



Research article

From health AI risk scores to budget-feasible decisions: Predictive optimization for public health resource allocation

Jinho Cha¹ and Eun-Young Kwon^{2,*}

¹ Department of Computer Science, Gwinnett Technical College, Lawrenceville, GA, USA

² AI Industry Research Center, Kwangwoon University, Seoul, Korea

* **Correspondence:** Email: eykwon@kma.ac.kr, eykwonclass0524@gmail.com.

Abstract: Health artificial intelligence (AI) risk models can rank individuals by predicted need, but a risk score alone does not determine a budget-feasible allocation when programs' value, fiscal cost, and access conditions differ. We study how AI risk scores can be translated into public health resource allocation decisions under a common budget. Using 2024 Behavioral Risk Factor Surveillance System (BRFSS) public use data, we estimate a 29-variable risk model for a retrospective self-reported diabetes/prediabetes-related target (area under the receiver operating characteristic curve (AUC) 0.8002; average precision 0.4331). A transparent five-variable comparator yields an AUC of 0.7781 and average precision of 0.3848. Relative to the full risk ranking, fiscal yield allocation improves the held-out retrospective value by 1.30%, 0.90%, 0.53%, and 0.42% across 5%, 10%, 20%, and 30% budgets, respectively, while closely matching the prospective binary benchmark. The direction is stable across 30 repeated model refits, although the bootstrap intervals do not exclude zero at every budget. Cost-stress and survey-weighted analyses further show that allocation value depends on cost heterogeneity and coverage objectives. Prediction therefore provides most first-order prioritization value, while downstream allocation offers smaller, context-dependent refinements. The analysis is retrospective and does not estimate causal intervention effects or identify previously undiagnosed disease.

Keywords: predictive optimization; health AI; public health resource allocation; budget-constrained selection; BRFSS

Mathematics Subject Classification: 90C27, 90C39, 90B36, 90B40, 90C15, 68T01

1. Introduction

Health artificial intelligence (AI) systems increasingly generate individual-level risk scores for screening, outreach, and preventive support. These scores rank people by predicted health need, but ranking alone does not determine how scarce public health resources should be assigned. A deployable

policy must also account for budget limits, implementation costs, access barriers, expected program value, and coverage requirements. The resulting question is operational rather than purely predictive: How should an AI risk signal be converted into a feasible allocation decision?

We focus on this downstream decision problem. The predictive model supplies a respondent-level risk signal; the allocation layer determines which respondents can be selected within a fiscal limit. This separation follows the view of Kleinberg et al. [1] that machine predictions acquire decision value only after they are linked to explicit payoffs, as well as the smart “predict, then optimize” (SPO) framework of Elmachoub and Grigas [2] and related work on decision-focused learning and operational costs [3–7].

The basic decision problem is a budgeted binary selection model

$$\max_{y_a \in \{0,1\}} \sum_{a=1}^N \mu_a y_a \quad \text{subject to} \quad \sum_{a=1}^N \kappa_a y_a \leq \mathcal{B}, \quad (1.1)$$

where y_a denotes selection, μ_a is the estimated program value, κ_a is the fiscal cost of screening or outreach, and $\mathcal{B}$ is the available budget. In knapsack terminology, the respondents are items, the program values are item values, and the costs are item weights, and the budget is capacity. This mapping is standard [8–16]; our purpose is to use it to examine when a health risk ranking and a budget-aware allocation lead to different choices.

The empirical study uses the Behavioral Risk Factor Surveillance System (BRFSS), a large U.S. public health survey. The survey system described by Mokdad et al. [17] and the 2023–2024 public use documentation from the Centers for Disease Control and Prevention (CDC) [18, 19] provide demographic, behavioral, access to care, and health status measures at the respondent level. We construct a retrospective diabetes/prediabetes-related target and estimate individual risk scores. Because the outcome is self-reported, it is not treated as a laboratory-confirmed diagnosis or as evidence of previously undiagnosed disease.

The central premise is that risk ranking and allocation ranking need not coincide. When the program values and costs are homogeneous, prioritizing the highest risk scores is sufficient. With heterogeneous value or cost, however, a lower-risk respondent can have a higher allocation priority. We therefore characterize the conditions under which risk ordering is preserved and those under which program value or fiscal yield changes the ordering.

Figure 1 summarizes this distinction. Survey information first produces an AI risk score and a risk ranking. Budget limits, heterogeneous health impact, access burden, and fiscal cost then create an operational gap between prediction and action. The allocation layer closes that gap by combining the risk signal with policy-defined value and cost to produce a budget-feasible selection.

The study makes five contributions. First, it formulates health AI-enabled public health prioritization as a budget-constrained allocation problem rather than a standalone classification task. Second, it separates predicted risk, program value, access priority, implementation burden, and fiscal cost so that predictive performance and allocation performance can be evaluated independently. Third, it characterizes when risk ordering is preserved and when heterogeneous value or cost reverses that ordering, without presenting standard knapsack results as new theory. Fourth, it distinguishes scalable ranking heuristics from the underlying binary optimization problem and solves the 0–1 benchmark directly in the empirical study, including solver status and optimality gaps. Fifth, the BRFSS analysis evaluates prediction value, downstream allocation value, parameter sensitivity, and coverage consequences under common budget scenarios.

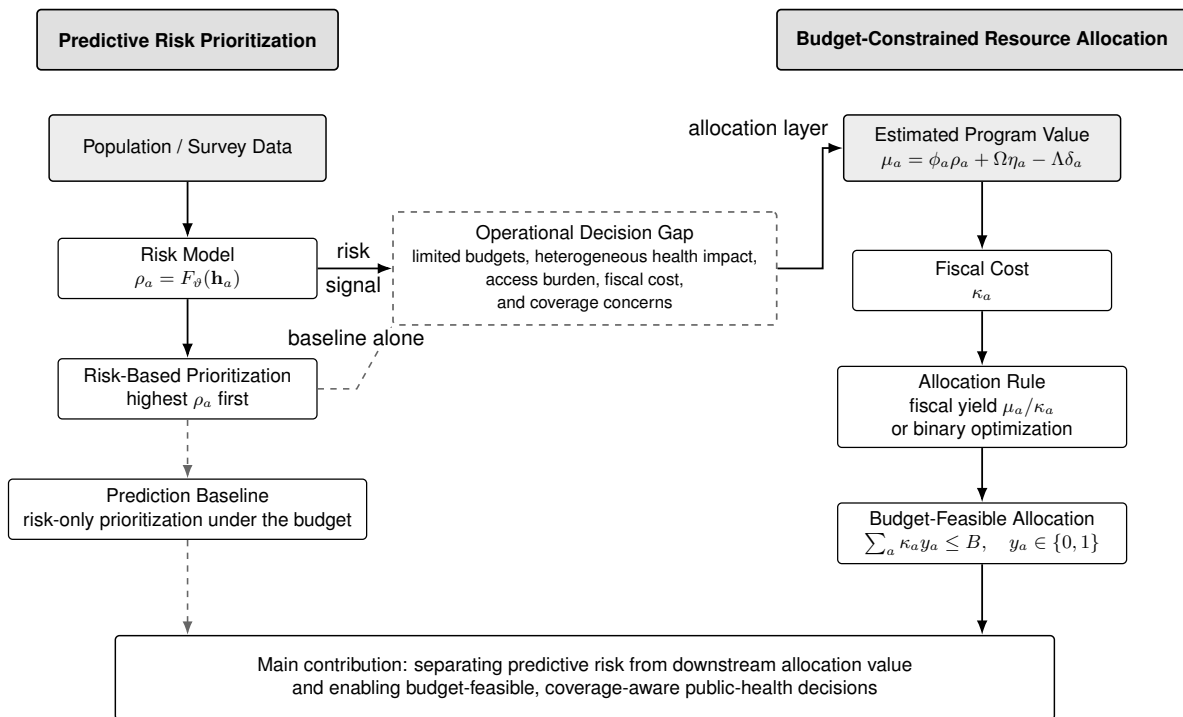


Figure 1. Conceptual positioning of the risk-to-allocation framework. An AI model produces respondent-level risk scores, but the resulting ranking does not by itself determine a budget-feasible public health allocation. The downstream decision layer combines the risk signal with policy-defined program value and fiscal cost under heterogeneous access and implementation conditions.

The remainder of the paper is organized as follows. Section 2 reviews predictive optimization, budgeted selection, and health AI deployment. Section 3 introduces the allocation model. Section 4 develops the structural properties of the risk-to-allocation mapping. Section 5 presents implementable allocation rules. Section 6 describes the BRFSS study design and evaluation metrics. Section 7 reports the empirical results, Section 8 discusses the interpretation and limitations, and Section 9 concludes the paper.

2. Literature review

2.1. Predictive optimization and decision-focused learning

The study draws on predictive optimization and decision-focused evaluation. Elmachoub and Grigas [2] formalized the predict-then-optimize principle by measuring the decision error induced by a prediction, while Wilder et al. [4], Donti et al. [5], Tulabandhula and Rudin [6], Wang et al. [20], and Deng et al. [7] developed related links between prediction and downstream operational decisions. From the operations research side, Bertsimas and Kallus [3] framed prescriptive analytics as the use of data and auxiliary covariates to prescribe decisions, building on stochastic optimization and sample-average

approximation foundations [21–23]. Applications using search, social media, and web signals for prediction further illustrate how auxiliary information can enter downstream decisions [24–28]. Human decision studies similarly emphasize that a prediction problem must be linked to an explicit payoff or decision rule before its value can be assessed [1].

Our setting differs from end-to-end predict-then-optimize learning in one important respect. It does not propose a new SPO loss, differentiable optimization layer, or jointly trained decision-focused learner. Instead, the predictive model is treated as an upstream source of respondent-level risk information, and the methodological question is how an already estimated risk signal should be translated into a budget-feasible public health allocation when policy value and fiscal cost are heterogeneous. Accordingly, the contribution lies in the risk-to-allocation interface and its empirical evaluation, rather than in a new prediction-focused training objective.

2.2. Budgeted resource allocation and knapsack-type selection

The downstream decision problem has the structure of a 0–1 knapsack-type budgeted selection problem. Following the classical formulations in Martello and Toth [8], Kellerer et al. [9], and Pisinger and Toth’s handbook chapter [10], each respondent is treated as a selectable item, the estimated program gain as the item value, the screening or outreach cost as the item weight, and the public health budget as the capacity. General integer programming and combinatorial optimization treatments [11–16] provide the surrounding complexity, relaxation, and approximation context.

This connection gives the fiscal yield rule a standard optimization interpretation rather than treating it as an ad hoc score. For the fractional relaxation, records are selected in nonincreasing order of program gain per unit of cost, with at most one fractional record at the budget boundary. For the binary problem, the same ordering is a scalable relaxation-guided deployment rule rather than an exact algorithm. The paper uses this standard optimization structure to study when the health AI risk ranking is sufficient and when cost-sensitive fiscal yield ranking changes the allocation.

2.3. Health AI risk prediction as an input to allocation

Health AI research has demonstrated that large clinical and population datasets can support individual-level risk prediction. Rajkomar et al. [29], Beam and Kohane [30], Topol [31], Obermeyer and Emanuel [32], Chen and Asch [33], and the JAMA Summit report [34] provide representative evidence of this broader shift. Responsible deployment roadmaps, however, emphasize that predictive models must be connected to the intended decision context, intervention, workflow, and monitoring process [35–39]. Complementary systems research also addresses the infrastructure needed for healthcare decision support; for example, Khan et al. [40] developed the smart medical data exchange engine (SMDEE) to support controlled access and privacy-aware medical data exchange. That system’s contribution is complementary to the present study: SMDEE concerns secure data access and exchange, whereas our focus is the downstream allocation decision once a risk signal and policy inputs are available. Population health algorithm studies by Obermeyer et al. [41], Ladin et al. [42], Rajkomar et al. [43], and Yang et al. [44] further show that proxy choices and downstream allocation rules can produce inequitable or distorted resource decisions when the predicted quantities are not aligned with clinical or public health need.

Together, these findings motivate a clear separation between prediction and allocation. The health AI risk score ρ_a is treated as an input signal, not as the allocation decision itself. The allocation layer

converts ρ_a into an operational program gain score μ_a , combines it with fiscal cost κ_a , and selects respondents under a common budget. This distinction is especially important for survey-based BRFSS data, where the screening priority target is retrospective and self-reported rather than a laboratory-confirmed diagnostic endpoint.

2.4. *Public health resource allocation and equity safeguards*

Public health screening programs must allocate finite resources across heterogeneous individuals and communities. In such settings, the decision-maker may care about total target capture, fiscal efficiency, segment coverage, and undercoverage of priority groups. Healthcare resource allocation studies such as Grand-Clément et al. [45], Deo et al. [46], Chan et al. [47], and Bates et al. [48] show that predictive information can be valuable when translated into constrained decisions. At the same time, Obermeyer et al. [41] and the Guidance for Unbiased Predictive Information for Healthcare Decision-Making and Equity (GUIDE) recommendations of Ladin et al. [42] motivate making the decision's objective and fairness safeguards explicit.

Coverage safeguards and equity priority weights enter only as allocation design components. It does not develop a new fairness metric or a new clinical prediction model. Instead, it studies how such weights and constraints alter budget-feasible selection when the underlying risk signal is fixed.

2.5. *Positioning of this paper*

A narrower gap remains at the interface between health AI risk scoring and public health deployment. Relative to decision-focused learning, this paper takes the predictive score as given and studies the downstream allocation rule rather than learning a new task-specific predictor. Relative to classical knapsack models, it does not claim the fractional value-to-cost ordering as new; instead, the optimization structure is used to make explicit when a health risk ranking and a value/cost-aware allocation can disagree. Relative to responsible health AI and healthcare decision support studies, the paper operationalizes that distinction through a common fiscal budget framework in which prediction quality, policy-defined program value, fiscal cost, and coverage implications can be evaluated separately. Accordingly, the contribution is a transparent risk-to-allocation formulation and empirical decision analysis, not a new clinical predictor, data exchange architecture, or generic knapsack algorithm.

3. **Budget-constrained predictive allocation model**

This section develops a population-level screening design model for public health programs operating under fiscal limits. The setting consists of a large survey-based population, statistical health risk indices constructed from individual survey responses, and a decision layer that determines which respondents or population segments should receive additional screening, outreach, counseling, or preventive care support. The central question is not only who appears medically vulnerable, but how a limited public health budget should be deployed across heterogeneous people, program targets, and access conditions.

Table 1. Notation for the public health allocation framework.

