

# Equity-aware stochastic optimization of mammography screening facility location under demand and capacity uncertainty

Mahboubeh Madadi

`Mahboubeh.madadi@uth.tmc.edu`

The University of Texas Health Science Center at Houston

Caroline Wilson

The University of Texas Health Science Center at Houston

Cici Bauer

The University of Texas Health Science Center at Houston

Shayan Shams

The University of Texas Health Science Center at Houston

---

## Article

**Keywords:** Mammography screening, Healthcare accessibility, Facility location, Stochastic optimization, Health equity, Geographic information systems

**Posted Date:** October 13th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-10844614/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# Equity-aware stochastic optimization of mammography screening facility location under demand and capacity uncertainty

Mahboubeh Madadi<sup>1,\*</sup>, Caroline Wilson<sup>1</sup>, Cici Bauer<sup>2</sup>, Shayan Shams<sup>1</sup>

<sup>1</sup>McWilliams School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, TX, USA

<sup>2</sup>School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA

\*Corresponding author: Mahboubeh Madadi (mahboubeh.madadi@uth.tmc.edu)

## Abstract

Uneven geographic access to mammography can limit participation in breast cancer screening, particularly in rural and underserved communities. We developed a statewide planning framework that combines geospatial accessibility analysis with stochastic facility-location optimization to identify locations for new mammography facilities under uncertainty in screening demand and existing-facility capacity. The analysis included 5.58 million Texas women aged 40–74 years, 615 existing facilities, and 222 candidate locations. The model first maximized expected screening coverage and then reduced population-weighted differences in potential access across eight race and ethnicity groups without sacrificing coverage. Opening 62 facilities across 58 counties increased expected coverage from 92.47% to 98.09% and reduced expected unmet demand from 420,089 to 106,822 women annually. Between-group variance (BGV) in access decreased by 32%, while access for the worst-served group increased from 0.858 to 0.953 effective screening slots per eligible woman. All selected locations were in baseline access-gap areas; 44% were in nonmetropolitan counties and 39% established the first mammography facility in their county. Explicitly modeling uncertainty reduced expected unmet demand by 29,643 women compared with deterministic planning. These results show that a modest, equity-aware expansion of fixed mammography capacity can substantially widen screening access and narrow both racial and geographic disparities—by placing facilities where under-served populations live rather than where volume is highest—while the most remote communities are better served by mobile screening than by additional fixed sites.

**Keywords:** Mammography screening, Healthcare accessibility, Facility location, Stochastic optimization, Health equity, Geographic information systems

## 1. Introduction

Female breast cancer is the second most common cancer diagnosis and the fourth leading cause of cancer death in the United States.<sup>1</sup> Screening can identify disease at an earlier, more treatable stage, but its population benefit depends on women being able to reach appropriately equipped facilities. The 2024 U.S. Preventive Services Task Force recommends biennial screening mammography for women aged 40–74 years, increasing the population for whom timely geographic access to screening services is relevant.<sup>2</sup>

Mammography access is not distributed evenly. Long travel times, transportation limitations, and shortages of accredited facilities disproportionately affect rural communities. National analyses found that geographic accessibility changed little between 2006 and 2022 and remained substantially more limited in rural than urban areas.<sup>3</sup> Screening utilization also varies with socioeconomic conditions and local service availability; insurance coverage, educational attainment, rurality, and the number of nearby mammography facilities have been associated with geographic variation in screening rates.<sup>4</sup> These access differences are clinically relevant because GIS-based measures that account for both travel barriers and competition for services have been associated with neighborhood risk of late-stage breast cancer diagnosis.<sup>5</sup>

Geographic information systems (GIS) provide a framework for integrating population need with the location and capacity of healthcare services. Two-step floating catchment area methods measure accessibility by considering both the supply available within a catchment and the competing population demand for that supply.<sup>6</sup> Enhanced two-step floating catchment area methods further incorporate distance-decay weights so that facilities contribute less to access as travel burden increases.<sup>7</sup> These measures provide more information than provider counts or nearest-facility

distance alone, but descriptive accessibility analyses identify where gaps exist rather than determine which new locations would yield the greatest system-wide improvement.

Mobile mammography can supplement fixed facilities, although its contribution depends on where and how units are deployed. A Texas study found that mobile units contributed unevenly to spatial accessibility for uninsured women and suggested that geospatial planning could improve their placement.<sup>8</sup> More recently, an Oklahoma analysis used two-step floating catchment area methods to identify rural populations with limited or no mammography access and to inform mobile-service deployment.<sup>9</sup> Mobile services are therefore an important complementary strategy, but health systems and public agencies must still decide where permanent screening capacity should be added under limited budgets and workforce constraints.

Location-allocation optimization offers a systematic approach for evaluating many feasible configurations while explicitly representing travel, capacity, and budget constraints.<sup>10</sup> In mammography planning, bi-objective models have been used to reconcile examination coverage with the number of units deployed.<sup>11</sup> Broader healthcare-allocation research has also shown how fairness can be incorporated directly into resource-allocation models rather than assessed only after decisions are made.<sup>12</sup> During the COVID-19 pandemic, multi-objective test-site models demonstrated that facility placement could jointly improve population coverage and equity in access,<sup>13</sup> and adaptive formulations subsequently balanced access, statistical precision, and sociodemographic equity as needs changed over time.<sup>14</sup>

Despite these advances, important gaps remain for statewide mammography planning. Prior studies have generally been descriptive, focused on a single metropolitan area or smaller geographic region, evaluated mobile deployment, or assumed deterministic demand and capacity. These limitations are especially consequential in Texas, where large metropolitan centers coexist with extensive rural and sparsely populated areas and where facilities serve populations across county boundaries. A statewide model must therefore distinguish geographic reachability from capacity-limited service, account for competition for shared facilities, and avoid improving average access at the expense of groups that are already less well served.

