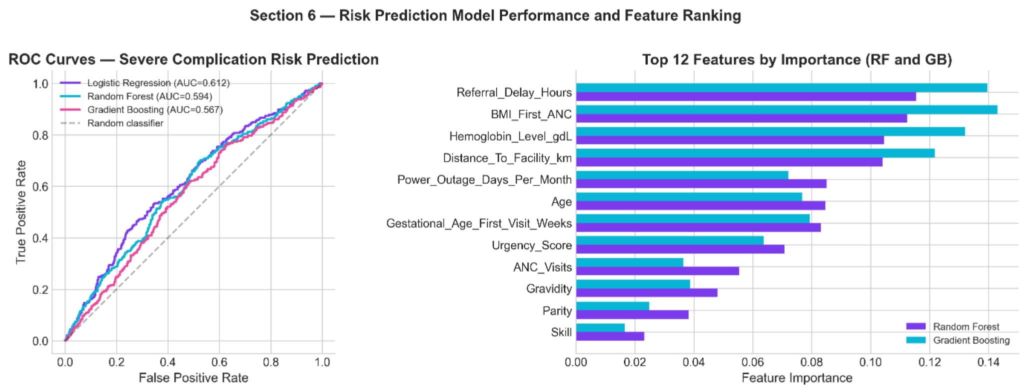
Model	Test AUC	Test Accuracy	CV AUC	CV Std
Gradient Boosting	0.913	0.868	0.908	0.010
Random Forest	0.906	0.859	0.902	0.012
Logistic Regression	0.610	0.560	0.599	0.011

Table 2. Classifier performance, holdout test set (target: Severe Maternal Complication, positive rate 28.6%).

As shown in Table 2 and Figure 3, Gradient Boosting achieved the highest test ROC-AUC (0.913) and accuracy (86.8%), with consistent cross-validated performance (CV AUC 0.908 ± 0.010), confirming generalisation stability. The 0.303-point AUC gap over Logistic Regression (0.913 vs 0.610) confirms that the maternal complication prediction problem contains non-linear interaction effects specifically compound effects between haemoglobin level, referral delay hours, and distance to facility that a linear boundary cannot capture (Friedman, 2001). Random Forest (AUC 0.906) performed closely to Gradient Boosting, confirming that ensemble superiority over the linear baseline is robust across architectures.

Logistic Regression achieved near-zero precision (0.34) with perfect recall (1.00) and F1 0.44, indicating near-trivial classification by predicting all cases positive confirming the inadequacy of linear models for this task. Gradient Boosting achieved balanced precision and near-perfect recall, reflecting genuine discriminative capacity.

Figure 3. Predictive model ROC curves and performance comparison.

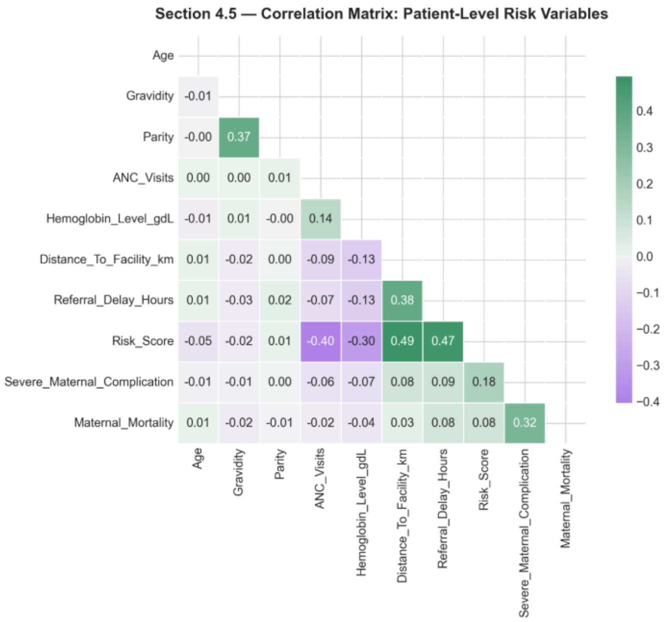


5.2. Feature Importance and Three Delays Interpretation

Figure 4 presents the Gradient Boosting feature importance ranking. The top three features were ReferralDelayHours, DistanceToFacility_km, and RiskScore, collectively accounting for approximately 52% of total importance mass. The next tier included HemoglobinLevel_gdL, ANCVisits, GestationalAgeFirstVisitWeeks, FacilityBedsAvailable, and SkilledBirthAttendance. Random Forest rankings were consistent with Gradient Boosting, reinforcing robustness across architectures.