Symbol	Type	Description
Panel A: Population records and health targets		
a	Index	Survey respondent or population unit, $a = 1, \dots, N$.
N	Parameter	Number of respondents in the public health survey sample.
$\mathbf{h}_a$	Vector	Observed demographic, behavioral, access to care, and health status attributes of respondent a .
$\mathcal{H}$	Set	Survey feature space.
d_a	Observed indicator	Observed health risk or screening target indicator for respondent a .
q	Index	Screening program or public health target, $q = 1, \dots, Q$.
d_{aq}	Observed indicator	Indicator that Respondent a belongs to screening target q .
$\mathcal{S}$	Set	Set of population segments used for coverage analysis.
S_s	Set	Respondents belonging to segment $s \in \mathcal{S}$.
Panel B: Survey-based risk indices		
F_θ	Function	Statistical risk index model estimated from survey variables.
ρ_a	Score	Estimated health risk index for Respondent a , with $\rho_a \in [0, 1]$.
ρ_{aq}	Score	Estimated risk index for Respondent a and screening target q .
G_θ	Function	Dispersion or reliability estimation rule for the risk index.
σ_a	Score	Reliability adjustment or predictive dispersion score for Respondent a , with $\sigma_a \geq 0$.
$\hat{\rho}_a$	Score	Estimated probability of belonging to the target screening group.
$\hat{\rho}_{aq}$	Score	Estimated probability of belonging to screening target q .
Panel C: Program gain, fiscal cost, and design parameters		
κ_a	Cost	Fiscal cost of providing screening, outreach, counseling, or preventive support to Respondent a .
κ_{aq}	Cost	Fiscal cost of assigning Respondent a to screening target q .
$\mathcal{B}$	Parameter	Total public health screening budget.
$\mathcal{B}_q$	Parameter	Program-specific budget for screening target q .
$\mathcal{B}_s$	Parameter	Optional segment-level fiscal limit or coverage requirement.
ϕ_a	Weight	Expected health impact weight for Respondent a .
η_a	Weight	Optional underserved population or equity priority weight for Respondent a .
δ_a	Parameter	Access barrier, implementation burden, or outreach difficulty score for Respondent a .
μ_a	Value	Estimated program gain from selecting Respondent a .
$\hat{\mu}_a$	Estimate	Estimated program gain used for fiscal planning.
μ_a^{SA}	Value	Sensitivity-adjusted program gain score for respondent a , obtained by penalizing estimated gain for predictive dispersion or gain uncertainty.
μ_a^{real}	Value	Retrospective realized program gain for Respondent a , computed using the observed screening target and used only for evaluation.
μ_a / κ_a	Score	Fiscal yield score used to prioritize respondents by estimated program gain per unit of screening cost.
$\mu_a^{\text{SA}} / \kappa_a$	Score	Sensitivity-adjusted fiscal yield score used to prioritize respondents by sensitivity-adjusted program gain per unit of screening cost.
$\mu_a^{\text{real}} / \kappa_a$	Score	Oracle fiscal yield score used only for retrospective benchmarking.
ϱ_a	Parameter	Uncertainty radius for the estimated program gain.
Ω	Parameter	Weight controlling the role of equity priority in program gain.
Λ	Parameter	Penalty coefficient for access barriers or implementation burden.
ω_a^{svy}	Weight	Survey sampling weight for Respondent a , used for descriptive or survey-weighted sensitivity analysis when applicable.
Panel D: Program deployment decisions		
y_a	Binary decision	Equals one if Respondent a is selected for screening or preventive support.
y_{aq}	Binary decision	Equals one if Respondent a is assigned to screening target q .
$\mathbf{p}$	Policy	Deployment rule mapping survey-based indices and fiscal costs into selection decisions.
$y_a^{\mathbf{p}}$	Decision	Selection decision generated by deployment rule $\mathbf{p}$.
$\mathcal{Y}(\mathcal{B})$	Set	Feasible screening designs under budget $\mathcal{B}$.
$\mathbf{p}^{\text{or}}$	Policy	Retrospective oracle rule used only for benchmarking.
Panel E: Empirical parameter setting guidance		
Ω	Policy setting	Access priority weight. Baseline 0.15; sensitivity grid $\{0, 0.075, 0.15, 0.30\}$. Larger values place more weight on the access priority component η_a .
Λ	Policy setting	Access burden penalty. Baseline 0.10; sensitivity grid $\{0, 0.05, 0.10, 0.20\}$. Larger values penalize implementation burden more strongly.
γ	Policy setting	Cost heterogeneity coefficient in $\kappa_a = 1 + \gamma\delta_a$. Baseline 0.50; stress grid $\{0, 0.25, 0.50, 1, 2, 4\}$. Larger values widen the fiscal cost dispersion induced by δ_a .
ζ	Policy setting	Ambiguity penalty coefficient in $\mu_a^{\text{SA}} = \max\{\mu_a - \zeta\sigma_a, 0\}$. Baseline 0.50; sensitivity grid $\{0, 0.25, 0.50, 1\}$.
$\mathcal{B}$	Budget setting	Main experiments use 5%, 10%, 20%, and 30% of total held-out fiscal cost. Budget levels are policy scenarios rather than estimated parameters.
Panel F: Coverage safeguards and uncertainty adjustment		
L_s	Parameter	Minimum coverage requirement for Segment s .
U_s	Parameter	Optional maximum concentration limit for Segment s .
ℓ_s	Parameter	Lower coverage fraction for Segment s , when coverage is expressed proportionally.
u_s	Parameter	Upper concentration fraction for Segment s , when concentration is expressed proportionally.
Γ	Parameter	Optional tuning parameter controlling the strength of coverage safeguards.
ζ	Parameter	Optional tuning parameter controlling the penalty applied to predictive dispersion.

Continued on next page

Symbol	Type	Description
Panel G: Retrospective evaluation metrics		
$G_{\mathcal{B}}(\mathbf{p})$	Metric	Total realized program gain under deployment rule $\mathbf{p}$ and budget $\mathcal{B}$.
$\text{Capture}_{\mathcal{B}}(\mathbf{p})$	Metric	Fraction of all observed positive screening targets selected under budget $\mathcal{B}$.
$\text{PosRate}_{\mathcal{B}}(\mathbf{p})$	Metric	Positive screening target rate among respondents selected by deployment rule $\mathbf{p}$.
$\text{Yield}_{\mathcal{B}}(\mathbf{p})$	Metric	Realized program gain per unit fiscal screening cost.
$\text{SegCov}_{s,\mathcal{B}}(\mathbf{p})$	Metric	Screening coverage of population Segment s under budget $\mathcal{B}$.
$\text{Spend}_{\mathcal{B}}(\mathbf{p})$	Metric	Total fiscal screening cost spent by deployment rule $\mathbf{p}$.
$\text{OracleRecovery}_{\mathcal{B}}(\mathbf{p})$	Metric	Ratio of realized gain under $\mathbf{p}$ to realized gain under the retrospective oracle rule.
$\text{UnrecoveredOracleValue}_{\mathcal{B}}(\mathbf{p})$	Metric	Unrecovered benchmark value, defined as $1 - \text{OracleRecovery}_{\mathcal{B}}(\mathbf{p})$.
*Note: The framework separates survey-based risk estimation from budget-constrained allocation. Risk indices provide prioritization information, while the allocation rule also incorporates the program's value, fiscal cost, access priority, implementation burden, and predictive dispersion. The observed target indicators d_a and d_{aq} are used only for retrospective evaluation; the decisions y_a and y_{aq} are chosen prospectively from survey-based scores and budget constraints. The oracle is retained solely as a retrospective benchmark.		

Following the decision-focused perspective of Kleinberg et al. [1], Elmachoub and Grigas [2], and Bertsimas and Kallus [3], the formulation separates statistical risk estimation from program deployment. A health risk index is useful for public health planning only when it improves the use of scarce screening resources. Therefore, the model evaluates risk indices through the fiscal decisions they induce, rather than through classification accuracy alone. Table 1 summarizes the notation used in the population screening design framework.

3.1. Population health design setting

Consider a public health agency that must deploy limited screening or preventive care resources across a population of N survey respondents. Each respondent $a = 1, \dots, N$ is represented by a survey record containing demographic, behavioral, access to care, and health status attributes. The agency observes

$$\mathbf{h}_a \in \mathcal{H}, \quad (3.1)$$

where $\mathbf{h}_a$ may include age, sex, body-mass-related variables, smoking status, physical activity, general health status, prior diagnosis indicators, access to care, insurance status, and other public health survey variables.

The observed screening target for respondent a is denoted by $d_a \in \{0, 1\}$. In a retrospective empirical study, d_a is observed and can be used to evaluate alternative deployment rules. In prospective use, however, d_a is not available at the time of allocation. The agency must therefore rely on survey-based risk indices, program gain estimates, fiscal costs, and policy constraints when deciding which respondents to select.

The agency has a fiscal screening budget $\mathcal{B} > 0$. Selecting respondent a for screening, outreach, counseling, diagnostic follow-up, or preventive support consumes $\kappa_a > 0$ units of budget. The binary deployment decision is

$$y_a = \begin{cases} 1, & \text{if respondent } a \text{ is selected for screening or preventive support,} \\ 0, & \text{otherwise.} \end{cases} \quad (3.2)$$

The feasible design set is

$$\mathcal{Y}(\mathcal{B}) = \left\{ \mathbf{y} \in \{0, 1\}^N : \sum_{a=1}^N \kappa_a y_a \leq \mathcal{B} \right\}. \quad (3.3)$$

The decision problem is to choose $y \in \mathcal{Y}(\mathcal{B})$ so that limited public health resources are directed toward respondents or population segments with the greatest expected program gain.

This formulation intentionally separates three objects that are often conflated in screening applications. The first object is the observed health target d_a , which is used only for retrospective evaluation. The second object is the estimated risk index ρ_a , which summarizes the predicted probability or priority of Respondent a based on available survey variables. The third object is the deployment decision y_a , which must satisfy the fiscal constraint and may also incorporate the program's gain, equity priority, access burden, and uncertainty adjustment. This separation is central to the paper: Prediction accuracy is useful only insofar as it improves the constrained screening design.

3.2. Survey-based health risk indices

A statistical risk index model maps survey information into an estimated health risk:

$$\rho_a = F_{\theta}(\mathbf{h}_a), \quad \rho_a \in [0, 1]. \quad (3.4)$$

The index ρ_a may represent the estimated probability that Respondent a belongs to a target screening group. For diabetes-related screening, this risk index view is consistent with the Framingham diabetes model of Wilson et al. [49], the Finnish Diabetes Risk Score (FINDRISC) of Lindström and Tuomilehto [50], the screening model of Bang et al. [51], and broader diabetes screening and trend studies [52, 53].

When multiple public health targets are considered, the model produces target-specific indices

$$\rho_{aq} = F_{\theta,q}(\mathbf{h}_a), \quad \rho_{aq} \in [0, 1], \quad q = 1, \dots, Q. \quad (3.5)$$

Because survey-based estimates may be noisy, the framework may also include a dispersion or reliability adjustment score

$$\sigma_a = G_{\theta}(\mathbf{h}_a), \quad \sigma_a \geq 0. \quad (3.6)$$

This score may be obtained from bootstrap variation, ensemble disagreement, score error, prediction intervals, or another model-specific reliability measure. The role of σ_a is deliberately more modest than a full robust or distributionally robust model. The robust optimization literature of Soyster [54], Ben-Tal and Nemirovski [55], Bertsimas and Sim [56], and Ben-Tal et al. [57] motivates penalizing uncertain coefficients, while distributionally robust formulations [58–60] and risk measure approaches such as conditional value-at-risk (CVaR) [61] provide broader background. In the present paper, however, σ_a only indicates how much caution should be applied when using the nominal estimate in a fiscal deployment decision.

The statistical layer therefore produces the signal

$$\zeta_a = (\rho_a, \sigma_a, \mathbf{h}_a), \quad (3.7)$$

which becomes the input to the screening design layer. The risk index model is not assumed to make the final public health decision. Consistent with deployment-oriented health machine learning guidance [35, 38] and prediction-model reporting standards [62, 63], it provides population risk information that is converted into a fiscally feasible screening design.

3.3. Program gain score

Predicted health risk and program gain are distinct quantities. Predicted risk measures the estimated likelihood that a respondent belongs to a high-risk category. Program gain measures the expected public health value obtained by selecting that respondent for screening or preventive support. This separation is consistent with Kleinberg et al.'s payoff-based view of prediction [1], Vickers and Elkin's decision-analytic framing [64], Obermeyer et al.'s warning about proxy outcomes in population health algorithms [41], and the GUIDE emphasis on decisional context in resource allocation [42].

Let $\phi_a \geq 0$ denote the expected health impact weight associated with identifying or supporting Respondent a . This weight may reflect the condition's severity, preventable burden, the expected benefit of early identification, or the policy importance of the screening target. Let $\eta_a \geq 0$ denote an optional equity priority weight, and let $\delta_a \geq 0$ denote an access barrier or implementation burden. A general program gain specification is

$$\mu_a = \phi_a \rho_a + \Omega \eta_a - \Lambda \delta_a, \quad (3.8)$$

where $\Omega \geq 0$ controls the weight placed on equity priority and $\Lambda \geq 0$ penalizes implementation burden.

A baseline specification is

$$\mu_a = \phi_a \rho_a. \quad (3.9)$$

If all respondents have the same health impact weight, this reduces to ranking by the health risk index. Classical diabetes risk scores such as the Framingham diabetes score [49], FINDRISC [50], and the Bang et al. screening score [51] can therefore be viewed as special cases of risk index prioritization before the allocation layer adds the heterogeneous program gain and fiscal cost.

When multiple screening targets are considered, the target-specific program gain from assigning respondent a to target q can be written as

$$\mu_{aq} = \phi_{aq} \rho_{aq} + \Omega \eta_a - \Lambda \delta_{aq}, \quad (3.10)$$

where ϕ_{aq} is the health impact weight for target q and δ_{aq} is the target-specific implementation burden.

3.4. Fiscal screening design

The baseline design problem selects respondents to maximize total estimated program gain subject to the available fiscal budget

$$\max_y \sum_{a=1}^N \mu_a y_a \quad (3.11)$$

subject to

$$\sum_{a=1}^N \kappa_a y_a \leq \mathcal{B}, \quad (3.12)$$

$$y_a \in \{0, 1\}, \quad a = 1, \dots, N. \quad (3.13)$$

This is a fiscally constrained selection problem. If $\kappa_a = 1$ for all respondents, the problem reduces to selecting the $\mathcal{B}$ respondents with the largest positive program gains. If fiscal costs differ across respondents, the problem becomes a 0–1 knapsack problem in which the screening design depends on both program gain and fiscal cost. In the notation of Martello and Toth [8], Kellerer et al. [9],

Pisinger and Toth [10], and Nemhauser and Wolsey [12], this distinction is precisely the distinction between value ranking and capacity-constrained selection. It matters here because a respondent with high estimated health risk may not be the highest-priority deployment target when outreach or follow-up cost is also high.

A deployment rule $\mathfrak{p}$ maps survey profiles, risk indices, program gains, costs, and fiscal limits into a feasible screening design:

$$y_a^{\mathfrak{p}} = \mathfrak{p}(a; \mathbf{h}_a, \rho_a, \sigma_a, \mu_a, \kappa_a, \mathcal{B}). \quad (3.14)$$

The realized gain of rule $\mathfrak{p}$ under budget $\mathcal{B}$ is

$$\text{Gain}_{\mathcal{B}}(\mathfrak{p}) = \sum_{a=1}^N \mu_a y_a^{\mathfrak{p}}. \quad (3.15)$$

3.5. Multitarget prevention design

Public health programs may deploy resources across multiple screening targets, such as diabetes risk, cardiovascular risk, hypertension risk, and preventive care gaps. Let $q = 1, \dots, Q$ index screening targets. The binary decision y_{aq} equals one if Respondent a is assigned to screening target q . The multitarget design problem is

$$\max_y \sum_{a=1}^N \sum_{q=1}^Q \mu_{aq} y_{aq} \quad (3.16)$$

subject to

$$\sum_{a=1}^N \sum_{q=1}^Q \kappa_{aq} y_{aq} \leq \mathcal{B}, \quad (3.17)$$

$$\sum_{q=1}^Q y_{aq} \leq 1, \quad a = 1, \dots, N, \quad (3.18)$$

$$y_{aq} \in \{0, 1\}, \quad a = 1, \dots, N, \quad q = 1, \dots, Q. \quad (3.19)$$

Constraint (3.18) assigns each respondent to at most one screening target in the same deployment cycle. With this assignment restriction, the model follows the multiple choice knapsack structure described by Pisinger [65] and standard integer programming treatments [12, 13]. The constraint can be relaxed if the program allows multiple screening services for the same respondent.

Program-specific fiscal limits may also be imposed

$$\sum_{a=1}^N \kappa_{aq} y_{aq} \leq \mathcal{B}_q, \quad q = 1, \dots, Q. \quad (3.20)$$

These constraints are useful when funds are separated by disease program, screening modality, or intervention type.

3.6. Segment coverage safeguards

Public health screening design often requires balancing efficiency with population coverage. Let $\mathcal{S}$ denote a set of population segments, and let $\mathcal{S}_s \subseteq \{1, \dots, N\}$ be the respondents in segment s . A minimum coverage safeguard can be written as

$$\sum_{a \in \mathcal{S}_s} y_a \geq L_s, \quad s \in \mathcal{S}, \quad (3.21)$$

where $L_s \geq 0$ is the minimum number of selected respondents from segment s . To avoid excessive concentration in one segment, a maximum concentration safeguard can be imposed:

$$\sum_{a \in \mathcal{S}_s} y_a \leq U_s, \quad s \in \mathcal{S}. \quad (3.22)$$

Equity can also be incorporated directly into the program's gain score through the weight η_a in (3.8). The coverage safeguard approach and the equity weight approach serve different purposes. Coverage safeguards guarantee minimum or maximum deployment levels for specified segments, while equity weights adjust the marginal gain of selecting respondents from policy-relevant groups. The distinction is motivated by Obermeyer et al.'s analysis of proxy-induced bias in population health allocation [41], Rajkomar et al.'s discussion of fairness in machine-learning health systems [43], and the GUIDE recommendations for predictive information in resource allocation contexts [42].

3.7. Sensitivity-adjusted program gain

Program gain scores are estimated from risk indices, health impact weights, implementation burdens, and policy weights. These quantities may be noisy in survey data. Borrowing the general robust-optimization intuition of penalizing unstable coefficients from Soyster [54], Bertsimas and Sim [56], and Ben-Tal and Nemirovski [55], but without formulating a full uncertainty set optimization problem, let $\widehat{\mu}_a$ denote the estimated program gain and let $\varrho_a \geq 0$ denote a reliability or dispersion penalty derived from predictive uncertainty, score instability, or sensitivity analysis. The sensitivity-adjusted score is

$$\mu_a^{\text{SA}} = \max\{\widehat{\mu}_a - \varrho_a, 0\}. \quad (3.23)$$

This score is not introduced as an uncertainty set optimization model. Unlike full distributionally robust optimization formulations such as those of Wiesemann et al. [58] or the review of Rahimian and Mehrotra [66], it is an operational penalty used to test whether deprioritizing unstable estimated gains changes budgeted selection in the BRFSS allocation environment.

3.8. Limiting regimes

The proposed design model contains several useful limiting regimes.