We address this need by developing an equity-aware, two-stage stochastic facility-location model for mammography screening in Texas. The planning problem is to decide which of a finite set of candidate locations to open, under a fixed facility budget and uncertainty in screening demand and existing-facility capacity, so as to first serve as much of the eligible population as possible and then distribute access as equitably as possible across race and ethnicity groups. Facility openings are selected before demand and existing-facility capacity are known, while patient allocation and unmet demand adapt to the realized scenario, consistent with a two-stage stochastic decision structure.<sup>15</sup> Because facility-level screening capacity is not publicly reported, we first estimate it for each existing facility from observed procedure volume, using facility-class priors where volume is unavailable. Candidate locations are then generated from tract-level spatial-access gaps, and the model is solved lexicographically: it first minimizes expected unmet screening demand—maximizing coverage—and then minimizes population-weighted disparities in potential access across eight mutually exclusive race and ethnicity groups while preserving near-optimal coverage.

Estimated capacity of the existing facilities and the need-informed candidate set that define the optimization inputs are first presented in the Results section. We then evaluate the resulting plan along four dimensions. First, we quantify the gains in screening coverage and potential access relative to the existing network. Second, we examine how these gains are distributed across race and ethnicity groups and along the rural–urban gradient, and how well the selected sites align with underlying need and social vulnerability. Third, we measure the value of explicitly modeling uncertainty by comparing the stochastic plan against deterministic and perfect-information benchmarks. Finally, we test the robustness of the coverage and equity outcomes to alternative facility budgets, distance-decay assumptions, and existing-facility capacity estimates.

## **2. Results**

We first present the two model inputs derived from the data, facility-level screening-capacity estimates and the generated candidate sites, followed by the optimal siting plan and its effects on screening coverage, potential access, and race/ethnicity equity. We then report the value of explicitly modeling uncertainty and the sensitivity of the findings to key assumptions.

## 2.1. Facility-capacity estimates

The existing supply comprised 615 Texas mammography facilities certified under the U.S. FDA Mammography Quality Standards Act (MQSA)<sup>16</sup>, for which we estimated annual screening capacity. The 2024 Texas Health Care Information Collection (THCIC) outpatient data<sup>17</sup> contained 1,321,990 screening mammograms and 320,996 diagnostic-only mammograms across 311 reporting facilities (screening = 80.5% of volume); observed screening volume was assigned to 133 of the MQSA facilities, while class-based unit priors were used for the remaining 482. Under the conservative specification, estimated existing screening capacity was  $\approx 5.16$  million examinations per year, below the annual eligible-woman demand of 5,581,261.<sup>18,19</sup>

The capacity estimates were evaluated against independent benchmarks. County facility capacity under the conservative specification was strongly correlated with the county population estimated to have been screened using CDC PLACES<sup>20</sup> (Pearson  $r = 0.97$ ;  $n = 254$ ), providing a spatial capacity-versus-demand check. A separate demand-source consistency check compared expected annual screening volume (eligible population  $\times$  two-year mammography prevalence  $\div 2$ ) with observed residence-based county mammography volume and yielded  $r = 0.94$  ( $n = 254$ ). Because both comparisons scale partly with county population, they primarily validate geographic distribution rather than absolute capacity calibration. Absolute plausibility was assessed separately: implied utilization was approximately 31%, close to the national benchmark of approximately 32%.<sup>21</sup>

## 2.2. Candidate sites and baseline spatial access

Candidate sites were generated from tract-level Enhanced Two-Step Floating Catchment Area (E2SFCA)<sup>7</sup> access gaps computed from the existing network ( $\beta = 0.025 \text{ km}^{-1}$ , 60-km Haversine catchment, conservative  $\rho = 1.0$ ), classifying tracts with potential access below 0.80 effective screening slots per eligible woman as access-gap tracts. Sites were placed by population-weighted spatial clustering of gap tracts and retained only when at least 2,500 screening-eligible women lived within 60 km, with the number of sites per county scaled to the local access shortfall (nominal design capacity 7,500 examinations) and capped at six (refer to Methods).

Under this specification, 2,060 of the 6,896 census tracts<sup>22</sup>, containing approximately 1.58 million screening-eligible women<sup>19</sup>, were classified as access-gap tracts. The finalized candidate pool contained 222 sites across 201 of the 254 counties. Most of these counties (186) received a single candidate, with additional sites in more populous or spatially extensive counties, up to the six-per-county cap. The remaining 53 counties received no candidate, having either no access-gap tracts or gap tracts too sparsely populated to meet the 2,500-woman viability threshold. Candidate coordinates represent tract-centroid planning locations rather than specific parcels, and the pool defines the options evaluated by the optimization rather than the final siting.