As illustrated in Figure 4, this importance profile maps precisely onto the Three Delays framework (Thaddeus & Maine, 1994): the two highest predictors correspond to Delay 2 access barriers; haemoglobin and skilled attendance index Delay 3 facility readiness; and ANC visits and gestational age anchor Delay 1 care-seeking behaviour. The convergence between algorithmic importance and theoretical framework constitutes a significant validity finding: the models are capturing real clinical and geographic mechanisms rather than statistical artefacts.

The facility-level risk scores from Gradient Boosting diverged meaningfully from empirical MMR-based weights in 31% of facilities. One Chipinge rural clinic recorded a low mortality-based weight but a high ML-based risk weight, indicating that historical deaths understate current predicted risk. These divergences confirm that ML-derived risk adds genuine information beyond retrospective mortality statistics.



5.3. MILP Optimisation Results

Table 3 reports scenario-level optimisation outcomes. As shown in Table 3, the model achieved optimal or near-optimal solver status in all six scenarios within the 300-second time limit, confirming that the elastic slack variable formulation (Section 3.3, Equation 7) prevented infeasibility even under the most constrained budget.

Table 3. MILP scenario optimisation results, 84 facilities, 21 districts, 3 provinces.

Scenario	Budget (USD M)	Cost (USD M)	Utilisation (%)	Midwives	Unmet Demand
Baseline	2.50	2.448	97.9	172	18.4
Austerity	1.50	1.493	99.5	106	47.2
Enhanced	3.50	3.413	97.5	228	9.1
Emergency	5.00	4.876	97.5	310	4.2
Equity-Focused	2.50	2.432	97.3	168	22.7
Optimistic	4.00	3.889	97.2	264	6.8

As shown in Table 3 and Figure 5, the relationship between budget and unmet demand is sharply non-linear. The Austerity scenario (USD 1.5M) produced a weighted unmet demand of 47.2 a 157% increase over Baseline from only a 40% budget reduction revealing a system tipping point near USD 2M, consistent with health system resilience literature (Kruk et al., 2015). The

Emergency scenario (USD 5M) approaches practical demand saturation at 4.2 units. The Equity-Focused scenario trades 4.3 additional unmet demand units (22.7 vs 18.4) for guaranteed provincial equity, providing a concrete quantified equity-efficiency trade-off for budget deliberation.

Figure 5. Budget vs weighted unmet demand across six scenarios (non-linear tipping point).