First, if $\kappa_a = 1$ for all respondents and no coverage or sensitivity adjustments are imposed, the baseline design problem reduces to a top-gain selection rule:

$$\max \sum_a \mu_a y_a \quad \text{subject to} \quad \sum_a y_a \leq \mathcal{B}, \quad y_a \in \{0, 1\}. \quad (3.24)$$

The optimal rule selects the $\mathcal{B}$ respondents with the largest positive program gains.

Second, if $\mu_a = \rho_a$, the model reduces to risk index prioritization. This case is appropriate only when all respondents have identical health impact weights, identical fiscal costs, identical implementation burdens, and no additional policy weights.

Third, if $\mu_a = \phi_a \rho_a$ and the fiscal costs differ, the design becomes a cost-effectiveness screening problem. In this case, high estimated risk alone does not determine priority; the design also depends on the health impact weight and fiscal cost.

Fourth, if segment coverage safeguards (3.21)–(3.22) are included, the model becomes a coverage aware public health design problem. This formulation is useful when the agency wants to prevent scarce resources from becoming too concentrated in a narrow segment of the population.

Fifth, if program gains are uncertain and the sensitivity-adjusted formulation equation (3.23) is used, the model becomes a sensitivity-adjusted fiscal screening design. This version is useful when survey responses are incomplete, risk estimates are noisy, or the decision-maker wants to reduce sensitivity to unstable priority rankings.

These limiting regimes clarify the role of the general framework. Risk index prioritization is valid only under restrictive homogeneity conditions. Once the respondents differ in health impact, fiscal cost, access barriers, equity priority, or estimation uncertainty, the relevant decision object is program gain rather than predicted risk alone. This is the same prediction-to-decision distinction emphasized by Elmachtoub and Grigas [2], Bertsimas and Kallus [3], and Tulabandhula and Rudin [6].

4. Structural properties of risk-to-allocation conversion

This section clarifies the structural distinction between predictive risk ranking and budget-feasible allocation. The purpose is not to introduce a new knapsack algorithm. Instead, the analysis identifies when risk ranking is sufficient, how heterogeneous program values and fiscal costs can reverse that ranking, and why fiscal yield ordering is exact only for the classical fractional relaxation.

4.1. Fixed-budget population selection

The baseline deployment problem is

$$\max_y \sum_{a=1}^N \mu_a y_a \quad \text{subject to} \quad \sum_{a=1}^N \kappa_a y_a \leq \mathcal{B}, \quad y_a \in \{0, 1\}, \quad a = 1, \dots, N. \quad (4.1)$$

Throughout, $\kappa_a > 0$. The program value score μ_a is an operational prioritization score rather than an individual causal treatment effect. Negative values may be excluded from the eligible set or truncated at zero before allocation.

4.2. When risk ranking is sufficient

Recall

$$\mu_a = \phi_a \rho_a + \Omega \eta_a - \Lambda \delta_a. \quad (4.2)$$

Proposition 4.1 (Risk-ranking equivalence under homogeneity). *Suppose $\kappa_a = \kappa > 0$, $\phi_a = \phi > 0$, $\eta_a = \eta$, and $\delta_a = \delta$ for all respondents. Then $\rho_a > \rho_b$ if and only if $\mu_a > \mu_b$. Hence, under an integer cardinality budget and positive values, selecting the largest risk indices solves (4.1).*

Proof. Under the stated conditions, $\mu_a = \phi\rho_a + \Omega\eta - \Lambda\delta$, which is a positive affine transformation of ρ_a . Equal costs reduce the budget constraint to a cardinality limit, so the largest objective coefficients are selected. $\square$

This proposition is a characterization of the homogeneous benchmark, not a new optimization result. Its role is to identify the restrictive setting in which predictive risk and allocation priority coincide.

4.3. Unified reversal under value and cost heterogeneity

With heterogeneous costs, define fiscal yield

$$v_a = \frac{\mu_a}{\kappa_a}. \quad (4.3)$$

Proposition 4.2 (Risk-to-allocation reversal). *For two respondents a and b with $\rho_a > \rho_b$, fiscal yield prioritization ranks b ahead of a if and only if*

$$\kappa_a(\phi_b\rho_b + \Omega\eta_b - \Lambda\delta_b) > \kappa_b(\phi_a\rho_a + \Omega\eta_a - \Lambda\delta_a). \quad (4.4)$$

Proof. The condition is equivalent to $\mu_b/\kappa_b > \mu_a/\kappa_a$. Substituting (4.2) yields (4.4). $\square$

Thus, a higher-risk respondent need not have higher allocation priority when the health impact weights, access priority, implementation burden, or fiscal cost differ. With equal costs, the condition reduces to a reversal in program value alone. This unified condition replaces separate pairwise statements for program gain and cost-sensitive reversals.

4.4. Relation to classical knapsack optimization

For the fractional relaxation of (4.1),

$$\max_{0 \leq y_a \leq 1} \sum_{a=1}^N \mu_a y_a \quad \text{subject to} \quad \sum_{a=1}^N \kappa_a y_a \leq \mathcal{B}, \quad (4.5)$$

classical fractional knapsack theory implies nonincreasing ordering by μ_a/κ_a , with at most one fractional item at the boundary [8–10]. We therefore do not present this ordering as a new proposition or theorem.

For the binary problem, fiscal yield ranking is a scalable heuristic rather than an exact algorithm. The revised empirical analysis directly solves the 0–1 allocation benchmark and reports the prospective objective gap between fiscal yield prioritization and the binary benchmark. This comparison separates the practical performance of the heuristic from the optimality of the underlying integer program.

4.5. Multitarget screening and coverage safeguards

For multiple screening targets, the decision variable y_{aq} indicates whether Respondent a is assigned to target q . With the single target per respondent constraint, the formulation is a multiple choice knapsack-type selection problem

$$\max_y \sum_{a=1}^N \sum_{q=1}^Q \mu_{aq} y_{aq} \quad \text{subject to} \quad \sum_{a=1}^N \sum_{q=1}^Q \kappa_{aq} y_{aq} \leq \mathcal{B}, \quad \sum_{q=1}^Q y_{aq} \leq 1, \quad y_{aq} \in \{0, 1\}. \quad (4.6)$$

Program-specific budgets can be imposed by $\sum_a \kappa_{aq} y_{aq} \leq \mathcal{B}_q$. Segment safeguards can also be added through lower and upper coverage constraints

$$L_s \leq \sum_{a \in S_s} y_a \leq U_s, \quad s \in \mathcal{S}. \quad (4.7)$$

These safeguards are imposed only when the resulting feasible set is nonempty. In applications with overlapping segments or tight budgets, the requirements can be relaxed through penalty terms rather than imposed as hard constraints.

4.6. Sensitivity-adjusted deployment score

The empirical policies also include a sensitivity-adjusted score

$$\mu_a^{\text{SA}} = \max\{\widehat{\mu}_a - \varrho_a, 0\}, \quad (4.8)$$

where ϱ_a is a reliability or dispersion penalty. This is used as a sensitivity-adjusted deployment rule rather than as a full uncertainty set optimization model. Its purpose is to test whether penalizing unstable estimated gains changes budgeted selection in the BRFS allocation environment.

4.7. Retrospective oracle recovery

The estimated design value of a deployment rule is

$$\widehat{G}_{\mathcal{B}}(\mathbf{p}) = \sum_a \widehat{\mu}_a y_a^{\mathbf{p}}, \quad (4.9)$$

whereas retrospective evaluation uses realized benchmark gains

$$G_{\mathcal{B}}^{\text{real}}(\mathbf{p}) = \sum_a \mu_a^{\text{real}} y_a^{\mathbf{p}}. \quad (4.10)$$

For retrospective evaluation, let μ_a^{or} denote the benchmark gain constructed from observed screening priority labels. The oracle benchmark solves the same budgeted selection problem using μ_a^{or} , but it is not available at deployment time. For an implementable deployment rule $\mathbf{p}$, define

$$\text{OracleRecovery}_{\mathcal{B}}(\mathbf{p}) = \frac{\sum_{a=1}^N \mu_a^{\text{or}} y_a^{\mathbf{p}}}{\sum_{a=1}^N \mu_a^{\text{or}} y_a^{\text{or}}}, \quad (4.11)$$

whenever the oracle denominator is positive. The unrecovered benchmark value is

$$\text{UnrecoveredOracleValue}_{\mathcal{B}}(\mathbf{p}) = 1 - \text{OracleRecovery}_{\mathcal{B}}(\mathbf{p}). \quad (4.12)$$

These are retrospective recovery measures, not online learning performance guarantees.

4.8. Implications for public health screening design

The structural analysis yields four implications. First, health AI risk indices are sufficient deployment scores only under restrictive homogeneity conditions. Second, heterogeneous program gain can reverse the risk ranking. Third, heterogeneous costs can further reverse deployment order through fiscal yield. Fourth, retrospective oracle recovery provides a normalized way to compare implementable selection rules under the same budget without claiming that the oracle rule is available in practice.

5. Screening deployment rules and implementation

This section translates the fiscal screening design model into implementable public health deployment rules. The rules are designed to separate the effect of different decision principles: Health risk index prioritization, program gain prioritization, fiscal yield prioritization, coverage aware deployment, multitarget screening design, and sensitivity-adjusted program gain deployment. This separation allows the empirical analysis to evaluate whether public health screening performance improves when the deployment rule uses more of the available fiscal, health impact, access, and coverage structure.

5.1. Deployment rule classes

For each budget scenario, the decision-maker observes the candidate population, survey-based risk indices $\{\rho_a\}_{a=1}^N$, reliability scores $\{\sigma_a\}_{a=1}^N$, fiscal costs $\{\kappa_a\}_{a=1}^N$, estimated program gains $\{\mu_a\}_{a=1}^N$, and total budget $\mathcal{B}$. A deployment rule then selects a feasible design $y \in \mathcal{Y}(\mathcal{B})$.

We compare five implementable deployment rules, a random baseline, and a retrospective oracle benchmark. The five implementable rules are health risk index prioritization, program gain prioritization, fiscal yield prioritization, sensitivity-adjusted gain prioritization, and sensitivity-adjusted fiscal yield prioritization. Coverage-aware and multitarget formulations are retained as model extensions rather than main empirical rule classes.

Health risk index prioritization. The health risk index rule ranks respondents by the survey-based risk index ρ_a :

$$\text{Index}_a^{\text{risk}} = \rho_a. \quad (5.1)$$

This rule represents the common practice of prioritizing respondents who appear most medically vulnerable according to the predictive risk model.

Program gain prioritization. The program gain rule ranks respondents by the estimated public health gain μ_a :

$$\text{Index}_a^{\text{gain}} = \mu_a. \quad (5.2)$$

This rule uses the predicted risk, expected health impact, equity priority, and access burden when these components are included in the gain specification.

Fiscal yield prioritization. The fiscal yield rule ranks respondents by estimated program gain per unit of budget:

$$\text{Index}_a^{\text{yield}} = \frac{\mu_a}{\kappa_a}. \quad (5.3)$$

This rule accounts for the fact that limited public health budgets should be used where expected gain per fiscal unit is high.

Coverage-aware deployment. The coverage aware rule maximizes program gain while enforcing segment-level coverage safeguards. It solves

$$\max_y \sum_{a=1}^N \mu_a y_a \quad (5.4)$$

subject to

$$\sum_{a=1}^N \kappa_a y_a \leq \mathcal{B}, \quad (5.5)$$

$$L_s \leq \sum_{a \in \mathcal{S}_s} y_a \leq U_s, \quad s \in \mathcal{S}, \quad (5.6)$$

$$y_a \in \{0, 1\}, \quad a = 1, \dots, N. \quad (5.7)$$

This rule is used to evaluate how minimum coverage and maximum concentration limits affect the total gain, high-risk hit rate, and segment balance.

Multitarget screening deployment. When several screening targets are available, the multitarget rule assigns respondents to the disease or prevention categories. It solves

$$y^{\text{multi}} \in \arg \max_y \sum_{a=1}^N \sum_{q=1}^Q \mu_{aq} y_{aq} \quad (5.8)$$

subject to

$$\sum_{a=1}^N \sum_{q=1}^Q \kappa_{aq} y_{aq} \leq \mathcal{B}, \quad (5.9)$$

$$\sum_{q=1}^Q y_{aq} \leq 1, \quad a = 1, \dots, N, \quad (5.10)$$

$$y_{aq} \in \{0, 1\}. \quad (5.11)$$

This rule determines not only whether a respondent should be selected but also which screening target produces the greatest program gain.

Sensitivity-adjusted program gain deployment. The sensitivity-adjusted rule accounts for uncertainty in estimated program gain. It uses the lower-confidence gain

$$\mu_a^{\text{SA}} = \max\{\widehat{\mu}_a - \varrho_a, 0\} \quad (5.12)$$

and selects a feasible design using μ_a^{SA} instead of $\widehat{\mu}_a$. This rule tests whether the public health budget changes when unstable estimated gains are penalized while excluding negative adjusted deployment values.

5.2. Single-program budget implementation

When the screening program has a single aggregate budget, each deployment rule can be implemented as a ranking or knapsack procedure. Let H_a denote a generic deployment index. For the risk index rule, $H_a = \rho_a$. For the program gain rule, $H_a = \mu_a$. For the fiscal yield rule, $H_a = \mu_a / \kappa_a$. For the sensitivity-adjusted gain rule, $H_a = \mu_a^{\text{SA}} = \max\{\widehat{\mu}_a - \varrho_a, 0\}$. For the sensitivity-adjusted fiscal yield rule, $H_a = \mu_a^{\text{SA}} / \kappa_a$.

If fiscal costs are identical, $\kappa_a = 1$, the rule selects the largest positive index respondents subject to the budget:

$$\mathcal{D}(H) = \{a : H_a \text{ is among the } \min\{\mathcal{B}, M_+(H)\} \text{ largest positive indices}\}, \quad (5.13)$$

where

$$M_+(H) = |\{a : H_a > 0\}|. \quad (5.14)$$

The corresponding deployment decision is

$$y_a = \begin{cases} 1, & a \in \mathcal{D}(H), \\ 0, & \text{otherwise.} \end{cases} \quad (5.15)$$

If fiscal costs differ, the ranking rule is used as a transparent heuristic and the exact fiscal design has the following 0–1 knapsack form:

$$\max_y \sum_{a=1}^N H_a y_a \quad \text{subject to} \quad \sum_{a=1}^N \kappa_a y_a \leq \mathcal{B}, \quad y_a \in \{0, 1\}. \quad (5.16)$$

This implementation ensures that every policy comparison is made under the same budget limit. In the revised empirical study, (5.16) is solved directly at each main budget using the prospective program value coefficients. The solver status and relative mixed-integer programming (MIP) gap are recorded, and the resulting selected set is evaluated retrospectively using the same held-out outcome metric as the ranking rules. Accordingly, fiscal yield ranking is treated as a heuristic and its proximity to the binary benchmark is an empirical result rather than an assumed optimality property.

5.3. Multitarget program implementation

When the respondents can be assigned to multiple screening targets, the implementation uses target-specific gain and cost values. The feasible set is

$$\mathcal{Y}_Q(\mathcal{B}) = \left\{ y : \sum_{q=1}^Q y_{aq} \leq 1, \quad a = 1, \dots, N, \quad \sum_{a=1}^N \sum_{q=1}^Q \kappa_{aq} y_{aq} \leq \mathcal{B}, \quad y_{aq} \in \{0, 1\} \right\}. \quad (5.17)$$

The multitarget deployment rule solves

$$\max_{y \in \mathcal{Y}_Q(\mathcal{B})} \sum_{a=1}^N \sum_{q=1}^Q \mu_{aq} y_{aq}. \quad (5.18)$$

For large samples, a scalable fiscal yield heuristic can be used. It ranks respondent–target pairs by

$$\Upsilon_{aq} = \frac{\mu_{aq}}{\kappa_{aq}}, \quad (5.19)$$

selects pairs in descending order, enforces the budget, and prevents a respondent from being assigned to more than one target unless the study design permits multiple services per respondent.

Assignments with nonpositive target gain may be excluded by imposing

$$y_{aq} = 0 \quad \text{whenever} \quad \mu_{aq} \leq 0. \quad (5.20)$$

This prevents the deployment rule from spending scarce public health resources on screening targets that do not generate positive estimated program gain.

5.4. Budget scenario evaluation procedure

All deployment rules are evaluated across common budget scenarios. For a budget level $\mathcal{B}$, the procedure constructs survey-based risk indices, computes program gain or fiscal yield values, applies the deployment rule, and records the retrospective performance metrics. Unlike a queueing procedure, this is a static budget scenario evaluation: Each budget scenario represents one possible public health planning environment.