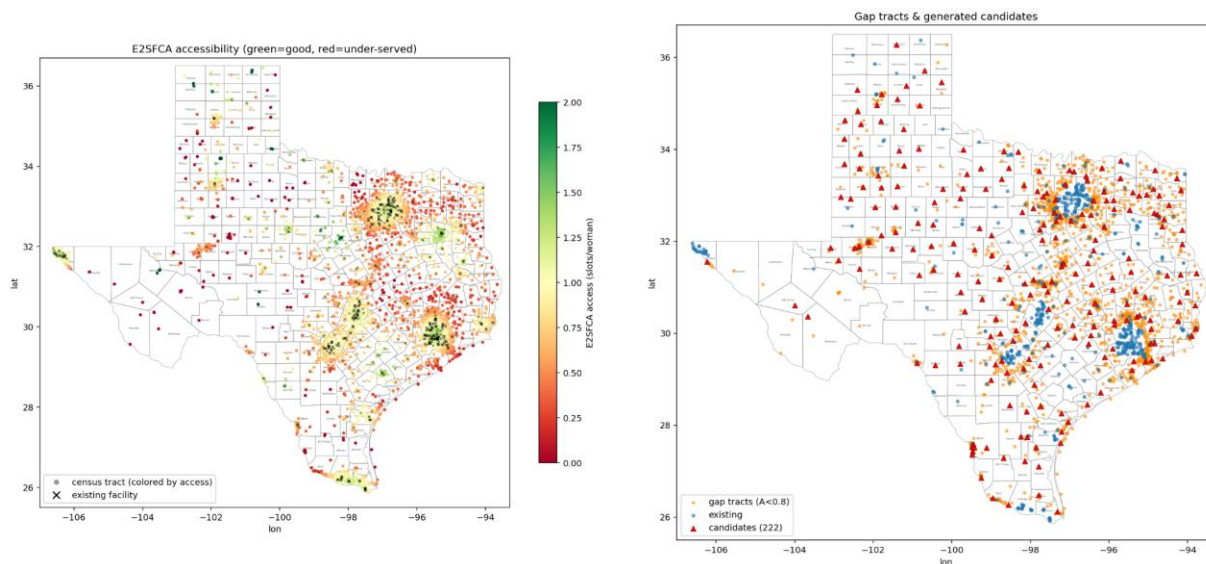


**Figure 1.** Spatial accessibility gaps and generated mammography-facility candidates in Texas. (A) Census-tract accessibility under the primary specification ( $\beta = 0.025 \text{ km}^{-1}$ ;  $d_0 = 60 \text{ km}$ ); black crosses

denote existing facilities. (B) Gap tracts ( $< 0.80$ ), existing facilities, and the 222 generated candidate sites. Candidate points represent tract-centroid planning locations rather than exact parcels.

When all existing and candidate facilities were assumed available, 99.87% of statewide screening-eligible demand fell within the 60-km Haversine catchment of a facility, confirming that the candidate pool spans essentially the entire state; the small residual was concentrated in sparsely populated far-West Texas counties. This audit assessed only the geographic breadth of the candidate pool and did not determine which facilities would be opened. Figure 1 presents the baseline tract-level accessibility surface and the geographic distribution of the candidate sites.

### 2.3. Optimal stochastic facility-location solution

In the baseline analysis, the model could open up to 62 new facilities, an approximately 10% expansion of existing network). To quantify the value of the siting intervention, we compared the optimized plan against a do-nothing baseline in which only the 615 existing facilities operate. Under the existing network, expected coverage was 92.47%, leaving approximately 420,089 eligible women per year unserved across the nine demand-and-capacity scenarios. Adding the 62 optimally sited facilities raised expected coverage to 98.09%—a 5.6-percentage-point gain that reduced expected unmet demand by roughly 313,000 women per year. These coverage gains were accompanied by a simultaneous reduction in race and ethnicity access disparity: the population-weighted between-group variance (BVG) of access fell 32%, from  $3.170 \times 10^{-3}$  to  $2.160 \times 10^{-3}$ . Table 1 summarizes the results.

**Table 1.** Results comparison under the existing-facility baseline and optimized siting plan

| Metric                                  | Baseline (existing facilities only) | Optimized: + 62 new facilities (Change) |
|-----------------------------------------|-------------------------------------|-----------------------------------------|
| Expected coverage                       | 92.47%                              | 98.09% (+5.62%)                         |
| Expected unmet demand (women/year)      | 420,089                             | 106,822 (−74.57%)                       |
| BVG of potential access (base scenario) | $3.170 \times 10^{-3}$              | $2.160 \times 10^{-3}$ (−32%)           |

The model used the full statewide budget, opening 62 facilities across 58 counties. Four counties, Webb, Grayson, McLennan, and Johnson, received two facilities, while the remaining 54 received one. No county received more than two facilities; therefore, the per-county cap of three was nonbinding. Expected unmet demand was 106,822 women per year, corresponding to 98.09% coverage. The probability-weighted mean of the nine scenario-specific coverage ratios was 98.24%.

As described, the model was solved in two lexicographic stages—first maximizing expected coverage, then minimizing access disparity within the set of coverage-optimal plans (Methods). Under this structure, the expected population-weighted BGV of potential access was minimized to  $2.379 \times 10^{-3}$ .

Under the baseline-demand, conservative-capacity scenario, 228 of 254 counties (89.8%) achieved complete service coverage. Of the 26 counties without complete coverage, 23 remained below 80%, one fell between 80% and 90%, and two between 90% and 100%. The counties below 80% are almost all small, remote areas of far-West and South Texas and the rolling plains—for example Brewster, Presidio, Jeff Davis, Pecos, Hudspeth, Kenedy, and Zapata.

On the county-level, the siting plan increased population-weighted mean potential access from 0.925 to 1.008 effective screening slots per eligible woman. The number of counties with potential access below 0.80 decreased from 178 to 141, and roughly 401,705 eligible women live in counties lifted above this threshold. Of the 141 counties that remained below 0.80, 118 received no new facility, largely because they were isolated, sparsely populated areas that no fixed candidate site could reach within the 60-minute access rule; the remaining 23 received at least one facility but stayed below 0.80 because their large eligible populations dilute the added capacity (for example Grayson, Ellis, Hunt, Kaufman, and Midland). Thus, service coverage and potential access identify largely different underserved counties: the counties below 80% service coverage are overwhelmingly small and remote, whereas the 141 below the 0.80 potential-access threshold include populous counties that received a facility yet remained capacity-constrained—underscoring that reachable capacity, not nominal coverage, is the binding constraint across much of rural Texas.