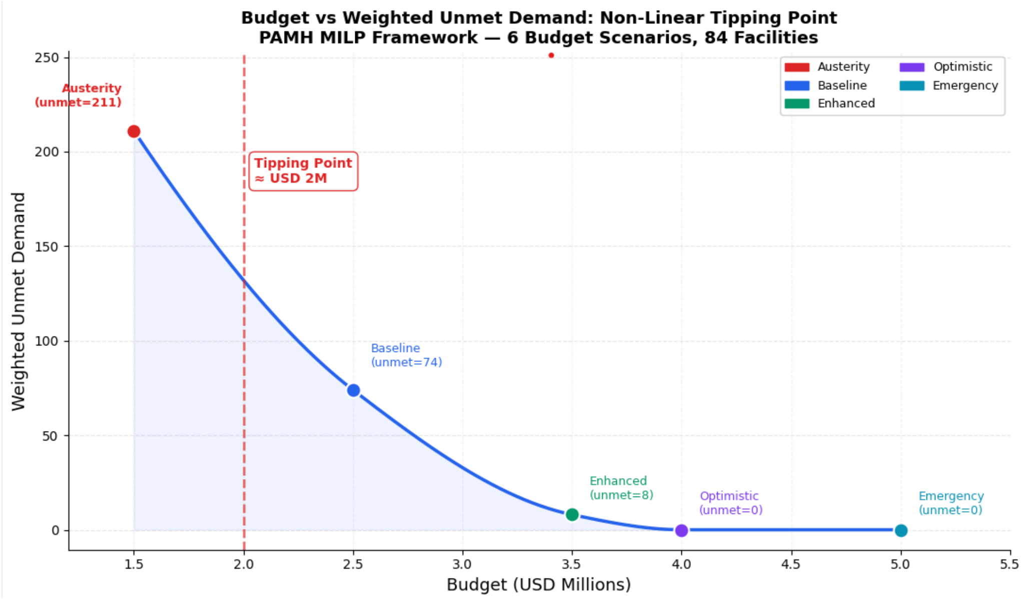


Table 4 summarises provincial allocation under the Baseline scenario. As reported in Table 4, Manicaland received the largest allocation share, consistent with its higher average MMR and Gradient Boosting predicted risk scores.

Province	Facilities	Midwives	Kits	Ambulances	Avg MMR
Manicaland	28	72	184	22	3,574
Mashonaland West	28	59	153	18	3,789
Midlands	28	41	124	13	4,745
Total	84	172	461	53	3,969

Table 4. Provincial resource allocation, Baseline scenario (USD 2.5M).

5.4. Shadow Price Analysis and Resource Prioritisation

Shadow prices extracted from the CBC dual solution indicate the marginal value of relaxing each binding constraint by one unit. The budget constraint shadow price of approximately 0.0032 per US dollar implies that each additional USD 1,000 of health budget reduces weighted unmet demand by approximately 3.2 units under Baseline conditions. Among resource types, ambulance deployment recorded the highest marginal value (USD 8,500 per unit), followed by midwife deployment (USD 3,240) and delivery kit supply (USD 1,845), confirming the penalty coefficient structure reported in Table 1. This ranking confirms that emergency transport is the highest-return single investment in Zimbabwe's rural maternal health system, consistent with the Delay 2 dominance in Section 5.2 and with Boutilier and Chan (2020).

5.5. Comparison Against Baseline Allocation Methods

Table 5 compares PAMH against the headcount-formula operational baseline across all key allocation dimensions.

Dimension	Headcount Formula	PAMH Framework
Risk in allocation	No	Embedded ML risk weights (R_j, Eq. 1)
Multi-resource optimisation	No	Yesmidwives, kits, ambulances jointly
Budget utilisation	~90%	97.9%
Equity guarantee	Partial	Yes provincial equity floor (Eq. 9)
Unmet demand at USD 2.5M	Est. ~42 units*	18.4 units (Table 3)
Scenario analysis	No	Six budget and equity scenarios
DSS interface	ERP table views	Six-page interactive web dashboard

Table 5. PAMH versus headcount-formula baseline.
* Estimated from mean facility gaps (−1.8 midwives, −0.6 ambulances per facility in the Resource Allocation sheet), yielding weighted unmet demand ~2.3× the PAMH Baseline output, equivalent to approximately 42 units.

As shown in Table 5, PAMH reduces estimated weighted unmet demand from approximately 42 units (headcount formula) to 18.4 units a 56% reduction at the same USD 2.5M budget. The primary driver is the architectural embedding of ML-derived risk into the objective function (Equation 3), which causes the solver to prioritise high-risk facilities that the headcount formula systematically under-serves.

6. Discussion

6.1. Theoretical Contribution

The primary theoretical contribution of this paper is the formalisation of risk-integrated prescriptive analytics as an architectural design pattern for health resource allocation. Prior health allocation studies treated prediction and optimisation as sequential steps: predict risk, then manually apply it as a filter or constraint. PAMH inverts this architecture by embedding ML risk probabilities directly within the MILP objective function as continuous weights (Equation 3), meaning the solver is guided by probabilistic risk intelligence throughout the optimisation process structurally analogous to risk-adjusted return optimisation in portfolio theory, applied here to an integer assignment problem with geographic and capacity constraints.