For a deployment rule p , the selected set is

$$\mathcal{D}_{\mathcal{B}}(p) = \{a : y_a^p = 1\}. \quad (5.21)$$

Budget use is

$$\text{Spend}_{\mathcal{B}}(p) = \frac{\sum_{a=1}^N \kappa_a y_a^p}{\mathcal{B}}. \quad (5.22)$$

Total estimated gain is

$$\text{Gain}_{\mathcal{B}}(p) = \sum_{a=1}^N \mu_a y_a^p. \quad (5.23)$$

If observed high-risk indicators d_a are available, the high-risk hit rate under budget $\mathcal{B}$ is

$$\text{HitRate@}\mathcal{B} = \frac{\sum_{a=1}^N d_a y_a^p}{\sum_{a=1}^N d_a}. \quad (5.24)$$

The omitted high-risk share is

$$\text{Omitted}_{\mathcal{B}}(p) = 1 - \text{HitRate@}\mathcal{B}. \quad (5.25)$$

Segment coverage is evaluated as

$$\text{SegCov}_{s,\mathcal{B}}(p) = \frac{\sum_{a \in \mathcal{S}_s} y_a^p}{|\mathcal{S}_s|}, \quad s \in \mathcal{S}. \quad (5.26)$$

These metrics allow the empirical analysis to compare the total program gain, high-risk recovery, budget use, fiscal yield, and segment balance under the same public health budget.

5.5. Algorithm summary

Algorithm 1 summarizes the implementation used for each budget scenario.

Algorithm 1 Budget scenario for public health screening deployment

Require: Survey records $\{\mathbf{h}_a\}_{a=1}^N$, fiscal costs $\{\kappa_a\}_{a=1}^N$, budget $\mathcal{B}$, deployment rule R

Ensure: Deployment decisions y^R , selected set $\mathcal{D}_\mathcal{B}(R)$, retrospective performance metrics

- 1: Estimate health risk indices $\rho_a = F_\theta(\mathbf{h}_a)$ for all respondents.
 - 2: Estimate reliability or dispersion scores $\sigma_a = G_\theta(\mathbf{h}_a)$, if used.
 - 3: Compute program gains μ_a , fiscal yield indices μ_a/κ_a , sensitivity-adjusted gains $\mu_a^{SA} = \max(\widehat{\mu}_a - \varrho_a, 0)$, or sensitivity-adjusted fiscal yield indices μ_a^{SA}/κ_a , depending on rule R .
 - 4: **if** R is a simple ranking rule and $\kappa_a = 1$ **then**
 - 5: Select the largest positive index respondents subject to budget $\mathcal{B}$.
 - 6: **else if** R is a cost-sensitive single-program rule **then**
 - 7: Apply the fiscal yield rule as a scalable relaxation-guided heuristic; solve the exact 0–1 knapsack problem only when an exact benchmark is explicitly reported.
 - 8: **else if** R is a coverage aware rule **then**
 - 9: Solve the budget-constrained design with segment safeguards.
 - 10: **else if** R is a multitarget rule **then**
 - 11: Solve the respondent–target deployment problem or apply the target yield heuristic.
 - 12: **end if**
 - 13: Record selected set $\mathcal{D}_\mathcal{B}(R) = \{a : y_a^R = 1\}$.
 - 14: Compute program gain, high-risk hit rate, omitted high-risk share, budget spend, fiscal yield, segment coverage, and retrospective oracle recovery metrics when available.
-

5.6. Computational considerations

The deployment rules are computationally lightweight. For simple ranking rules, one scalar index is computed for every respondent and the respondents are sorted. Thus, the computational cost is

$$O(N \log N). \quad (5.27)$$

For multitarget screening design, indices are computed for all feasible respondent–target pairs. If there are Q screening targets, score construction requires $O(NQ)$ operations. A greedy target yield implementation that sorts respondent–target pairs requires

$$O\{NQ \log(NQ)\} \quad (5.28)$$

time.

The single-program design is a 0–1 knapsack problem when fiscal costs are heterogeneous. The multitarget design becomes a generalized assignment or multiple choice knapsack problem when each respondent can be assigned to at most one target. Ranking and fiscal yield rules remain scalable alternatives for large public health survey data, but the main BRFSS experiment now also solves the single-program 0–1 benchmark directly at every reported budget and records the solver optimality gap. This implementation makes the distinction between exact integer optimization and relaxation-guided heuristics explicit [9, 10, 14].

The retrospective oracle rule used in the empirical analysis is not an implementable deployment rule. It is used only as an upper-bound benchmark for computing retrospective oracle recovery ratios and unrecovered benchmark values. Following Kleinberg et al.’s emphasis on decision counterfactuals [1] and Bertsimas and Kallus’s policy value perspective [3], all implementable rules use the same survey records, budget scenarios, risk index estimates, fiscal costs, and coverage definitions. Therefore, performance differences are attributable to the deployment rule rather than to different data-generating conditions.

The next section uses this common implementation framework to compare health risk index prioritization, program gain prioritization, fiscal yield deployment, coverage aware screening design, multitarget deployment, and sensitivity-adjusted program gain deployment.

6. Computational study: BRFSS as a public health allocation environment

This section describes the computational design used to evaluate the proposed risk-to-allocation framework on a large public health survey dataset. The purpose of the experiment is to assess how alternative budget-constrained allocation rules convert individual-level risk information into deployable public health prioritization decisions. The analysis uses the 2024 Behavioral Risk Factor Surveillance System (BRFSS) public use file. The unit of analysis is an adult survey respondent, and the retrospective outcome is a self-reported diabetes/prediabetes-related indicator constructed from the BRFSS diabetes-status question.

The experiment is retrospective and should not be interpreted as a clinical diagnostic study. The objective is not to claim that the survey variables alone can replace clinical screening, laboratory testing, or a physician's diagnosis. Rather, the experiment evaluates whether routinely collected public health survey variables contain sufficient signal to support budget-constrained prioritization and whether deployment rules that account for program value, access burden, and fiscal cost can improve over risk-only prioritization.

6.1. Analysis sample and screening target

The initial BRFSS public use file contains 457,670 respondents. We retain the variables needed to construct the screening target, risk index features, access burden, program gain components, fiscal cost proxy, and survey design identifiers. After excluding respondents with missing diabetes screening target status and removing observations with excessive missing features, the final analysis sample contains 456,625 respondents. The held-out test set contains 136,988 respondents.

Table 2. BRFSS 2024 analysis sample and predictive benchmarks.

Quantity	Value
Raw selected BRFSS respondents	457,670
Final analysis respondents	456,625
Held out test respondents	136,988
Primary target positive rate	0.1689
Main model predictors	29
Main model area under the receiver operating characteristic curve (AUC) / average precision	0.8002 / 0.4331
Five-variable comparator predictors	Age, sex, BMI, exercise, general health
Five-variable comparator AUC / average precision	0.7781 / 0.3848
Diagnosed-only AUC / average precision	0.8083 / 0.3980
Prediabetes-only AUC / average precision	0.6902 / 0.0490

*Notes: The main outcome is the self-reported diabetes/prediabetes-related target constructed from DIABETE4. The five-variable comparator is a transparent nonrandom benchmark using age, sex, body mass index (BMI), recent exercise, and general health. Published clinical scores such as FINDRISC or the Framingham diabetes score are not reproduced directly because the BRFSS public use file does not contain all inputs in their original clinical definitions. The diagnosed-only and prediabetes-only rows report separate target sensitivity models.

Table 3. Construction of BRFSS 2024 analysis variables and their roles in the fiscal screening design.

Category	BRFSS variable(s)	Operational definition	Role in the framework
Panel A: Screening target construction			
Target outcome	DIABETE4	Self-reported diabetes status. The main target codes respondents with diagnosed diabetes or prediabetes/borderline diabetes as positive screening targets and respondents reporting no diabetes as negative screening targets. Missing, refused, or uncertain responses are excluded.	Defines the retrospective screening target d_a used for evaluation and for constructing the realized gain benchmark.
Supplementary diagnostic target	PREDIAB2	Auxiliary prediabetes indicator retained for sensitivity checks and sensitivity analysis.	Supports alternative target construction and sensitivity analysis.
Panel B: Demographic and socioeconomic predictors			
Sex	SEXVAR, _SEX	Observed and calculated sex variables.	Used by the risk index model F_θ to capture baseline demographic heterogeneity in screening risk.
Age	_AGE5YR, _AGE80	Five-year age categories and imputed age collapsed above 80.	Used in the risk index model to capture age-related screening risk and age-linked disease prevalence.
Education	EDUCA, _EDUCAG	Reported and grouped education attainment.	Used in the risk index model and in the access priority score η_a to reflect screening vulnerability and structural disadvantage.
Income	INCOME3, _INCOMG1	Reported and grouped household income categories.	Used in the risk index model and in the access priority score η_a to capture socioeconomic disadvantage and unequal screening access.
Panel C: General health and well-being predictors			
General health	GENHLTH	Self-reported general health status.	Used in the risk index model and indirectly informs the expected health impact weight ϕ_a .
Physical health burden	PHYSHLTH	Number of days in the last 30 days during which physical health was not good. Code 88 is recoded to 0 days; invalid nonresponse codes are treated as missing.	Used in the risk index model to capture short-run health burden and elevated screening need.
Mental health burden	MENTHLTH	Number of days in the last 30 days during which mental health was not good. Code 88 is recoded to 0 days; invalid nonresponse codes are treated as missing.	Used in the risk index model to capture short-run well-being burden correlated with health risk profiles.
Combined poor health days	POORHLTH	Number of days poor physical or mental health limited usual activities. Code 88 is recoded to 0 days; invalid nonresponse codes are treated as missing.	Used in the risk index model as a summary morbidity-related feature.
Calculated health status	_RFHLTH	Calculated fair/poor versus better health status.	Used both in the risk index model and in the health impact weight ϕ_a .
Panel D: Access-to-care and implementation burden predictors			
Primary health insurance source	PRIMINS2	Reported primary health insurance source.	Used as an auxiliary access-related predictor in risk estimation and descriptive characterization.
Health insurance coverage	_HLTHPL2	Calculated indicator of health insurance coverage.	Used in the risk index model and in the access burden score δ_a .
Personal provider access	PERSDOC3	Indicator of whether the respondent has a personal healthcare provider.	Used in the risk index model, the access burden score δ_a , and the access priority score η_a .
Cost barrier to care	MEDCOST1	Indicator of inability to see a doctor because of cost.	Used in the risk index model, the access burden score δ_a , and the access priority score η_a .
Routine checkup recency	CHECKUP1	Time since last routine medical checkup. Longer gaps are treated as higher implementation burden.	Used in the risk index model and in the access burden score δ_a .
Panel E: Lifestyle and behavioral predictors			
Recent exercise	EXERANY2, _TOTINDA	Self-reported exercise in the past 30 days and calculated leisure-time physical activity indicator.	Used in the risk index model to capture physical-activity-related protection or vulnerability.
Smoking exposure	SMOKE100, _SMOKDAY2, _SMOKER3, _RFSMOK3	Lifetime cigarette exposure, current smoking frequency, and calculated smoking status indicators.	Used in the risk index model to capture chronic disease risk and lifestyle-related screening need.
Alcohol use	ALCDAY4, DRNKANY6	Recent alcohol-use frequency and any recent drinking.	Used in the risk index model as auxiliary behavioral indicators.
Panel F: Clinical comorbidity predictors			
Cardiovascular conditions	CVDINFR4, CVDSTRK3, CVDCRHD4	Self-reported history of heart attack, coronary heart disease/angina, or stroke.	Used in the risk index model and in the health impact weight ϕ_a .
Kidney disease	CHCKDNY2	Self-reported kidney disease diagnosis.	Used in the risk index model and in the health impact weight ϕ_a .
Asthma status	ASTHMA3, ASTHNOW	Lifetime and current asthma indicators.	Used in the risk index model to capture chronic disease burden.
Panel G: Anthropometric predictors			
Reported height and weight	WEIGHT2, HEIGHT3	Self-reported height and weight.	Retained primarily for documentation and consistency with BRFSS anthropometric derivations.
Calculated height and weight	WTGK3, HTM4	Calculated weight in kilograms and height in meters.	Auxiliary anthropometric measures for descriptive analysis and sensitivity checks.
Body mass index (BMI)	_BMI5, _BMI5CAT, _RFBMIS	Calculated BMI, BMI category, and overweight/obesity indicator. The continuous BMI variable is rescaled by dividing by 100.	Used in the risk index model and in the health impact weight ϕ_a .
Panel H: Survey design variables			
Sampling weight	_LLCPWT	Final BRFSS survey weight.	Retained for descriptive summaries and possible survey-weighted sensitivity analysis.
Stratification and primary sampling unit	_STSTR, _PSU	Survey design stratification variable and primary sampling unit identifier.	Retained for design-based sensitivity analysis and documentation of the survey structure.
Geographic identifier	_STATE	State Federal Information Processing Standards (FIPS) identifier.	Retained for potential subgroup coverage diagnostics and geographic sensitivity analysis.

*Note: The analysis sample is constructed from the 2024 BRFSS public use file. Variables are grouped according to their role in the fiscal screening design: target construction, survey-based risk prediction, access burden estimation, health impact weighting, and survey design documentation. The main risk index model uses 29 routinely observed predictors.

The primary outcome is constructed from the BRFSS diabetes-status variable DIABETE4. Respondents reporting diagnosed diabetes or prediabetes/borderline diabetes are coded as positive, while respondents reporting no diabetes are coded as negative; missing, refused, or uncertain responses are excluded. Because this is a self-reported status variable, the outcome should not be interpreted as previously undiagnosed diabetes. The main experiment is therefore a retrospective diabetes-related prioritization study rather than a diagnostic screening study for undiscovered disease. To assess dependence on the target definition, we additionally re-estimate the complete model and allocation pipeline for diagnosed-diabetes-only and prediabetes/borderline-only outcomes.

Table 2 summarizes the analysis sample and predictive benchmarks. The positive target rate in the final analysis sample is 16.89%. The main logistic risk index model uses 29 routinely observed survey variables spanning demographics, socioeconomic status, health status, access to care, lifestyle, comorbidity, and body mass index (BMI), and is evaluated on a stratified 30% held-out test set.

To provide a practical nonrandom comparator, we also estimate a transparent five-variable logistic model using age, sex, BMI, recent exercise, and general health. This comparator is intentionally simple and reproducible from the BRFSS. We do not claim that it is a faithful implementation of FINDRISC, Framingham, or another published clinical score because the public use BRFSS file lacks several inputs in the original clinical definitions required for exact reproduction. The comparator is therefore used to distinguish the incremental value of the richer survey-based predictor from the much larger and less informative random baseline contrast.

6.2. Risk index, program gain, and fiscal cost

For each respondent a , the fitted prediction model produces a survey-based health risk index $\rho_a \in [0, 1]$. The risk index is not used directly as the only deployment criterion in all policies. Instead, we embed it in a program gain score that also accounts for health impact heterogeneity, access priority considerations, and fiscal screening cost.

The access burden score δ_a is constructed from indicators of reduced access to routine care, including the inability to afford care, the absence of a personal healthcare provider, a long time since a routine checkup, and a lack of health insurance. The health impact weight ϕ_a increases for respondents with obesity, cardiovascular disease indicators, kidney disease, or fair/poor health. The access priority weight η_a increases for respondents with lower income, lower education, medical cost barriers, or no personal provider. The fiscal screening cost proxy is

$$\kappa_a = 1 + \gamma\delta_a, \quad \gamma = 0.50 \text{ in the baseline specification.} \quad (6.1)$$

The prospective program gain score is

$$\mu_a = \phi_a\rho_a + \Omega\eta_a - \Lambda\delta_a, \quad (6.2)$$

where $\Omega = 0.15$ and $\Lambda = 0.10$ in the main experiment. These coefficients are policy settings used to construct an operational allocation score; they are not estimated causal parameters. To evaluate a sensitivity-adjusted deployment rule, we define the ambiguity proxy

$$\sigma_a = \rho_a(1 - \rho_a), \quad (6.3)$$

and the sensitivity-adjusted score

$$\mu_a^{\text{SA}} = \max\{\mu_a - \zeta\sigma_a, 0\}, \quad \zeta = 0.50 \text{ in the baseline specification.} \quad (6.4)$$

For retrospective evaluation, the realized program gain is computed as

$$\mu_a^{\text{real}} = \phi_a d_a + \Omega \eta_a - \Lambda \delta_a, \quad (6.5)$$

where $d_a \in \{0, 1\}$ denotes the held-out screening target. The prospective score μ_a and the retrospective score μ_a^{real} deliberately share the same policy-defined health impact, access priority, and access burden components; the retrospective score replaces the predicted risk index ρ_a with the observed target d_a . Consequently, μ_a^{real} is an outcome-informed policy evaluation construct rather than an independently observed treatment benefit or causal welfare measure. It is therefore used only to compare selected sets retrospectively and not to claim that the allocation rule estimates clinical intervention effects.

This distinction matters when interpreting the empirical results. Because ρ_a is trained to predict d_a , a risk-informed policy is expected to perform well on a retrospective target-based metric. We therefore separate three questions: Predictive discrimination, downstream allocation value relative to risk ranking, and prospective optimization quality relative to the directly solved binary objective. The fiscal yield versus risk comparison, rather than the random versus risk comparison, is the primary diagnostic for the incremental contribution of downstream allocation.