Figure 2 shows the geographic distribution of the 62 selected facilities together with tract-level E2SFCA access under the optimal siting plan. New sites were distributed across metropolitan peripheries, mid-size communities, and rural access gaps rather than being concentrated solely in the largest urban centers. On the straight-line tract layer used for this map, the number of underserved tracts ( $A < 0.80$ ) declined from 2,060 to 1,341. Remaining low-access areas were concentrated in sparsely populated parts of far-West Texas and the Panhandle, where the 60-minute reachability constraint limits the effectiveness of fixed-site expansion; notably, the Big Bend counties of Brewster, Jeff Davis, and Presidio—together roughly 4,200 eligible women—received no facility and remain at zero access, as no single fixed site can serve their dispersed, low-density demand within the budget, an area better suited to a mobile-screening program. Note that the map uses tract-level straight-line distances rather than the county-level OSRM drive times used in the optimization. Since straight-line distance understates travel where road networks are sparse, the mapped access estimates are optimistic relative to the county-level drive-time access used in the model, particularly in rural areas.

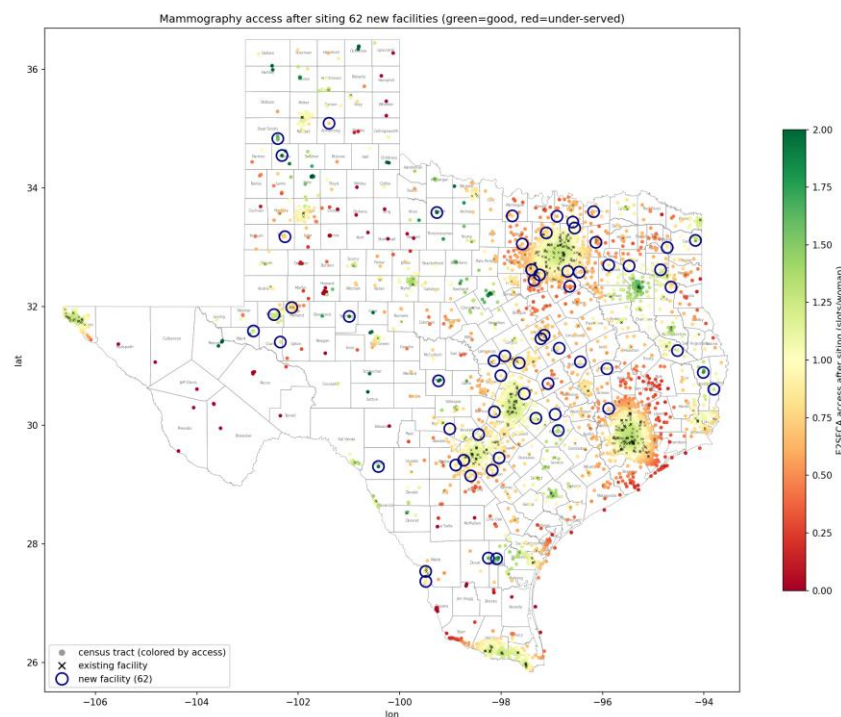


**Figure 2.** Geographic distribution of mammography screening access after siting the 62 new facilities selected by two-stage stochastic model. Census tracts are colored by post-siting potential E2SFCA access, with red indicating lower access and green indicating higher access. Black crosses mark the 615 existing MQSA-certified facilities, and hollow navy circles mark the 62 optimized new facilities.

**Scenario-specific coverage.** Table 2 reports scenario-specific coverage and unmet demand across the nine demand–capacity combinations. For all six scenarios with demand at or below baseline, coverage exceeded 99.6%; the residual unmet demand (~11,400–12,700 women, scaling with the demand multiplier) reflects a structural floor of remote far-West and South Texas counties (e.g., Brewster, Pecos) that no fixed facility can reach within 60 minutes and that go unserved in every scenario, largely independent of capacity. Existing capacity became decisive only under 10%-higher demand, where coverage ranged from 91.26% (most conservative capacity,  $\rho = 1.0$ ) to 98.71% (highest inferred capacity,  $\rho = 0.6$ ), since lower  $\rho$  implies additional latent units. Because this floor persists across all scenarios, the far-West gap is a geographic limitation of fixed-site expansion—better addressed by a complementary mobile-screening program—than a capacity shortfall. The probability-weighted disparity objective (BGV) across scenarios was  $2.379 \times 10^{-3}$ .

**Table 2.** Scenario-specific coverage and unmet demand for the optimal solution. Scenario probabilities are products of the demand and capacity-state marginal probabilities.

| Demand multiplier | Capacity parameter ( $\rho$ ) | Probability | Coverage | Unmet women |
|-------------------|-------------------------------|-------------|----------|-------------|
| 0.9               | 1.0                           | 12.5%       | 99.77%   | 11,402.1    |
| 0.9               | 0.8                           | 7.5%        | 99.77%   | 11,402.1    |
| 0.9               | 0.6                           | 5.0%        | 99.77%   | 11,402.1    |
| 1.0               | 1.0                           | 25.0%       | 99.65%   | 19,357.0    |
| 1.0               | 0.8                           | 15.0%       | 99.77%   | 12,669.0    |
| 1.0               | 0.6                           | 10.0%       | 99.77%   | 12,669.0    |
| 1.1               | 1.0                           | 12.5%       | 91.26%   | 536,289.2   |
| 1.1               | 0.8                           | 7.5%        | 94.58%   | 333,009.6   |
| 1.1               | 0.6                           | 5.0%        | 98.71%   | 79,057.1    |

## 2.4. Equity in potential access across race and ethnicity groups.

We assessed equity across the eight mutually exclusive race and ethnicity groups using the population-weighted BGV of potential access. Figure 3 presents the results. Under the existing network, access was already relatively balanced but showed a distinct rural gradient: the lowest-access groups were the more rurally distributed non-Hispanic White (0.858), Pacific Islander (0.879), and Native American (0.899) populations, whereas the urban-concentrated Hispanic (0.972) and non-Hispanic Black (0.973) populations already had higher access. Siting the 62 facilities improved access for every group and preferentially benefited those starting lowest—gains ranged from +0.099 for Pacific Islanders and +0.095 for non-Hispanic Whites to +0.052 for the highest-baseline group, non-Hispanic Asians—compressing the distribution: the worst-off group rose from 0.858 to 0.953 and the population-weighted mean from 0.925 to 1.008, narrowing the between-group variance by nearly a third (Figure 3). Because coverage is fixed before disparity is minimized, these gains came from adding capacity where under-served groups live rather than from leveling access down—every group improved and none was made worse off. The residual gap is therefore geographic rather than a racial-minority disadvantage: the groups furthest from parity are those in sparsely populated far-West counties beyond a 60-minute reach of any fixed facility.