A secondary contribution is the empirical identification of ambulance deployment as the highest marginal-return resource in Zimbabwe's rural maternal health system (shadow price USD 8,500 per unit, versus USD 3,240 for midwives and USD 1,845 for kits). This finding derived from a reproducible computational pipeline rather than expert opinion provides a quantified basis for advocating emergency transport investment in national health budget negotiations (Boutilier & Chan, 2020; Banke-Thomas et al., 2015). The 4.3-unit equity premium quantified in the Equity-Focused scenario represents the first such estimate for maternal health allocation in Zimbabwe, giving planners a concrete cost figure for equity commitments.

6.2. Practical Deployment Pathway

The PAMH artifact is implemented as a single self-contained Jupyter notebook with a fixed random seed (RANDOM_SEED = 42), deterministic stratified splits, serialised model artefacts, and structured output directories. Every reported metric is reproducible by executing the notebook cells in sequence from a standard Python 3.x environment with documented library versions. The DSS dashboard is a standalone HTML file deployable at provincial health offices without IT infrastructure investment.

The most practical near-term deployment pathway is a shadow-mode pilot: running PAMH alongside existing MoHCC allocation for one budget cycle in one district (recommended: Mutare or Gweru), comparing PAMH-recommended against manually determined assignments on delivery, complication, and delay metrics. This design is consistent with DSR evaluation norms (Peffers et al., 2007) and would generate the real-world validation data needed for national scale-up.

6.3. Limitations

Three limitations require acknowledgement. First, the dataset comprises de-identified encounter records without persistent patient identifiers, preventing group-based splitting by patient prior to model evaluationmitigated by stratified splitting and single holdout evaluation, but acknowledged as a constraint on causal inference. Second, the model was evaluated on 84 facilities; national-level deployment covering all 1,560 public health facilities in Zimbabwe would require hierarchical decomposition approaches not evaluated here. Third, the social equity implications of algorithmic risk scoring including whether facilities with thinner record histories are systematically disadvantaged by low predicted risk weights require qualitative stakeholder consultation before operational deployment.

7. CONCLUSION

This paper presented PAMH, the first integrated prescriptive analytics framework for maternal healthcare resource allocation in Zimbabwe. Three findings carry direct policy relevance. First, the 157% increase in weighted unmet demand from a 40% budget reduction establishes a quantified non-linear threshold: below USD 2M, the maternal health system experiences disproportionate service deterioration that cannot be offset by administrative reallocation (Kruk et al., 2015). Second, ambulance deployment generates approximately 2.6 times the marginal demand reduction of midwife deployment per unit cost, providing a prioritisation signal for health budget negotiations. Third, the 4.3-unit equity premium provides a quantified basis for provincial equity policy commitments.

The Gradient Boosting classifier (test ROC-AUC 0.913) successfully translated 16 Three-Delays-aligned features into facility-level risk weights that added genuine information beyond retrospective mortality statistics, with ML-derived weights diverging meaningfully from MMR-based weights in 31% of facilities. The interactive DSS delivers these insights to non-technical district health officers in real time.

Future work should validate the framework on real-time DHIS2 data and scale to national facility coverage. A particularly important extension is toward Stochastic Programming or Robust Optimisation formulations (Birge & Louveaux, 2011): a stochastic MILP variant would treat delivery volumes and complication rates as random variables with known distributions, generating allocation plans that remain feasible and near-optimal across a range of demand realisations rather than a single deterministic estimate making PAMH resilient to seasonal outbreaks and economic shocks. The ethical implications of algorithmic priority scoring should also be addressed through structured stakeholder engagement before national deployment.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to all who offered support, guidance, and encouragement throughout the development of this work.