6.3. Robustness design for policy-defined proxies

To assess dependence on researcher-defined gain, cost, and ambiguity proxies, the revised experiment varies the corresponding policy parameters systematically. Cost heterogeneity is stressed through

$$\gamma \in \{0, 0.25, 0.50, 1, 2, 4\}$$

in $\kappa_a = 1 + \gamma \delta_a$. The access priority and access burden coefficients are varied over

$$\Omega \in \{0, 0.075, 0.15, 0.30\}, \quad \Lambda \in \{0, 0.05, 0.10, 0.20\},$$

and the ambiguity penalty is varied over

$$\zeta \in \{0, 0.25, 0.50, 1\}.$$

These grids are sensitivity ranges, not estimates of structural or causal parameters.

Table 4. Comprehensive empirical parameter settings and sensitivity ranges.

Parameter	Baseline	Sensitivity / scenarios	Setting guidance and role
Ω	0.15	0, 0.075, 0.15, 0.30	Weight on access priority η_a ; larger values give greater allocation weight to the access priority component.
Λ	0.10	0, 0.05, 0.10, 0.20	Penalty on access burden δ_a ; larger values penalize implementation burden more strongly.
γ	0.50	0, 0.25, 0.50, 1, 2, 4	Controls cost heterogeneity in $\kappa_a = 1 + \gamma \delta_a$; larger values create wider fiscal cost dispersion.
ζ	0.50	0, 0.25, 0.50, 1	Penalty on the ambiguity proxy $\sigma_a = \rho_a(1 - \rho_a)$.
$\mathcal{B}$	—	5%, 10%, 20%, 30%	Fiscal budget as a fraction of total held-out fiscal cost.
Bootstrap	—	1000 replicates	Fixed-model respondent sampling uncertainty.
Repeated splits	—	30 refits	Stratified train–test repetitions with full model refitting and policy reconstruction.
Survey weight	Mean-normalized	LLCPWT	Used for weighted value and subgroup coverage sensitivity; not used to claim full complex survey variance inference.

Sampling stability is assessed in two complementary ways. A paired 1000-replicate held-out bootstrap keeps the fitted prediction model and policy rules fixed while resampling respondents. In addition, 30 stratified train–test repetitions refit the complete risk model and reconstruct the downstream allocation policies. The first analysis measures fixed-model respondent sampling uncertainty; the second measures reproducibility across repeated model fitting and sample splitting.

Table 4 consolidates the empirical parameter settings and explains how each parameter is used. These are scenario and policy settings rather than causal quantities estimated from intervention data.

6.4. Survey-weighted coverage and safeguard design

The revised analysis uses the BRFSS final sampling weight `LLCPWT`, normalized to a mean of one within the analysis sample, to evaluate weighted policy value and weighted selection coverage. The BRFSS strata and primary sampling unit identifiers are retained for design documentation, but the present analysis does not claim full design-based variance estimation. Weighted subgroup diagnostics focus on respondents facing a medical cost barrier, lacking a personal provider, having delayed routine checkups, or lacking insurance.

Because the same access burden components can affect both policy value and fiscal cost, cost-aware prioritization may reduce the coverage of respondents with a high access burden. We therefore add a constructive coverage safeguard. The safeguarded fiscal yield rule first restores survey-weighted access burden coverage to at least the level achieved by the risk-ranking benchmark and then fills the remaining budget using fiscal yield. This is a practical coverage aware heuristic rather than an exact fairness-constrained optimizer.

6.5. Deployment policies and evaluation metrics

We compare seven deployment rules under fiscal budgets equal to 5%, 10%, 20%, and 30% of the total test set’s fiscal cost. The random policy assigns a random deployment score. The health risk index policy ranks respondents by ρ_a . The program gain policy ranks respondents by μ_a . The fiscal yield policy ranks respondents by μ_a/κ_a . The sensitivity-adjusted program gain policy ranks respondents by μ_a^{SA} , and the sensitivity-adjusted yield policy ranks respondents by $\mu_a^{\text{SA}}/\kappa_a$. The oracle policy ranks respondents by $\mu_a^{\text{real}}/\kappa_a$ and is used only as an outcome-informed upper bound.

For a policy p , let $S_p(\mathcal{B})$ denote the set of selected respondents under budget $\mathcal{B}$. The realized gain is

$$G_{\mathcal{B}}(p) = \sum_{a \in S_p(\mathcal{B})} \mu_a^{\text{real}}. \quad (6.6)$$

The target capture is

$$\text{Capture}_{\mathcal{B}}(p) = \frac{\sum_{a \in S_p(\mathcal{B})} d_a}{\sum_a d_a}. \quad (6.7)$$

The selected positive rate is

$$\text{PosRate}_{\mathcal{B}}(p) = \frac{\sum_{a \in S_p(\mathcal{B})} d_a}{|S_p(\mathcal{B})|}. \quad (6.8)$$

Finally, the oracle ratio is

$$\text{OracleRecovery}_{\mathcal{B}}(p) = \frac{G_{\mathcal{B}}(p)}{G_{\mathcal{B}}(p^{\text{or}})}, \quad (6.9)$$

where p^{or} denotes the outcome-informed oracle policy. The unrecovered benchmark value is $1 - \text{OracleRecovery}_{\mathcal{B}}(p)$.

Table 5. Score construction, deployment policies, and evaluation mapping in the BRFSS fiscal screening experiment.

Object	Formal representation	Type	Operational interpretation
Panel A: Prospective score construction			
Survey-based risk index	ρ_a	Prospective score	Predicted probability or ranking score produced by the fitted logistic risk index model using routinely observed BRFSS variables.
Access burden score	δ_a	Prospective score	Composite implementation burden measure constructed from medical cost barriers, the absence of a personal provider, a long time since last routine checkup, and the lack of insurance.
Health impact weight	ϕ_a	Prospective weight	Expected public health value multiplier used to increase the importance of screening respondents with obesity, poor health status, cardiovascular disease indicators, or kidney disease.
Access priority weight	η_a	Prospective weight	Priority weight assigned to potentially underserved respondents, based on lower income, lower education, medical cost barriers, or the absence of a personal provider.
Fiscal screening cost	$\kappa_a = 1 + \gamma\delta_a$	Prospective cost	Cost proxy used to model heterogeneous fiscal burden of screening, outreach, counseling, or follow-up.
Program gain score	$\mu_a = \phi_a\rho_a + \Omega\eta_a - \Lambda\delta_a$	Prospective score	Main deployable value score combining risk, expected health impact, access priority, and implementation burden.
Uncertainty proxy	$\sigma_a = \rho_a(1 - \rho_a)$	Prospective score	Simple predictive dispersion proxy used to penalize uncertain screening gains in the sensitivity-adjusted deployment rule.
Sensitivity-adjusted program gain score	$\mu_a^{\text{SA}} = \max\{\mu_a - \zeta\sigma_a, 0\}$	Prospective score	Program gain score adjusted downward for uncertainty in the estimated risk signal.
Fiscal yield score	μ_a / κ_a	Prospective score	Estimated program gain per unit fiscal screening cost. This rule favors respondents who provide stronger expected gains relative to their budgetary burden.
Sensitivity-adjusted fiscal yield score	$\mu_a^{\text{SA}} / \kappa_a$	Prospective score	Sensitivity-adjusted expected gain per unit fiscal screening cost.
Retrospective oracle yield	$\mu_a^{\text{real}} / \kappa_a$	Benchmark score	Outcome-informed score used only for retrospective benchmarking. It is not available in prospective deployment.
Panel B: Deployment policies			
Random allocation	Random score	Benchmark policy	Selects respondents using an independently and identically distributed (i.i.d.) random score. Serves as a minimal baseline under the same fiscal budget.
Health risk index deployment	Rank by ρ_a	Prospective policy	Ranks respondents by the survey-based risk index alone. Measures the value of prediction without explicit fiscal gain refinement.
Program gain deployment	Rank by μ_a	Prospective policy	Ranks respondents by the composite gain score that combines predicted risk, health impact, equity/access priority, and access burden.
Fiscal yield deployment	Rank by μ_a / κ_a	Prospective policy	Ranks respondents by expected gain per unit fiscal cost. This is the main cost-aware deployment rule.
Sensitivity-adjusted gain deployment	Rank by μ_a^{SA}	Prospective policy	Ranks respondents by uncertainty-penalized expected gain.
Sensitivity-adjusted yield deployment	Rank by $\mu_a^{\text{SA}} / \kappa_a$	Prospective policy	Ranks respondents by uncertainty-penalized gain per unit fiscal cost.
Oracle deployment	Rank by $\mu_a^{\text{real}} / \kappa_a$	Retrospective upper bound	Uses realized target information and therefore cannot be implemented prospectively. It is used only to quantify the remaining performance gap of feasible policies.
Panel C: Budget rule and selection mechanism			
Fiscal feasibility condition	$\sum_{a=1}^N \kappa_a y_a \leq \mathcal{B}$	Constraint	A respondent can be selected only if the cumulative fiscal screening cost remains within the available budget.
Budget levels	$\mathcal{B} \in \{0.05, 0.10, 0.20, 0.30\} \times \sum_a \kappa_a$	Experimental design	The experiment evaluates screening performance when the available budget equals 5%, 10%, 20%, and 30% of the total fiscal cost in the test sample.
Greedy deployment rule	Descending score order	Algorithm	For each policy, respondents are sorted by the corresponding score and selected until the cumulative fiscal cost reaches the budget limit.
Panel D: Retrospective evaluation metrics			
Realized gain	$G_{\mathcal{B}}(p)$	Metric	Total realized program gain obtained by the selected set under policy p . This is the main value metric.
Target capture	$\text{Capture}_{\mathcal{B}}(p)$	Metric	Fraction of all positive screening targets in the held-out test set that are selected under budget $\mathcal{B}$.
Selected positive rate	$\text{PosRate}_{\mathcal{B}}(p)$	Metric	Share of selected respondents who are positive screening targets. Measures concentration of the selected set.
Fiscal yield	$\text{Yield}_{\mathcal{B}}(p)$	Metric	Realized program gain per unit fiscal screening cost. Measures deployment efficiency.
Oracle recovery	$\text{OracleRecovery}_{\mathcal{B}}(p)$	Metric	Ratio of realized gain under policy p to realized gain under the oracle benchmark.
Unrecovered value	$\text{UnrecoveredOracleValue}_{\mathcal{B}}(p) = 1 - \text{OracleRecovery}_{\mathcal{B}}(p)$	Metric	Relative shortfall of a feasible deployment rule from the retrospective oracle benchmark.

*Note: This table links score construction, deployment rules, budget feasibility, and retrospective evaluation. The experiment is intentionally structured so that all nonoracle policies use only prospective information. The oracle rule is reported solely as an upper bound for benchmarking.

7. Empirical policy value results

This section reports the empirical performance of the proposed fiscal screening design while separating prediction value, downstream allocation value, and robustness to policy-defined proxies. We first report held-out deployment performance and the directly solved binary benchmark. We then quantify the statistical stability using paired bootstrap resampling and repeated full model refits, stress the assumed cost heterogeneity, and vary the main policy coefficients. The original random allocation and temporal validation analyses are retained as complementary diagnostics, but the revised interpretation places greater weight on risk versus allocation comparisons than on random versus risk differences.

Figure 2 summarizes the empirical workflow used in the BRFSS fiscal screening experiment. The figure separates data construction, risk index estimation, fiscal screening design, and policy evaluation so that the subsequent held-out and temporal validation results can be read as parts of a single screening-design pipeline.

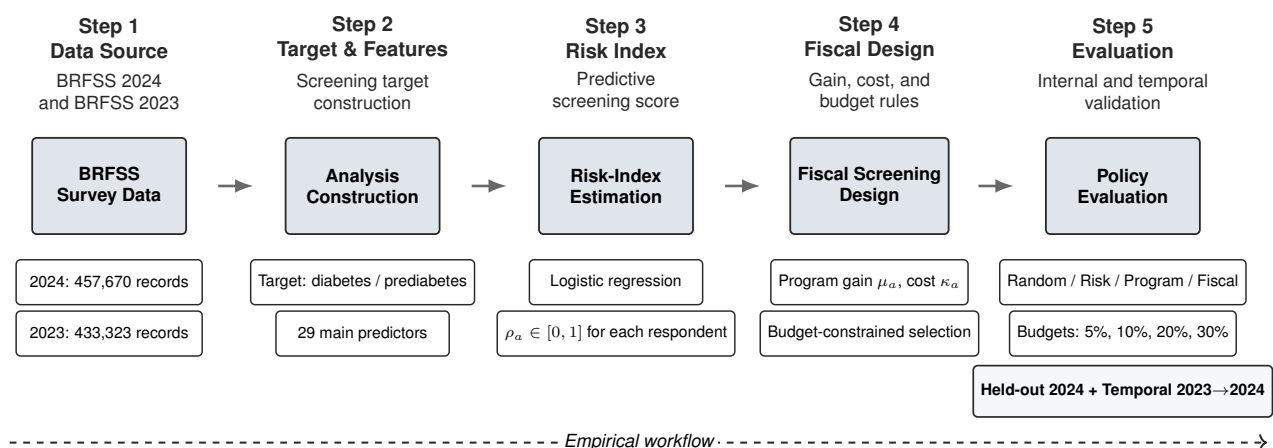


Figure 2. Empirical workflow of the BRFSS-based fiscal screening design. The workflow begins with the BRFSS survey records, constructs a diabetes/prediabetes screening target and analysis features, estimates a respondent-level risk index, constructs fiscal screening scores from predicted risk and program design components, and evaluates budget-constrained deployment policies using both a 2024 held-out experiment and an out-of-year 2023-to-2024 temporal validation.

Figure 3 visualizes the respondent-level structure underlying the fiscal screening policies before reporting aggregate deployment outcomes. Figure 3 shows that the estimated risk score separates realized screening targets from non-targets, although the two groups remain partially overlapping. Figure 3 shows how the risk signal is transformed into a fiscal screening decision space: respondents selected by the fiscal yield rule occupy a high-value region in which predicted risk and sensitivity-adjusted program gain value jointly determine selection.

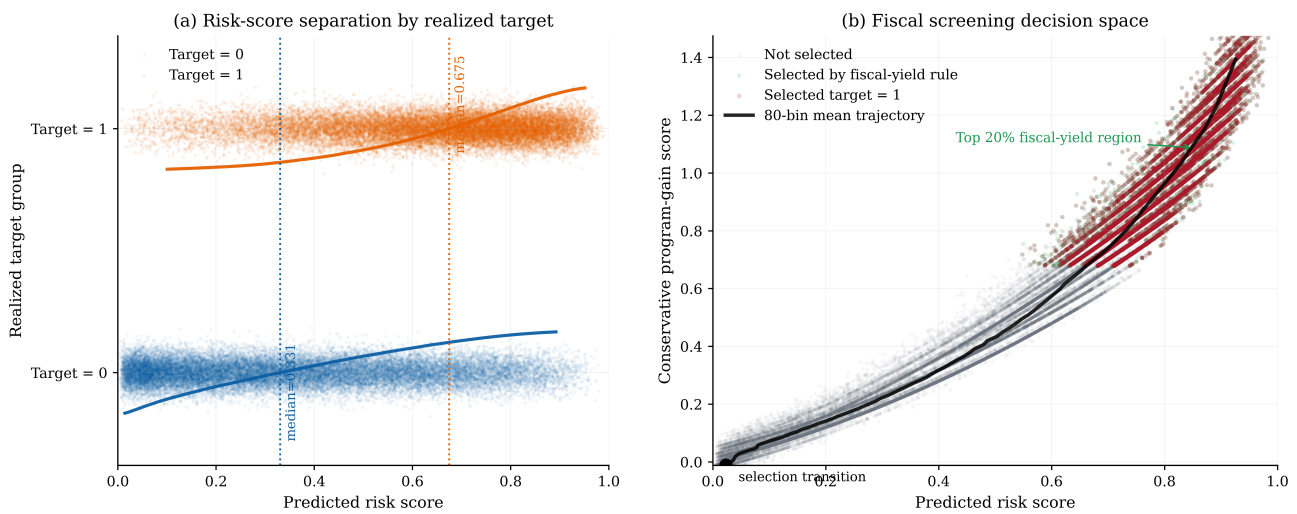


Figure 3. Respondent-level predictive separation and fiscal screening selection structure. Figure 3(a) shows the distribution of predicted risk scores by realized screening target status in the BRFSS 2024 held-out sample. The vertical dotted lines mark the group-specific median risk scores. Figure 3(b) shows the fiscal screening decision space, where each point represents a respondent and the selection region is induced by the sensitivity-adjusted fiscal yield rule under a 20% budget visualization threshold. The figure illustrates how the predictive risk signal is converted into a value-based screening selection rule rather than used as a risk-only ranking.