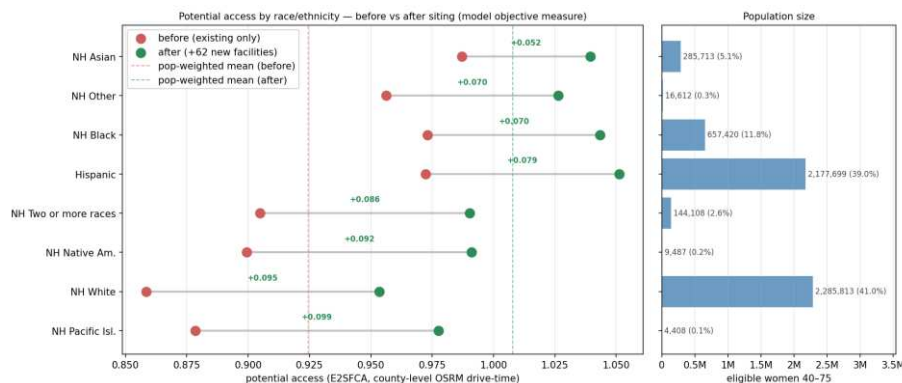


**Figure 3.** Potential access to mammography by race/ethnicity, before (existing facilities only) versus after siting 62 new facilities. Left: population-weighted mean county-level access (effective slots per eligible woman) for each of the eight groups (ordered by gain); dashed lines are the population-weighted group means. Right: each group's share of eligible women. Every group improved, with the largest gains for the initially lowest-access (rurally distributed) groups.

## 2.5. Geographic and social context of selected sites

To characterize where the optimized solution directs new capacity, we classified each of the 62 sited facilities along three dimensions: rurality (USDA Rural-Urban Continuum Codes (RUCC)<sup>23</sup>, by county), baseline under-service (the E2SFCA access of each site's census tract computed from existing facilities only—i.e., before the new facility was added; a site is classed as under-served if this pre-siting access was below 0.80), and social vulnerability (CDC/ATSDR Social Vulnerability Index

(SVI)<sup>24</sup>). Table 3 presents the results. Siting was distributed across the rural–urban gradient rather than concentrated in large cities—35 facilities (56%) in metropolitan counties, 15 (24%) in nonmetropolitan/intermediate-urbanization counties, and 12 (19%) in rural counties, sending 44% to non-metropolitan Texas. Placement was strongly need-driven: all 62 (100%) fell in tracts with baseline access below 0.80 (median 0.54), and 24 (39%) opened the first facility in their county (mammography deserts), with none directed to already-adequate areas. Baseline under-service is assessed on the tract-level E2SFCA surface (straight-line distance, the same layer used for the access maps). Sited tracts were not concentrated in the most socially vulnerable areas: only 11% of the sited tracts fell in the most-vulnerable quartile of Texas tracts, the 62 sited tracts had a mean overall SVI of 0.46 (near the statewide median of 0.50), and their racial and ethnic minority-status score (SVI Theme 3) was 0.33, below the median. The disparity closed is thus primarily geographic—rural and suburban distance gaps—rather than a social-vulnerability or minority-concentration gap; prioritizing high-SVI populations would require an explicit vulnerability-weighted objective.<sup>23,24</sup>

**Table 3.** Distribution of the 62 facilities across rurality, baseline under-service, and social vulnerability.

| Dimension     | Category                                                   | <i>n</i> | Share |
|---------------|------------------------------------------------------------|----------|-------|
| Rurality      | Metropolitan (RUCC 1–3)                                    | 35       | 56%   |
|               | Nonmetropolitan/intermediate urbanization (RUCC 4–7)       | 15       | 24%   |
|               | Rural (RUCC 8–9)                                           | 12       | 19%   |
| Under-service | Located in an access gap (E2SFCA < 0.80)                   | 62       | 100%  |
|               | First facility in the county (mammography facility desert) | 24       | 39%   |
|               | Median baseline access                                     | 0.54     | —     |
| SVI           | Q1: least vulnerable                                       | 12       | 19%   |
|               | Q2                                                         | 27       | 44%   |
|               | Q3                                                         | 16       | 26%   |
|               | Q4: most vulnerable                                        | 7        | 11%   |

## 2.6. Value of the stochastic solution

To quantify the benefit of accounting for uncertainty in demand and capacity, we evaluated the value of the stochastic solution (VSS) and the expected value of perfect information (EVPI) (See Methods). The stochastic plan (SP) left an expected SP = 106,822 women unserved per year, compared with EEV = 136,465 for the deterministic mean-value plan—optimized for average conditions, then evaluated across all nine scenarios—and WS = 100,369 for the wait-and-see plan, which adapts optimally to each realized scenario. The value of the stochastic solution, VSS = EEV – SP = 29,643 women (27.8% of SP), indicates that ignoring uncertainty would leave over 29,600 more women unserved each year, while the small expected value of perfect information, EVPI = SP – WS = 6,453 women (6.0%), places the stochastic plan within about 6% of the perfect-information bound.