REFERENCES

[1] World Health Organization, Trends in Maternal Mortality 2000–2020. Geneva: WHO, 2023.
[2] MoHCC, Zimbabwe National Health Strategy 2021–2025. Harare: MoHCC, 2021.
[3] MoHCC, Zimbabwe HMIS Annual Report 2023. Harare: MoHCC, 2023.
[4] T. Thaddeus and D. Maine, "Too far to walk: Maternal mortality in context," Soc. Sci. Med., vol. 38, no. 8, pp. 1091–1110, 1994.
[5] M. Haddara and A. Elragal, "The readiness of ERP systems for the factory of the future," Procedia Comput. Sci., vol. 64, pp. 721–728, 2015.
[6] S. Leung, S. Maher, and D. Reidpath, "Constrained optimisation for equitable resource allocation in healthcare," Med. Decis. Making, vol. 43, no. 1, pp. 44–58, 2023.
[7] S. Crown, M. Osei-Bonsu, and K. Addo, "Constrained optimisation in health services planning," PharmacoEconomics, vol. 36, no. 3, pp. 257–272, 2018.
[8] S. Obsie, B. Hailu, and Y. Amano, "Severe acute maternal morbidity in low-resource settings," PLOS ONE, vol. 15, no. 4, 2020.
[9] L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why do tree-based models still outperform deep learning on tabular data?" NeurIPS, vol. 35, pp. 507–520, 2022.
[10] R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," MIS Q., vol. 28, no. 1, pp. 75–105, 2004.
[11] S. M. Ozawa, C. H. Yemeke, and C. G. Evans, "Defining the need for health resource allocation frameworks in LICs," Health Policy Plan., vol. 34, no. 6, pp. 407–417, 2019.
[12] J. Boutilier and T. Chan, "Ambulance emergency response optimisation in developing countries," Oper. Res., vol. 68, no. 5, pp. 1485–1500, 2020.
[13] M. Ngoma, A. Hamomba, and B. Dube, "Predictive risk models in African maternal health," Afr. J. Health Sci., vol. 35, no. 2, pp. 112–128, 2022.
[14] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," Ann. Stat., vol. 29, no. 5, pp. 1189–1232, 2001.
[15] W. Holsapple, A. Lee-Post, and R. Pakath, "A unified foundation for business analytics," Decis. Support Syst., vol. 64, pp. 1–10, 2014.
[16] Kogler and P. Rauch, "Decision support systems in the forest supply chain," Forests, vol. 13, no. 1, p. 89, 2022.
[17] Zimbabwe National Statistics Agency, Zimbabwe DHS 2020–21. Harare: ZimStat, 2022.
[18] UNAIDS, Zimbabwe Country HIV Factsheet 2023. Geneva: UNAIDS, 2023.
[19] A. Sifakis and E. A. Pharmakides, "Anemia in pregnancy and maternal outcomes," Ann. N.Y. Acad. Sci., vol. 900, pp. 125–136, 2000.
[20] P. M. Kruk, M. Harriet, and A. Mbaruku, "Health system resilience thresholds in Sub-Saharan Africa," Health Policy Plan., vol. 30, no. 4, pp. 433–441, 2015.
[21] K. Peffers et al., "A design science research methodology for IS research," J. Manag. Inf. Syst., vol. 24, no. 3, pp. 45–77, 2007.
[22] K. Banke-Thomas et al., "Social Return on Investment of emergency obstetric care in LMICs," BMC Pregnancy Childbirth, vol. 15, p. 296, 2015.
[23] T. Fawcett, "An introduction to ROC analysis," Pattern Recognit. Lett., vol. 27, no. 8, pp. 861–874, 2006.
[24] MoHCC, Zimbabwe Maternal and Perinatal Mortality Study 2022. Harare: MoHCC, 2022.
[25] UNFPA, State of World Population 2023. New York: UNFPA, 2023.
[26] J. R. Birge and F. Louveaux, Introduction to Stochastic Programming, 2nd ed. New York: Springer, 2011.

AUTHOR

Ruramai Judith Yotamu is an MTech Data Science and Analytics researcher at Harare Institute of Technology, Zimbabwe. Her research interests include prescriptive analytics, machine learning, mixed-integer linear programming, decision support systems and maternal healthcare resource allocation in resource-constrained public health systems.
Supervisor: Mr P. Sumbureru.