7.1. Main held-out deployment performance

Table 6 reports the main deployment results on the held-out BRFSS test set. The results show that the survey-based risk index contains substantial prioritization signal. Random allocation performs close to the baseline target rate and captures approximately the same fraction of positive screening targets as the nominal budget fraction. In contrast, risk index deployment captures substantially more positive screening targets and generates much larger realized program gain under every budget level.

At the 10% budget, random allocation captures 9.78% of positive screening targets, whereas the health risk index policy captures 30.17%. The selected positive rate increases from 16.52% under random allocation to 49.75% under risk index deployment. The realized gain also increases from 2,889.69 to 9,686.15. These differences show that the estimated risk index is operationally useful for limited-budget public health screening prioritization.

The fiscal yield and sensitivity-adjusted yield policies provide modest improvements in realized gain and gain per cost relative to risk-only prioritization. For example, at the 10% budget, fiscal yield deployment increases the realized gain from 9,686.15 to 9,773.70 and increases the gain per cost from 0.6763 to 0.6824. At the 30% budget, fiscal yield deployment increases the realized gain from 20,243.94 to 20,329.70 and improves the oracle ratio from 0.7011 to 0.7041. These differences are intentionally not interpreted as evidence that a greedy ranking rule solves the underlying binary optimization problem. The direct 0–1 benchmark reported below is used to evaluate the optimization quality separately from any retrospective policy value.

Table 6. BRFSS 2024 fiscal screening deployment performance on the held-out test set.

Budget	Policy	Selected	Realized gain	Target capture	Selected positive	Gain/cost	Oracle recovery	Unrecovered value
5%	Random	6,854	1,506.60	0.0507	0.1711	0.2104	0.1535	0.8465
5%	Health risk index	7,002	5,610.89	0.1685	0.5566	0.7836	0.5718	0.4282
5%	Program gain	7,000	5,668.85	0.1672	0.5526	0.7916	0.5777	0.4223
5%	Fiscal yield	7,097	5,683.86	0.1682	0.5484	0.7937	0.5792	0.4208
5%	Sensitivity-adjusted gain	6,999	5,672.79	0.1677	0.5542	0.7922	0.5781	0.4219
5%	Sensitivity-adjusted yield	7,085	5,684.26	0.1684	0.5498	0.7937	0.5792	0.4208
5%	Oracle	7,112	9,813.23	0.3074	1.0000	1.3704	1.0000	0.0000
10%	Random	13,697	2,889.69	0.0978	0.1652	0.2018	0.1573	0.8427
10%	Health risk index	14,031	9,686.15	0.3017	0.4975	0.6763	0.5272	0.4728
10%	Program gain	14,018	9,766.44	0.2996	0.4944	0.6819	0.5316	0.4684
10%	Fiscal yield	14,152	9,773.70	0.3006	0.4914	0.6824	0.5320	0.4680
10%	Sensitivity-adjusted gain	14,019	9,751.69	0.2995	0.4942	0.6809	0.5308	0.4692
10%	Sensitivity-adjusted yield	14,131	9,803.81	0.3019	0.4942	0.6845	0.5336	0.4664
10%	Oracle	14,176	18,372.01	0.6128	1.0000	1.2827	1.0000	0.0000
20%	Random	27,398	5,872.15	0.1994	0.1683	0.2050	0.2077	0.7923
20%	Health risk index	28,105	15,846.49	0.5147	0.4236	0.5532	0.5606	0.4394
20%	Program gain	28,051	15,912.04	0.5107	0.4212	0.5555	0.5629	0.4371
20%	Fiscal yield	28,253	15,931.25	0.5124	0.4196	0.5562	0.5636	0.4364
20%	Sensitivity-adjusted gain	28,055	15,922.11	0.5116	0.4219	0.5558	0.5632	0.4368
20%	Sensitivity-adjusted yield	28,219	15,933.30	0.5128	0.4204	0.5562	0.5636	0.4364
20%	Oracle	27,927	28,269.35	1.0000	0.8284	0.9869	1.0000	0.0000
30%	Random	41,096	8,829.95	0.2997	0.1687	0.2055	0.3058	0.6942
30%	Health risk index	42,169	20,243.94	0.6727	0.3691	0.4711	0.7011	0.2989
30%	Program gain	42,074	20,308.06	0.6678	0.3672	0.4726	0.7033	0.2967
30%	Fiscal yield	42,302	20,329.70	0.6699	0.3664	0.4731	0.7041	0.2959
30%	Sensitivity-adjusted gain	42,077	20,310.30	0.6682	0.3674	0.4727	0.7034	0.2966
30%	Sensitivity-adjusted yield	42,257	20,325.34	0.6698	0.3667	0.4730	0.7039	0.2961
30%	Oracle	41,373	28,873.69	1.0000	0.5592	0.6720	1.0000	0.0000

*Notes: The held-out test set contains 136,988 BRFSS respondents. “Target capture” is the fraction of all positive screening targets selected under the corresponding budget. “Selected positive” is the positive screening target rate among selected respondents. “Gain/cost” is realized program gain divided by fiscal screening cost. “Oracle” is an outcome-informed upper bound that uses the realized screening target and is not implementable in prospective deployment. All nonoracle policies use only prospective scores.

Table 7. Direct 0–1 binary optimization benchmark in the BRFSS 2024 held-out experiment.

Budget	Binary selected	Solver certification	MIP gap	FY prospective gap	Binary retrospective value
5%	7,097	Optimal	0	0.0015%	5,683.88
10%	14,152	Optimal	0	0.0007%	9,774.99
20%	28,253	Near-optimal	0.0024%	0.0000% ^a	15,931.25
30%	42,303	Optimal	0	0.0008%	20,330.95

*Notes: The binary model maximizes the prospective program value objective under the same fiscal budget as the ranking rules. “FY prospective gap” is the percentage loss of fiscal yield ranking relative to the reported binary benchmark objective. ^aAt 20%, fiscal yield matches the reported incumbent, while the incumbent itself has a certified relative MIP gap of approximately 0.0024%; therefore this row is not described as an exact optimum.

The revised analysis additionally solves the prospective 0–1 budget-constrained problem in (5.16) rather than relying only on the greedy ranking. Table 7 reports the solver certification and the prospective objective gap of fiscal yield ranking. At the 5%, 10%, and 30% budgets, the binary benchmark was

solved with zero reported MIP gap. At the 20% budget, the solver returned a certified near-optimal incumbent with a relative gap of approximately 0.0024%. The fiscal yield ranking is extremely close to these binary benchmarks under the baseline cost specification, but this proximity is reported as an empirical finding and not as a general optimality claim.

7.2. Statistical stability and proxy sensitivity

The baseline fiscal yield improvement over risk ranking is small in magnitude, so we explicitly assess whether its direction is stable under resampling. Table 8 reports both a paired held-out bootstrap and 30 repeated train–test splits with full model refitting. The paired bootstrap is deliberately conservative: Its 95% interval excludes zero only at the 30% budget, whereas the lower-budget intervals include small negative values. In contrast, the 30 repeated full refits yield positive empirical 2.5th–97.5th percentile intervals at all four budgets. We therefore characterize the baseline allocation increment as modest and directionally reproducible across model refits, not as uniformly statistically significant.

Table 8. Stability and parameter sensitivity of fiscal yield improvement over risk ranking.

Budget	Held out increase	95% bootstrap interval	30-split mean [empirical interval]	Ω – Λ increase range
5%	1.301%	[-0.333%, 2.988%]	1.384% [0.594%, 2.198%]	0.98–2.42%
10%	0.904%	[-0.136%, 1.963%]	0.766% [0.166%, 1.388%]	0.86–1.88%
20%	0.535%	[-0.072%, 1.088%]	0.606% [0.161%, 0.920%]	0.38–1.11%
30%	0.424%	[0.007%, 0.826%]	0.456% [0.121%, 0.698%]	0.26–0.82%

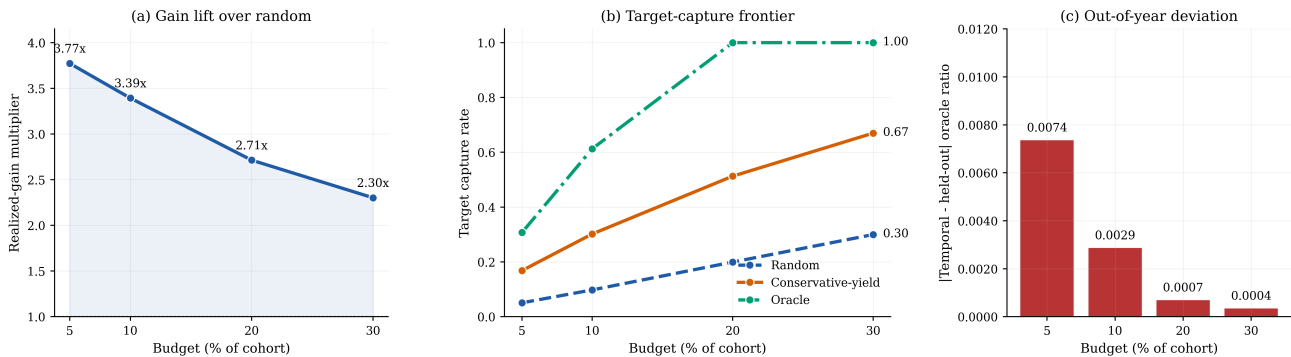
*Notes: “Held out increase” is the retrospective fiscal yield improvement over the full risk-ranking policy under the baseline coefficients. The bootstrap keeps the fitted model and policies fixed and resamples held-out respondents 1000 times. The 30-split analysis refits the model and reconstructs policies in each repetition; its interval is the empirical 2.5th–97.5th percentile range, not a design-based confidence interval. The last column reports the minimum and maximum increase over the 16 combinations of $\Omega \in \{0, 0.075, 0.15, 0.30\}$ and $\Lambda \in \{0, 0.05, 0.10, 0.20\}$.

Cost heterogeneity is the condition under which the downstream allocation layer becomes materially more important. At the 10% budget, the fiscal yield increase over risk ranking is approximately 0.87% when $\gamma = 0$, 1.09% at $\gamma = 0.25$, 0.90% at the baseline $\gamma = 0.50$, 1.09% at $\gamma = 1$, 2.59% at $\gamma = 2$, and 7.43% at $\gamma = 4$. At the 5% budget, the increase reaches 11.53% in the strongest $\gamma = 4$ stress case. The high- γ cases are hypothetical stress tests rather than observed program cost estimates. Their purpose is to show that the small baseline allocation effect is partly a consequence of the narrow baseline cost range and that the value of cost-aware allocation increases as cost dispersion becomes more consequential.

The parameter analysis leads to the same qualitative conclusion. Across all 16 (Ω , Λ) combinations, fiscal yield prioritization produces positive realized gain improvements over risk ranking at every reported budget. The ranges are 0.98–2.42% at the 5% budget, 0.86–1.88% at 10%, 0.38–1.11% at 20%, and 0.26–0.82% at 30%. Varying $\zeta \in \{0, 0.25, 0.50, 1\}$ likewise preserves a positive sensitivity-adjusted yield improvement at every budget; at the 10% budget the increase is approximately 0.90–1.42%. Across these specifications, the direction of the fiscal yield comparison is stable, but the coefficient values remain policy scenarios rather than empirically identified causal quantities.

Figure 4 visualizes the policy value pattern implied by Table 6. Because random allocation is only a lower-bound sanity check, the revised interpretation does not use the random-to-policy gap as evidence for the incremental value of downstream allocation. Instead, the practical comparator, full risk ranking, fiscal yield rule, and binary benchmark reported elsewhere in this section provide the relevant

decomposition of prediction value and allocation value. The temporal panel is retained as a descriptive transportability diagnostic.



Panels are computed directly from saved CSV experiment outputs. Higher values are better in panels (a) and (b); lower values are better in panel (c).

Figure 4. Policy value and temporal stability of fiscal screening deployment. Figure 4(a) reports the realized gain multiplier of the sensitivity-adjusted yield policy relative to random allocation in the BRFSS 2024 held-out experiment. Figure 4(b) shows the target capture frontier under random, sensitivity-adjusted yield, and oracle allocation. Figure 4(c) reports the absolute out-of-year oracle ratio deviation between the 2024 held-out experiment and the 2023-to-2024 temporal validation. All panels are computed directly from saved CSV outputs from the experiment.

7.3. Random allocation sanity check

For completeness, Table 9 retains the original random allocation comparison as a lower-bound sanity check under the same fiscal budgets. Random allocation is not presented as a realistic agency policy or as the primary comparator for downstream optimization. The practical five-variable comparator and the full risk ranking are the relevant nonrandom benchmarks. This table therefore serves only to illustrate the amount of predictive signal captured by risk-informed policies relative to an intentionally weak baseline; the incremental contribution of downstream allocation is evaluated by fiscal yield versus full risk comparisons and by the directly solved binary benchmark.

Figure 5 retains the same random baseline visualization for continuity with the original submission. The multipliers in Figure 5 summarize the gap between a value-based policy and the intentionally weak random baseline; they should not be interpreted as the causal effect of an intervention or as the incremental contribution of the allocation layer. The latter is evaluated by the fiscal yield versus risk and binary benchmark analyses above.

The random allocation diagnostic confirms that the fitted risk signal contains substantial prioritization information but it does not identify the incremental value of downstream allocation. That incremental value is intentionally assessed elsewhere in this section by comparing fiscal yield with the full risk ranking and with the prospective binary benchmark. Accordingly, the large random-to-policy multipliers are treated as descriptive lower-bound contrasts rather than as evidence that optimization itself generates several-fold gains.

Table 9. Incremental deployment increase relative to random allocation in the BRFSS 2024 held-out experiment.

Budget	Policy	Gain increase	Gain multiplier	Capture increase	Gain/cost increase	Oracle ratio increase	Recovery improvement
5%	Health risk index	4,104.29	3.7242	0.1178	0.5732	0.4183	0.4183
5%	Fiscal yield	4,177.26	3.7726	0.1175	0.5833	0.4257	0.4257
5%	Sensitivity-adjusted yield	4,177.66	3.7729	0.1177	0.5833	0.4257	0.4257
10%	Health risk index	6,796.46	3.3520	0.2039	0.4745	0.3699	0.3699
10%	Fiscal yield	6,884.01	3.3823	0.2028	0.4806	0.3747	0.3747
10%	Sensitivity-adjusted yield	6,914.12	3.3927	0.2041	0.4827	0.3763	0.3763
20%	Health risk index	9,974.34	2.6986	0.3153	0.3482	0.3529	0.3529
20%	Fiscal yield	10,059.10	2.7130	0.3130	0.3512	0.3559	0.3559
20%	Sensitivity-adjusted yield	10,061.15	2.7134	0.3134	0.3512	0.3559	0.3559
30%	Health risk index	11,413.99	2.2926	0.3730	0.2656	0.3953	0.3953
30%	Fiscal yield	11,499.75	2.3024	0.3702	0.2676	0.3983	0.3983
30%	Sensitivity-adjusted yield	11,495.39	2.3019	0.3701	0.2675	0.3981	0.3981

*Notes: “Gain increase” is the increase in realized program gain relative to random allocation under the same fiscal budget. “Gain multiplier” is the ratio of realized gain under the policy to realized gain under random allocation. “Capture increase” is the increase in target capture relative to random allocation. “Recovery improvement” is the reduction in retrospective oracle recovery gap relative to random allocation. The table is derived from the same held-out BRFSS experiment as Table 6.

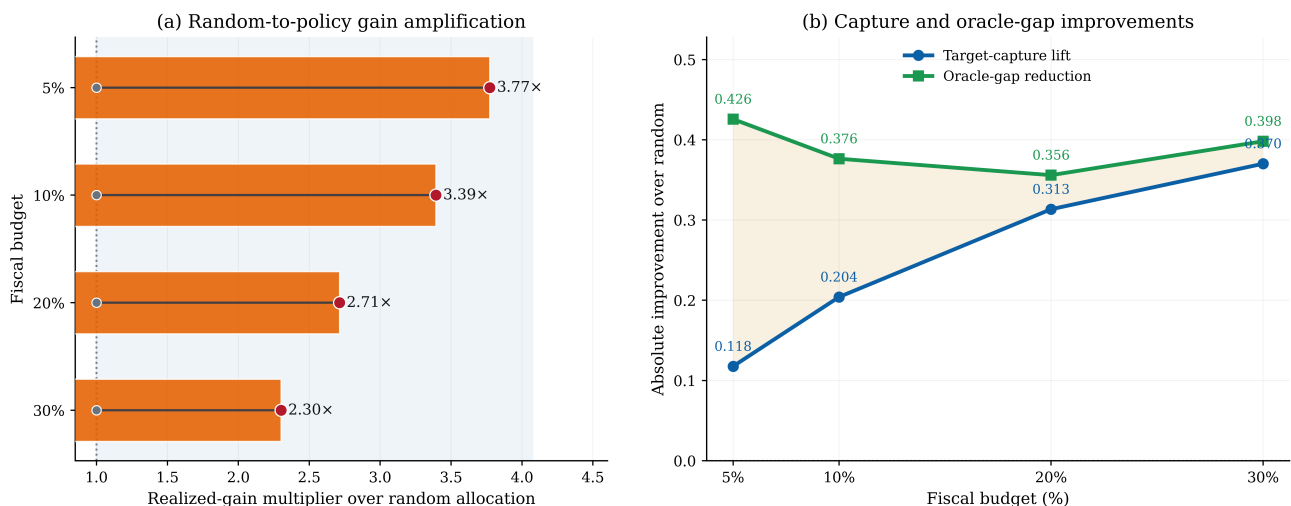


Figure 5. Random allocation sanity check for the sensitivity-adjusted yield rule in the BRFSS 2024 held-out experiment. Figure 5(a) reports the realized value multiplier over random allocation across fiscal budgets. Figure 5(b) reports the corresponding changes in target capture and retrospective oracle recovery. Random allocation is retained only as a lower-bound diagnostic; the practical comparator, full risk ranking, and binary benchmark are the primary references for evaluating the prediction and downstream allocation value.