## 2.7. Sensitivity analyses

We conducted three one-at-a-time sensitivity analyses by varying the new-facility budget  $B$ , the distance-decay parameter  $\beta$ , and the capacity scale  $\rho$  applied to existing facilities. In each analysis, the complete two-phase model was re-solved while all other parameters were held at their baseline values. Model performance was evaluated using expected coverage, tract-level population-weighted BGV reduction, worst-group access, and the number of women whose potential-access score increased from below 0.8 to at least 0.8. Candidate-facility capacity was held fixed at  $\kappa^C=7,500$  screening examinations per year throughout all analyses.

For the budget analysis, the model was re-solved for  $B \in \{40, 50, 62, 70, 80, 90\}$ , where  $B$  denotes the maximum number of new facilities. Figure 4 presents the results. Expected coverage increased monotonically but with diminishing returns: raising  $B$  from 40 to 62 increased coverage from 96.63% to 98.09% (+1.46 percentage points), whereas raising  $B$  from 62 to 90 produced only 0.87 additional points, reaching 98.96%. These results place the baseline budget  $B = 62$  near the knee of the coverage frontier. The number of women lifted above the 0.80 threshold increased steadily, from 380,152 at  $B = 40$  to 547,065 at  $B = 62$  and 737,267 at  $B = 90$ , and worst-group tract-level access rose monotonically from 0.928 to 0.966 to 1.004. The tract-level population-weighted BGV reduction

increased overall with budget (from 19% at  $B = 40$  to 63% at  $B = 90$ ) but was not strictly monotonic, dipping at some intermediate budgets; this reflects both the different coverage-optimal facility set selected at each budget and the use of tract-level straight-line E2SFCA for the displayed sensitivity measure, whereas the optimization minimizes county-level drive-time BGV.

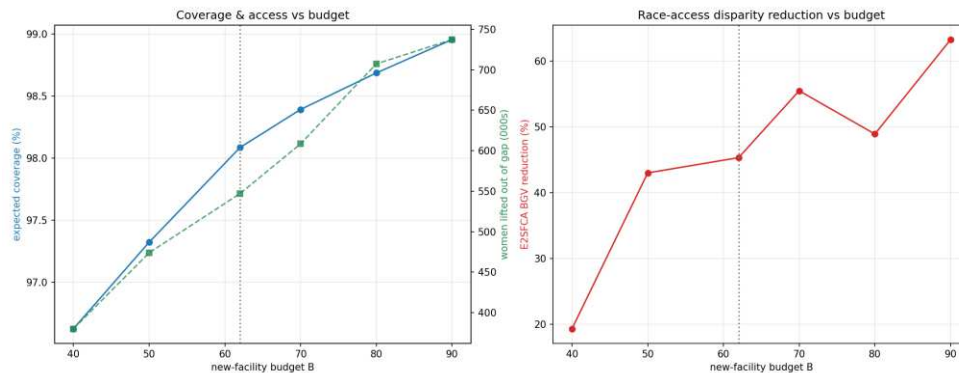


**Figure 4-** Sensitivity of the optimized solution to the new-facility budget  $B$ . Expected coverage rises with diminishing returns, with a knee near the baseline  $B = 62$ , while the number of women lifted above the 0.80 access grows nearly linearly; worst-group access increases monotonically with  $B$ .

The distance-decay analysis increased  $\beta$  from its baseline of 0.025 to 0.05 and 0.10 per minute (steeper decay). Figure 5 (bottom row) shows the results. Expected coverage remained essentially invariant at 98.09% across all three values, because coverage is determined by the binary 60-minute accessibility threshold rather than by the rate of decay within the catchment. Steeper decay made access more localized and increased the equity leverage of facility placement, raising the tract-level BGV reduction from 45% at  $\beta = 0.025$  to 48% at  $\beta = 0.05$  and 62% at  $\beta = 0.10$ . Worst-group access was essentially stable (0.965–0.966), while the number of women lifted above the threshold peaked at moderate decay (559,112 at  $\beta = 0.05$ ) and declined at  $\beta = 0.10$  (426,698), because steeper decay also lowers baseline access for populations located farther from facilities.

The capacity analysis varied estimated existing-facility capacity by  $\pm 20\%$  (scale factors 0.8 and 1.2) around the baseline, holding candidate-facility capacity fixed at 7,500 examinations per year. Figure 5 (top row) presents the results. Expected coverage was far more sensitive to this assumption than to  $\beta$ , rising from approximately 84.57% under the 0.8 scale to 98.09% at baseline and 99.91% under the 1.2 scale. As existing-facility capacity increased, both mean and worst-group access improved monotonically: population-weighted mean access reached 0.823, 1.008, and 1.193 effective screening slots per eligible woman under the 0.8, 1.0, and 1.2 capacity scales, respectively, with corresponding worst-group values of 0.784, 0.966, and 1.141. The marginal contribution of the new facilities, however, moved in the opposite direction. Because candidate capacity was held fixed, the number of women lifted above the 0.80 access threshold fell from 982,440 under the 0.8 scale to 547,065 at baseline and 377,520 under the 1.2 scale, as a better-resourced existing network had already brought more women above the threshold before any new facilities were sited. The tract-level BGV reduction was greatest near baseline (45%) and smaller at both extremes—24% under scarce (0.8) capacity and 37% under abundant (1.2) capacity—consistent with equity being capacity-limited when existing supply is scarce, since the fixed new capacity cannot fully close the gap, and increasingly geography-limited when supply is already abundant.

Across all evaluated budgets, decay parameters, and capacity scales, potential-access disparity declined (by roughly 19%–63% over the full range) and the worst-off group gained approximately 0.10 effective screening slots per woman, supporting the robustness of the equity conclusion. Expected coverage, in contrast, depended materially on the estimated capacity of the existing network (ranging from  $\sim 84.6\%$  to  $\sim 99.9\%$  across the capacity scale) but was insensitive to  $\beta$ . The existing-network capacity estimate is therefore the dominant determinant of the absolute headline coverage.

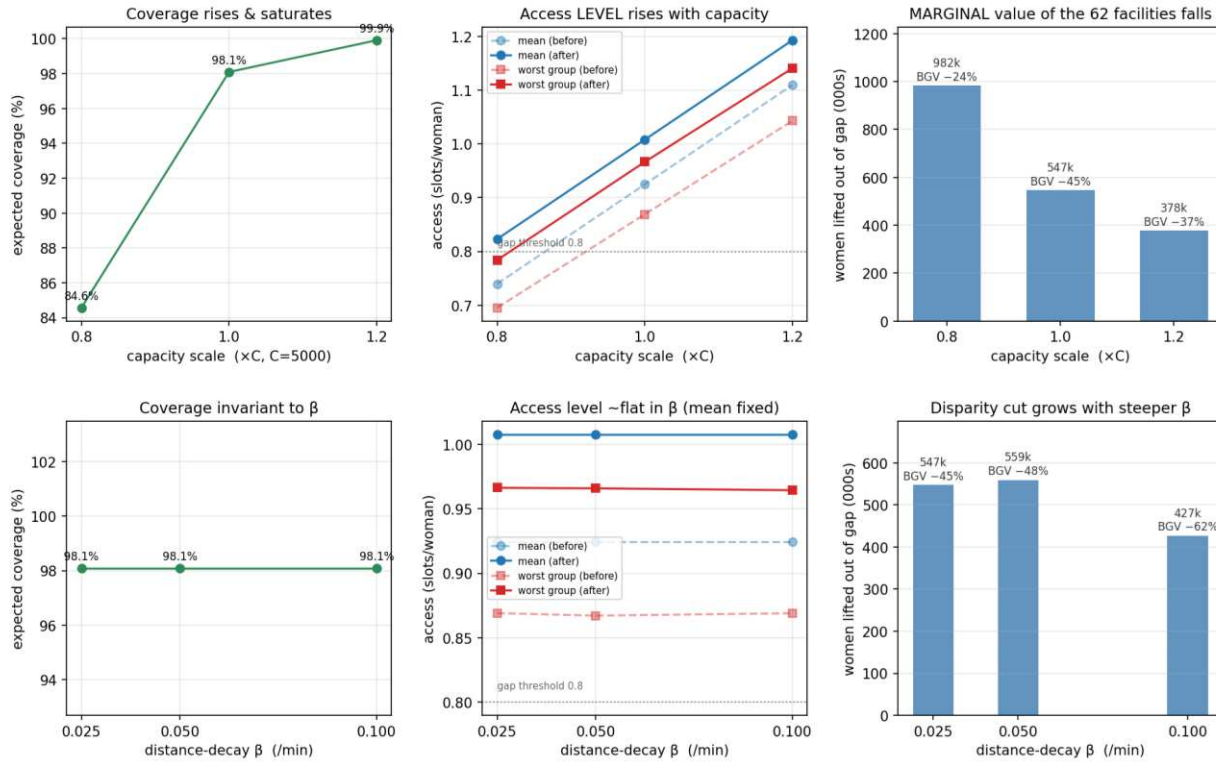


**Figure 5-** One-at-a-time sensitivity to existing-facility capacity (top row, scale  $\times C$  with candidate capacity fixed at 7,500/yr) and distance-decay  $\beta$  (bottom row). Left column: expected coverage—strongly increasing and saturating in capacity, invariant to  $\beta$ . Center column: mean and worst-group access levels ("before" = existing facilities only; "after" = +62 sited facilities). Right column: marginal value of the 62 facilities—women lifted above the 0.80 threshold and the tract-level BGV reduction—largest when existing capacity is scarce and increasing with steeper  $\beta$ . Access and disparity measures use tract-level straight-line E2SFCA.

### 3. Discussion

By measuring disparity using potential E2SFCA access as a function of the opening decisions—rather than of a system-optimal patient-routing pattern that a planner could not directly enforce—the model ensured that reducing disparity necessarily corresponds to placing capacity where under-served groups live. At the baseline budget, the solution raised expected coverage from 92.47% under the existing network to 98.09% (roughly 313,000 fewer unmet women per year) while improving potential access for every race and ethnicity group. The gain of approximately 0.10 in worst-group access was consistent across the county-level drive-time objective (0.858 to 0.953) and the tract-level visualization layer (0.869 to 0.966), supporting the robustness of the principal equity finding. Notably, these coverage and equity improvements were achieved jointly rather than traded against one another, a direct consequence of the two-phase lexicographic structure, which fixes coverage before minimizing disparity and thereby prevents equalization through "leveling down."

The existing network was already comparatively equitable across race and ethnicity: baseline potential access spanned a narrow band (0.858–0.987 screening slots per eligible woman across the eight groups, a coefficient of variation of roughly 6%). Moreover, the lowest-access groups were the more rurally distributed populations—non-Hispanic White, Pacific Islander, and Native American—whereas the urban-concentrated Hispanic and non-Hispanic Black populations already enjoyed higher access. The residual gap is therefore primarily geographic, reflecting rural and suburban distance to care, rather than a racial-minority disadvantage. The optimization narrows this gap by expanding rural and peri-urban access, and its equity contribution should be framed accordingly.