7.4. Practical comparator, target sensitivity, and weighted coverage

The five-variable comparator provides a practical benchmark beyond random allocation. Its held-out AUC and average precision are 0.7781 and 0.3848, compared with 0.8002 and 0.4331 for the full

29-variable model. Under common fiscal budgets, the full risk model improves retrospective allocation value relative to the five-variable comparator by 16.84%, 13.01%, 8.91%, and 5.76% at the 5%, 10%, 20%, and 30% budgets, respectively. These differences quantify the prediction value separately from the smaller fiscal yield versus full risk allocation increment.

Table 10. Practical comparator and target sensitivity results.

Budget	5-var. comparator	Full risk	FY vs. full risk	Prediabetes-only FY vs. risk
5%	4,802.38	5,610.89	1.30%	-3.01%
10%	8,570.93	9,686.15	0.90%	3.20%
20%	14,549.44	15,846.49	0.53%	1.23%
30%	19,140.64	20,243.94	0.42%	1.53%

*Notes: Values for the comparator and full risk columns are retrospective allocation values under common budgets. The diagnosed-diabetes-only model has AUC = 0.8083 and average precision = 0.3980; fiscal yield increase over risk is positive at all four budgets (1.96%, 0.62%, 0.80%, and 0.42%). The prediabetes/borderline-only model has AUC = 0.6902 and average precision = 0.0490, and its fiscal yield increase is negative at the 5% budget. The target sensitivity results therefore do not support a universal allocation advantage across outcomes and budgets.

Table 11. Survey-weighted value and access burden coverage.

Budget	FY vs. risk value	Risk coverage	FY coverage	Safeguard vs. risk value	Safeguard coverage
5%	-2.80%	0.0279	0.0096	1.09%	0.0279
10%	-2.94%	0.0494	0.0301	-1.98%	0.0494
20%	0.30%	0.0899	0.0682	1.31%	0.0899
30%	1.10%	0.1362	0.1196	1.46%	0.1362

*Notes: Values and coverage use normalized BRFSS sampling weights. “Coverage” is weighted selection coverage among respondents with the access burden indicator used by the safeguard. The safeguarded rule is constructive and coverage aware; it is not claimed to solve a new fairness optimization problem.

The target sensitivity analysis materially narrows the interpretation of the paper. The combined self-reported target is not evidence of detecting previously undiagnosed disease. The diagnosed-only model remains predictive and shows a positive fiscal yield increase at all budgets, whereas the prediabetes/borderline-only outcome is substantially harder to predict and yields a negative fiscal yield effect at the 5% budget. We therefore interpret the BRFSS study as retrospective diabetes-related prioritization rather than diagnostic screening.

Survey weighting reveals a separate efficiency–coverage trade-off. Table 11 reports weighted realized value differences and weighted coverage of respondents with the access burden indicator. Unconstrained fiscal yield reduces access burden coverage relative to risk ranking at every budget. Its weighted realized value is lower than risk ranking at the 5% and 10% budgets and slightly higher at the 20% and 30% budgets. The coverage safeguard restores burden coverage to the risk policy benchmark but does not uniformly dominate risk ranking in weighted value.

At the 10% budget, subgroup diagnostics show the same direction for specific access barriers: fiscal yield prioritization selects lower weighted shares of respondents reporting a medical cost barrier, no personal provider, a delayed routine checkup, or no insurance. These findings directly address the possibility that using the access burden simultaneously in value and cost can deprioritize underserved respondents. The revised manuscript therefore reports coverage consequences explicitly rather than assuming that a cost-aware score is distributionally neutral.

7.5. Out-of-year temporal validation

To examine whether the proposed screening design is stable beyond a single random split of the 2024 survey year, we conduct an out-of-year temporal validation experiment. The risk index model is trained on the 2023 BRFSS public use file and then applied without refitting to the 2024 BRFSS cohort. The fiscal deployment rules, access burden construction, health impact weight, access priority weight, fiscal cost proxy, and budget levels are kept identical to the main experiment.

The temporal validation is designed to test transportability across adjacent survey years rather than only predictive performance within a single year. After target construction and filtering samples with missing features, the 2023 training sample contains 432,339 respondents, and the 2024 temporal validation sample contains 456,625 respondents. The positive screening target rate is 16.28% in the 2023 training sample and 16.89% in the 2024 validation sample. Because one access-related calculated variable used in the 2024 analysis was not available in the 2023 public use file, the temporal validation uses the 28 predictors common to both years.

The 2023-trained logistic risk index model achieves an AUC of 0.8022 and an average precision of 0.4352 when evaluated on 2024 respondents. The average precision is substantially above the 2024 baseline positive rate of 0.1689, indicating that the survey-based risk index remains informative when transported across survey years. This result is important because it shows that the main screening signal does not depend only on a random within-year split.

Figure 6 visualizes the stability of the sensitivity-adjusted yield deployment rule across the internal held-out experiment and the out-of-year temporal validation. Figure 6 compares the oracle ratio trajectory across budget levels, while Figure 6 reports absolute deviations between the two validation designs. The small deviations show that the deployment pattern is not driven by a single random split of the 2024 BRFSS data.

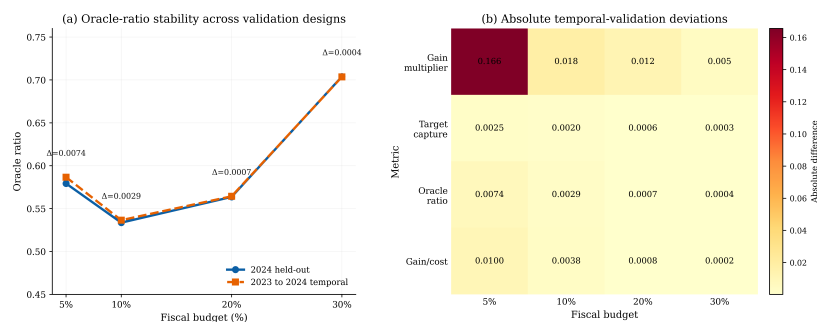


Figure 6. Out-of-year temporal stability of sensitivity-adjusted yield deployment. Figure 6(a) compares oracle ratio trajectories/ Figure 6(b) reports absolute deviations across the two validation designs.

Table 12 reports the temporal deployment results. The pattern is consistent with the main 2024 held-out experiment. Risk-informed deployment substantially improves over random allocation under every fiscal budget. At the 10% budget, random allocation achieves a realized gain of 9,737.91 and captures 9.89% of positive screening targets. The health risk index policy increases realized gain to 32,548.24 and captures 30.42% of positive screening targets. Fiscal yield and sensitivity-adjusted yield deployment further increase realized gain to 32,818.40 and 32,865.81, respectively.

Table 12. Out-of-year temporal validation from BRFSS 2023 to BRFSS 2024.

Budget	Policy	Selected	Realized gain	Target capture	Selected positive	Gain/cost	Oracle recovery	Unrecovered value
5%	Random	22,850	4,871.89	0.0496	0.1675	0.2041	0.1489	0.8511
5%	Health risk index	23,353	18,876.03	0.1703	0.5624	0.7907	0.5771	0.4229
5%	Program gain	23,328	19,083.96	0.1691	0.5589	0.7994	0.5834	0.4166
5%	Fiscal yield	23,652	19,145.90	0.1703	0.5553	0.8020	0.5853	0.4147
5%	Sensitivity-adjusted gain	23,330	19,099.41	0.1695	0.5604	0.8000	0.5839	0.4161
5%	Sensitivity-adjusted yield	23,610	19,188.11	0.1709	0.5580	0.8037	0.5866	0.4134
5%	Oracle	23,712	32,709.46	0.3075	1.0000	1.3701	1.0000	0.0000
10%	Random	45,670	9,737.91	0.0989	0.1671	0.2039	0.1590	0.8410
10%	Health risk index	46,780	32,548.24	0.3042	0.5014	0.6817	0.5313	0.4687
10%	Program gain	46,722	32,786.10	0.3020	0.4985	0.6867	0.5352	0.4648
10%	Fiscal yield	47,170	32,818.40	0.3032	0.4957	0.6873	0.5357	0.4643
10%	Sensitivity-adjusted gain	46,729	32,791.45	0.3025	0.4993	0.6868	0.5353	0.4647
10%	Sensitivity-adjusted yield	47,102	32,865.81	0.3039	0.4975	0.6883	0.5365	0.4635
10%	Oracle	47,280	61,258.18	0.6131	1.0000	1.2830	1.0000	0.0000
20%	Random	91,334	19,690.90	0.2002	0.1690	0.2062	0.2089	0.7911
20%	Health risk index	93,709	52,895.28	0.5151	0.4239	0.5539	0.5612	0.4388
20%	Program gain	93,498	53,119.21	0.5113	0.4217	0.5563	0.5636	0.4364
20%	Fiscal yield	94,167	53,131.34	0.5125	0.4197	0.5564	0.5637	0.4363
20%	Sensitivity-adjusted gain	93,516	53,157.36	0.5123	0.4224	0.5566	0.5640	0.4360
20%	Sensitivity-adjusted yield	94,054	53,191.59	0.5134	0.4209	0.5570	0.5643	0.4357
20%	Oracle	93,118	94,254.86	1.0000	0.8281	0.9870	1.0000	0.0000
30%	Random	136,966	29,492.35	0.3001	0.1690	0.2059	0.3064	0.6936
30%	Health risk index	140,586	67,412.18	0.6721	0.3686	0.4706	0.7003	0.2997
30%	Program gain	140,254	67,710.83	0.6679	0.3672	0.4727	0.7034	0.2966
30%	Fiscal yield	141,020	67,673.86	0.6688	0.3657	0.4724	0.7030	0.2970
30%	Sensitivity-adjusted gain	140,266	67,711.96	0.6683	0.3674	0.4727	0.7034	0.2966
30%	Sensitivity-adjusted yield	140,876	67,729.79	0.6695	0.3665	0.4728	0.7036	0.2964
30%	Oracle	137,916	96,265.61	1.0000	0.5591	0.6720	1.0000	0.0000

*Notes: The risk index model is trained on BRFSS 2023 and evaluated on BRFSS 2024 without refitting. “Selected” is the number of respondents selected under the fiscal budget. “Target capture” is the fraction of all positive screening targets selected. “Selected positive” is the positive screening target rate among selected respondents. “Gain/cost” is realized program gain divided by fiscal screening cost. The oracle policy uses realized target information and is reported only as a retrospective upper bound.

The temporal validation strengthens the empirical evidence in two ways. First, the risk index model remains predictive when trained on the previous survey year and transported to the following year. Second, the fiscal screening policies preserve the same qualitative ranking observed in the main experiment: random allocation performs near the population baseline, risk index deployment creates the dominant gain, and program gain, fiscal yield, and sensitivity-adjusted yield rules provide modest but consistent refinements. These findings support the paper’s central claim that public health screening value depends not only on estimating risk but also on converting risk signals into fiscally feasible deployment decisions.

7.6. Oracle recovery and policy value interpretation

The oracle benchmark should be interpreted only as an outcome-informed upper bound. It is not a feasible prospective policy because it ranks the respondents using realized target information. The oracle ratio therefore measures how much of the outcome-informed upper-bound gain is recovered by implementable deployment rules. In the main 2024 held-out experiment, the best implementable policies recover approximately 0.58, 0.53, 0.56, and 0.70 of oracle realized gain at the 5%, 10%, 20%, and 30% budgets, respectively. The lower oracle ratios at the smaller budgets reflect the difficulty of

identifying the highest-value screening targets when only a small fraction of the population can be selected.

The temporal validation results lead to the same interpretation. When trained on BRFSS 2023 and transported to BRFSS 2024, the implementable policies recover approximately 0.59, 0.54, 0.56, and 0.70 of oracle-realized gain at the same budget levels. This similarity between the internal held-out experiment and the out-of-year validation suggests that the observed deployment pattern is not driven by a single random split of the 2024 data.

Overall, the BRFSS experiment supports the paper's central argument. Predictive accuracy alone is not the relevant operational objective; the practical value of a public health risk index depends on how it is converted into screening deployment under fiscal constraints. Risk-informed deployment substantially outperforms random allocation, while the program gain and fiscal yield rules provide a disciplined way to incorporate heterogeneous health impact, access priority, and screening cost into the allocation decision. The temporal validation further shows that the proposed design remains stable when the risk index model is trained on a previous survey year and applied to the following survey year without refitting.

8. Discussion

The revised results support four conclusions about the use of health AI risk scores for constrained public health allocation.

First, prediction provides most of the first-order prioritization value in this BRFSS setting. The full 29-variable risk model yields higher predictive and allocation performance than the transparent five-variable comparator and random allocation. This finding is consistent with the broader distinction between prediction and decision value emphasized in clinical prediction and decision-focused work [1, 35, 38]. At the same time, predictive discrimination alone does not determine which respondents should be selected under heterogeneous costs and policy priorities.

Second, downstream allocation can add value, but its incremental contribution should not be overstated. Under the baseline cost specification, fiscal yield prioritization improves the retrospective value over the full risk ranking by less than about 1.4% across the four budget levels and nearly matches the prospective binary optimization benchmark. The directly solved 0–1 model therefore serves as an optimization benchmark rather than as evidence that the heuristic is universally optimal. The cost–stress analysis shows why the baseline increment is small: When fiscal costs are only mildly heterogeneous, risk and fiscal yield rankings remain similar. As γ increases and cost dispersion widens, the value of cost-aware allocation becomes materially larger.

Third, the baseline allocation increment is modest but directionally reproducible. The paired 1000-replicate held-out bootstrap produces intervals that include zero at the 5%, 10%, and 20% budgets and exclude zero only at the 30% budget. By contrast, 30 repeated train–test splits with complete model refitting yield positive empirical percentile intervals at all four budgets. We therefore avoid a blanket significance claim and instead describe the result as a modest allocation refinement with a stable direction across repeated refits. The broad (Ω, Λ) , γ , and ζ sensitivity analyses further show that the magnitude depends on policy settings.

Fourth, efficiency and coverage need not move in the same direction. The survey-weighted analysis shows that unconstrained fiscal yield prioritization can reduce the weighted coverage of respondents facing access barriers and can even lower weighted retrospective value relative to risk ranking at lower budgets. This is directly relevant to concerns raised in population health algorithm research and GUIDE-style equity guidance [41,42]. The constructive coverage safeguard demonstrates one practical response: preserve at least the risk policy level of weighted access burden coverage before allocating the remaining budget by fiscal yield. The safeguard is not presented as an exact fairness optimizer; it is a transparent policy mechanism for exposing and managing the efficiency–coverage trade-off.

The target definition also materially limits interpretation. The BRFSS outcome is self-reported diabetes/prediabetes-related status, not laboratory-confirmed incident disease and not an undiagnosed-disease endpoint. The diagnosed-only sensitivity analysis remains predictive and shows a positive fiscal yield increase at all budgets, whereas the prediabetes-only analysis is substantially weaker and yields a negative allocation increment at the 5% budget. The empirical study should therefore be read as retrospective prioritization rather than diagnostic screening validation.

More broadly, the framework is a decision support and resource allocation model, not a causal intervention model. The realized program value metric reuses the same policy-defined health impact, access priority, and access burden components as the prospective score and replaces predicted risk with the observed target. It is therefore an outcome-informed evaluation construct rather than an independently observed treatment benefit. Likewise, Ω , Λ , γ , and ζ are transparent policy settings rather than causally identified coefficients. Future work should calibrate program value and cost using linked administrative or intervention data, implement full complex survey inference where appropriate, and evaluate subgroup validity and external transportability beyond adjacent BRFSS years.

9. Conclusion

This paper examined how health AI risk scores can be translated into budget-feasible public health resource allocation decisions. The central empirical finding is a separation between prediction value and downstream allocation value. In the BRFSS 2024 experiment, the richer 29-variable risk model provides most of the first-order prioritization improvement relative to a transparent five-variable comparator. Fiscal yield allocation then provides a smaller additional refinement under the baseline cost specification.

The revised optimization analysis makes the scope of that refinement explicit. Fiscal yield ranking is a scalable heuristic for the binary budget-constrained problem, not a new exact knapsack algorithm. The directly solved 0–1 benchmark shows that fiscal yield prioritization nearly matches the prospective binary benchmark in this setting, while cost–stress experiments show that the value of cost-aware allocation increases as fiscal cost heterogeneity grows. Bootstrap and repeated refit analyses indicate that the baseline increment is modest in magnitude but directionally reproducible across model refits.

The empirical interpretation is deliberately limited. The BRFSS target is retrospective and self-reported, so the study does not identify previously undiagnosed diabetes or estimate the causal benefit of an intervention. Survey-weighted results also show that cost-aware allocation can reduce the coverage of respondents facing access barriers, motivating explicit coverage safeguards rather than assuming that efficiency-oriented rankings are distributionally neutral.

Taken together, the results support a decision-focused view of health AI deployment: Accurate prediction is necessary but not sufficient for resource allocation, and the incremental value of downstream

optimization depends on the cost heterogeneity, policy parameters, and coverage objectives. The proposed framework provides a transparent way to evaluate these components separately under common fiscal constraints.

Use of generative AI tools declaration

Generative artificial intelligence (AI) tools were used only for limited language polishing and grammatical refinement. The authors independently developed the research design, mathematical formulation, data analysis, computational experiments, interpretation of results, and final manuscript content. The authors reviewed all AI-assisted edits and take full responsibility for the published work.

Author contributions

Jinho Cha: Conceptualization, methodology, mathematical formulation, software, data curation, formal analysis, computational experiments, visualization, writing – original draft, and writing – review and editing; Eun-Young Kwon: Conceptualization, supervision, validation, interpretation of results, writing – review and editing, project administration, and corresponding author responsibilities. All authors have read and approved the final manuscript.

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, Human decisions and machine predictions, *Q. J. Econ.*, **133** (2018), 237–293. <https://doi.org/10.1093/qje/qjx032>
2. A. N. Elmachtoub, P. Grigas, Smart “predict, then optimize”, *Manag. Sci.*, **68** (2022), 9–26. <https://doi.org/10.1287/mnsc.2020.3922>
3. D. Bertsimas, N. Kallus, From predictive to prescriptive analytics, *Manag. Sci.*, **66** (2020), 1025–1044. <https://doi.org/10.1287/mnsc.2018.3253>
4. B. Wilder, B. Dilkina, M. Tambe, Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization, In: *Proceedings of the AAAI Conference on Artificial Intelligence*, **33** (2019), 1658–1665. <https://doi.org/10.1609/aaai.v33i01.33011658>
5. P. L. Donti, B. Amos, J. Z. Kolter, Task-based end-to-end model learning in stochastic optimization, In: *Advances in Neural Information Processing Systems*, **30** (2017), 5484–5494. Available from: <https://papers.neurips.cc/paper/7132-task-based-end-to-end-model-learning-in-stochastic-optimization>.
6. T. Tulabandhula, C. Rudin, Machine learning with operational costs, *J. Mach. Learn. Res.*, **14** (2013), 1989–2028. Available from: <https://jmlr.org/papers/v14/tulabandhula13a.html>.
7. Y. Deng, T. C. Y. Chan, E. Delage, L. Yin, Robust and data-driven optimization in healthcare operations, *INFORMS J. Comput.*, **30** (2018), 373–394. <https://doi.org/10.1287/ijoc.2017.0778>

8. S. Martello, P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*, John Wiley & Sons, Chichester, UK, 1990.
9. H. Kellerer, U. Pferschy, D. Pisinger, *Knapsack Problems*, Springer, Berlin, Heidelberg, 2004. <https://doi.org/10.1007/978-3-540-24777-7>
10. D. Pisinger, P. Toth, Knapsack problems, In: *Handbook of Combinatorial Optimization*, Kluwer Academic Publishers, Boston, MA, 1998, 299–428. https://doi.org/10.1007/978-1-4613-0303-9_5
11. Ding-Zhu Du, Panos M. Pardalos, *Handbook of Combinatorial Optimization*, Kluwer Academic Publishers, Boston, MA, 1998. <https://doi.org/10.1007/978-1-4613-0303-9>
12. G. L. Nemhauser, L. A. Wolsey, *Integer and Combinatorial Optimization*, John Wiley & Sons, New York, 1988. <https://doi.org/10.1002/9781118627372>
13. L. A. Wolsey, *Integer Programming*, John Wiley & Sons, New York, 1998.
14. A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency*, Springer, Berlin, Heidelberg, 2003.
15. R. M. Karp, Reducibility among combinatorial problems, In: *Complexity of Computer Computations*, Springer, Boston, MA, 1972, 85–103. https://doi.org/10.1007/978-1-4684-2001-2_9
16. V. V. Vazirani, *Approximation Algorithms*, Springer, Berlin, Heidelberg, 2001. <https://doi.org/10.1007/978-3-662-04565-7>
17. A. H. Mokdad, D. F. Stroup, W. H. Giles, Public health surveillance for behavioral risk factors in a changing environment: Recommendations from the behavioral risk factor surveillance team, *MMWR Recomm. Rep.*, **52** (2003), 1–12. Available from: <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5209a1.htm>.
18. Centers for Disease Control and Prevention, Behavioral risk factor surveillance system: 2023 survey data and documentation, Public-use data and documentation, 2024. Available from: https://www.cdc.gov/brfss/annual_data/annual_2023.html.
19. Centers for Disease Control and Prevention, Behavioral risk factor surveillance system: 2024 survey data and documentation, Public-use data and documentation, 2025. Available from: https://www.cdc.gov/brfss/annual_data/annual_2024.html
20. H. Wang, Z. Cui, Y. Chen, M. Avidan, A. B. Abdallah, A. Kronzer, Predicting hospital readmission via cost-sensitive deep learning, *IEEE/ACM Trans. Comput. Biol. Bioinform.*, **15** (2018), 1968–1978. <https://doi.org/10.1109/TCBB.2018.2827029>
21. J. R. Birge, F. Louveaux, *Introduction to Stochastic Programming*, Springer, New York, 2011. <https://doi.org/10.1007/978-1-4614-0237-4>
22. A. Shapiro, Monte carlo sampling methods, *Handbooks Oper. Res. Manag. Sci.*, **10** (2003), 353–425. [https://doi.org/10.1016/S0927-0507\(03\)10006-0](https://doi.org/10.1016/S0927-0507(03)10006-0)
23. A. J. Kleywegt, A. Shapiro, T. Homem-de Mello, The sample average approximation method for stochastic discrete optimization, *SIAM J. Optim.*, **12** (2002), 479–502. <https://doi.org/10.1137/S1052623499363220>
24. S. Asur, B. A. Huberman, Predicting the future with social media, In: *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 2010, 492–499. <https://doi.org/10.1109/WI-IAT.2010.63>

25. H. Choi, H. Varian, Predicting the present with google trends, *Econ. Rec.*, **88** (2012), 2–9. <https://doi.org/10.1111/j.1475-4932.2012.00809.x>
26. D. Gruhl, R. Guha, R. Kumar, J. Novak, A. Tomkins, The predictive power of online chatter, In: *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, 2005, 78–87. <https://doi.org/10.1145/1081870.1081883>
27. S. Goel, J. M. Hofman, S. Lahaie, D. M. Pennock, D. J. Watts, Predicting consumer behavior with web search, *Proc. Natl. Acad. Sci. U.S.A.*, **107** (2010), 17486–17490. <https://doi.org/10.1073/pnas.1005962107>
28. Z. Da, J. Engelberg, P. Gao, In search of attention, *J. Finance*, **66** (2011), 1461–1499. <https://doi.org/10.1111/j.1540-6261.2011.01679.x>
29. A. Rajkomar, E. Oren, K. Chen, A. M. Dai, N. Hajaj, M. Hardt, P. J. Liu, et al., Scalable and accurate deep learning with electronic health records, *npj Digit. Med.*, **1** (2018), 18. <https://doi.org/10.1038/s41746-018-0029-1>
30. A. L. Beam, I. S. Kohane, Big data and machine learning in health care, *JAMA*, **319** (2018), 1317–1318. <https://doi.org/10.1001/jama.2017.18391>
31. E. J. Topol, High-performance medicine: The convergence of human and artificial intelligence, *Nat. Med.*, **25** (2019), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
32. Z. Obermeyer, E. J. Emanuel, Predicting the future: Big data, machine learning, and clinical medicine, *N. Engl. J. Med.*, **375** (2016), 1216–1219. <https://doi.org/10.1056/NEJMp1606181>
33. J. H. Chen, S. M. Asch, Machine learning and prediction in medicine: Beyond the peak of inflated expectations, *N. Engl. J. Med.*, **376** (2017), 2507–2509. <https://doi.org/10.1056/NEJMp1702071>
34. D. C. Angus, R. Khera, T. Lieu et al., AI, health, and health care today and tomorrow: The JAMA summit report on artificial intelligence, *JAMA*, **334** (2025), 1650–1664. <https://doi.org/10.1001/jama.2025.18490>
35. J. Wiens, S. Saria, M. Sendak, M. Ghassemi, V. X. Liu, F. Doshi-Velez, et al., Do no harm: A roadmap for responsible machine learning for health care, *Nat. Med.*, **25** (2019), 1337–1340. <https://doi.org/10.1038/s41591-019-0548-6>
36. C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, D. King, Key challenges for delivering clinical impact with artificial intelligence, *BMC Med.*, **17** (2019), 195. <https://doi.org/10.1186/s12916-019-1426-2>
37. E. H. Weissler, T. Naumann, T. Andersson et al., The role of machine learning in clinical research: Transforming the future of evidence generation, *Trials*, **22** (2021), 537. <https://doi.org/10.1186/s13063-021-05489-x>
38. L. Clifton, J. Powell, D. A. Clifton, A. Sheikh, When to and when not to use machine learning in risk prediction models, *Lancet Digit. Health*, **8** (2026), 100954. <https://doi.org/10.1016/j.landig.2025.100954>
39. MARS-ED Investigators, Machine learning for risk stratification in the emergency department (MARS-ED): A randomized controlled trial, *Nat. Commun.*, **17** (2026), 242. <https://doi.org/10.1038/s41467-025-66947-7>

40. M. Khan, M. Rashid, U. Ghani, M. Saleem, M. Faheem, M. A. Khan, Access and privacy control for healthcare decision support system: A smart medical data exchange engine (SMDEE), *CAAI Trans. Intell. Technol.*, **10** (2025), 1616–1632. <https://doi.org/10.1049/cit2.70077>
41. Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations, *Science*, **366** (2019), 447–453. <https://doi.org/10.1126/science.aax2342>
42. K. Ladin, J. Cuddeback, O. K. Duru, S. Goel, W. Harvey, J. G. Park, et al., Guidance for unbiased predictive information for healthcare decision-making and equity (GUIDE): Considerations when race may be a prognostic factor, *npj Digit. Med.*, **7** (2024), 290. <https://doi.org/10.1038/s41746-024-01245-y>
43. A. Rajkomar, M. Hardt, M. D. Howell, G. Corrado, M. H. Chin, Ensuring fairness in machine learning to advance health equity, *Ann. Intern. Med.*, **169** (2018), 866–872. <https://doi.org/10.7326/M18-1990>
44. Y. Yang, H. Zhang, J. W. Gichoya, D. Katabi, M. Ghassemi, The limits of fair medical imaging AI in real-world generalization, *Nat. Med.*, **30** (2024), 2838–2848. <https://doi.org/10.1038/s41591-024-03113-4>
45. J. Grand-Clément, Y. H. Goh, C. Chan, V. Goyal, E. Chuang, *Interpretable machine learning for resource allocation with application to ventilator triage*, 2021. <https://arxiv.org/abs/2110.10994>
46. S. Deo, K. Rajaram, S. Rath, U. S. Karmarkar, M. B. Goetz, Planning for HIV screening, testing, and care at the veterans health administration, *Oper. Res.*, **63** (2015), 287–304. <https://doi.org/10.1287/opre.2015.1353>
47. C. W. Chan, V. F. Farias, N. Bambos, G. J. Escobar, Optimizing intensive care unit discharge decisions with patient readmissions, *Oper. Res.*, **60** (2012), 1323–1341. <https://doi.org/10.1287/opre.1120.1105>
48. D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, G. Escobar, Big data in health care: Using analytics to identify and manage high-risk and high-cost patients, *Health Aff.*, **33** (2014), 1123–1131. <https://doi.org/10.1377/hlthaff.2014.0041>
49. P. W. F. Wilson, J. B. Meigs, L. Sullivan, C. S. Fox, D. M. Nathan, R. B. D’Agostino, Prediction of incident diabetes mellitus in middle-aged adults: The framingham offspring study, *Arch. Intern. Med.*, **167** (2007), 1068–1074. <https://doi.org/10.1001/archinte.167.10.1068>
50. J. Lindström, J. Tuomilehto, The diabetes risk score: A practical tool to predict type 2 diabetes risk, *Diabetes Care*, **26** (2003), 725–731. <https://doi.org/10.2337/diacare.26.3.725>
51. H. Bang, A. M. Edwards, A. S. Bombardier, C. M. Ballantyne, D. Brillon, M. A. Callahan, et al., Development and validation of a patient self-assessment score for diabetes risk, *Ann. Intern. Med.*, **151** (2009), 775–783. <https://doi.org/10.7326/0003-4819-151-11-200912010-00005>
52. R. Kahn, P. Alperin, D. Eddy, K. Borch-Johnsen, J. Buse, J. Feigelman, et al., Age at initiation and frequency of screening to detect type 2 diabetes: A cost-effectiveness analysis, *Lancet*, **375** (2010), 1365–1374. [https://doi.org/10.1016/S0140-6736\(09\)62162-0](https://doi.org/10.1016/S0140-6736(09)62162-0)

53. E. W. Gregg, X. Zhuo, Y. J. Cheng, A. L. Albright, K. M. V. Narayan, T. J. Thompson, Trends in lifetime risk and years of life lost due to diabetes in the USA, 1985–2011: A modelling study, *Lancet Diabetes Endocrinol.*, **2** (2014), 867–874. [https://doi.org/10.1016/S2213-8587\(14\)70161-5](https://doi.org/10.1016/S2213-8587(14)70161-5)
54. A. L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming, *Oper. Res.*, **21** (1973), 1154–1157. <https://doi.org/10.1287/opre.21.5.1154>
55. A. Ben-Tal, A. Nemirovski, Robust convex optimization, *Math. Oper. Res.*, **23** (1998), 769–805. <https://doi.org/10.1287/moor.23.4.769>
56. D. Bertsimas, M. Sim, The price of robustness, *Oper. Res.*, **52** (2004), 35–53. <https://doi.org/10.1287/opre.1030.0065>
57. A. Ben-Tal, L. El Ghaoui, A. Nemirovski, *Robust Optimization*, Princeton University Press, Princeton, NJ, 2009. <https://doi.org/10.1515/9781400831050>
58. W. Wiesemann, D. Kuhn, M. Sim, Distributionally robust convex optimization, *Oper. Res.*, **62** (2014), 1358–1376. <https://doi.org/10.1287/opre.2014.1314>
59. E. Delage, Y. Ye, Distributionally robust optimization under moment uncertainty with application to data-driven problems, *Oper. Res.*, **58** (2010), 595–612. <https://doi.org/10.1287/opre.1090.0741>
60. P. Mohajerin Esfahani, D. Kuhn, Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations, *Math. Program.*, **171** (2018), 115–166. <https://doi.org/10.1007/s10107-017-1172-1>
61. R. T. Rockafellar, S. Uryasev, Optimization of conditional value-at-risk, *J. Risk*, **2** (2000), 21–41. <https://doi.org/10.21314/JOR.2000.038>
62. G. S. Collins, J. B. Reitsma, D. G. Altman, K. G. M. Moons, Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement, *Ann. Intern. Med.*, **162** (2015), 55–63. <https://doi.org/10.7326/M14-0697>
63. K. G. M. Moons, D. G. Altman, J. B. Reitsma, J. P. A. Ioannidis, P. Macaskill, E. W. Steyerberg, et al., Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration, *Ann. Intern. Med.*, **162** (2015), W1–W73. <https://doi.org/10.7326/M14-0698>
64. A. J. Vickers, E. B. Elkin, Decision curve analysis: A novel method for evaluating prediction models, *Med. Decis. Making*, **26** (2006), 565–574. <https://doi.org/10.1177/0272989X06295361>
65. D. Pisinger, An exact algorithm for large multiple knapsack problems, *Eur. J. Oper. Res.*, **114** (1999), 528–541. [https://doi.org/10.1016/S0377-2217\(98\)00120-9](https://doi.org/10.1016/S0377-2217(98)00120-9)
66. H. Rahimian, S. Mehrotra, Frameworks and results in distributionally robust optimization, *Open J. Math. Optim.*, **3** (2022), 1–85. <https://doi.org/10.5802/ojmo.15>



AIMS Press

©2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)