These findings extend earlier spatial studies that documented persistent geographic disparities in mammography access but primarily described the existing distribution of services. Prior GIS-based

analyses linked reduced mammography accessibility to rural residence and greater risk of late-stage diagnosis, and identified underserved communities in states such as Delaware and Illinois; however, these studies did not determine how limited new capacity should be allocated under uncertainty.<sup>3,5,25,26</sup> Our study also builds on location-allocation research showing that mammography coverage can be improved under budget constraints and on broader healthcare optimization models that balance accessibility, capacity, and equity.<sup>10–14</sup> The present framework advances this literature by combining empirically informed facility-capacity estimates, statewide candidate generation, uncertain demand and capacity, and a lexicographic objective that preserves near-optimal coverage before reducing race- and ethnicity-based differences in potential access. Unlike models in which equity depends on a centrally imposed patient-allocation pattern, the disparity reduction here is produced directly through facility placement. Finally, the persistent gaps observed in remote West Texas are consistent with prior research on mobile mammography, suggesting that fixed facilities and mobile services should be viewed as complementary strategies for geographically dispersed populations.<sup>8,9</sup> The magnitude of the BGV reduction depended on the measurement layer: 32% on the model's county-level drive-time metric and 45% on the finer tract-level straight-line proxy. This difference demonstrates that disparity estimates are sensitive to spatial resolution and to the representation of travel. Straight-line distance understates travel in areas with sparse or indirect road networks, inflating the proportional reduction. Accordingly, all reported statistics use the county-level drive-time measure—the quantity the model actually optimizes—while the tract-level map and group comparison are retained only to convey where gains occurred.

The value-of-information analysis supports the use of a stochastic formulation. The expected-value (deterministic) plan left approximately 29,600 more women unserved per year than the stochastic plan (VSS = 29,643, or 27.8% of the stochastic optimum), indicating a substantial benefit from explicitly modeling demand and capacity uncertainty. The comparatively modest EVPI (6,453, or 6.0% of the optimum) indicates that the stochastic solution captured most of the benefit available even from perfect foresight of the realized scenario. The scenario-specific results suggest that the benefit of hedging against uncertainty is particularly relevant under high-demand conditions, in which capacity constraints become binding. However, the reported VSS is an aggregate expected value and was not decomposed by scenario.

The budget analysis placed  $B = 62$  near the knee of the coverage frontier: expected coverage rose from 96.63% at  $B = 40$  to 98.09% at  $B = 62$ , but only to 98.96% at  $B = 90$ . Worst-group access, however, increased monotonically with budget (0.928 at  $B = 40$  to 1.004 at  $B = 90$ ), indicating that additional capacity continues to benefit the least-served group even after aggregate coverage saturates. The dominant quantitative limitation was the existing-capacity assumption: a 20% reduction in estimated existing-facility per-unit capacity lowered expected coverage from 98.1% to approximately 84.6%, whereas a 20% increase raised it to 99.9%, even with candidate capacity fixed at 7,500 examinations per year. By contrast, coverage was essentially invariant to the distance-decay parameter, which is governed by the binary 60-minute threshold. The capacity analysis also illustrates why threshold-crossing counts must be interpreted with care: more women crossed the 0.80 access threshold under the low-capacity setting (approximately 982,000, versus 547,000 at baseline) precisely because the baseline gap population was larger, despite substantially lower absolute access. The disparity reduction remained substantial across all settings (roughly 19%–63% over the full range of budgets, decay values, and capacity scales), reinforcing that the equity conclusion is robust even though the absolute coverage figure depends materially on the existing-capacity estimate.

Characterizing where new capacity was directed further clarifies the intervention's mechanism. Placement was strongly need-driven for access. Every sited facility fell in a tract with baseline access below 0.80, and 39% opened the first facility in their county (mammography deserts)—and was distributed across the rural–urban gradient rather than concentrated in large cities, with 44% of facilities in non-metropolitan counties. By contrast, sited tracts were not disproportionately the most socially vulnerable (mean SVI 0.46 versus 0.50 statewide; only 11% in the top national vulnerability quartile). This is consistent with the disparity being closed being geographic rather than a social-vulnerability or minority-concentration gap. A policy that also prioritized socially vulnerable populations

would require an explicit vulnerability-weighted objective; the present model optimizes access and race/ethnicity equity but does not target social vulnerability directly.

Persistent access gaps were concentrated in isolated, sparsely populated counties—141 counties remained below the 0.80 drive-time access threshold after siting, 118 of which received no new facility. Several far-West counties, including Brewster, Jeff Davis, and Presidio (together roughly 4,200 eligible women), remained at zero access because no single fixed site can serve their dispersed, low-density demand within a 60-minute travel limit. In these areas, fixed-facility expansion alone is structurally insufficient, and complementary strategies—mobile mammography, periodic outreach, transportation assistance, or relaxed travel thresholds—should be examined; the persistence of these gaps across all scenarios indicates that the limitation is geographic rather than a shortfall of capacity.

Several limitations warrant emphasis. First, existing-facility capacity was inferred from observed screening volume under an assumed per-unit throughput, and the sensitivity analysis identifies this as the principal driver of the absolute coverage result; more precise facility-level capacity data would sharpen the estimates. Second, candidate capacity was fixed at a single design value (7,500 examinations per year), and heterogeneous or endogenous facility sizing was not considered. Third, access was optimized at the county level using drive-time from population-weighted centroids; sub-county and intra-urban variation, congestion, and time-of-day effects are not captured, and the tract-level surface used for mapping relies on straight-line distance. Fourth, travel times were obtained from a public routing service reflecting contemporaneous road data. Additionally, demand was represented as one annual screening examination per woman aged 40–74 years and was not adjusted for current screening participation, individual risk, prior screening history, or differences among screening guidelines. The model therefore represents planning demand under a uniform annual-access standard rather than forecasted realized utilization. Future work could extend the framework to a vulnerability-weighted or geographically explicit equity objective, incorporate mobile-unit deployment as an additional decision, allow variable facility capacity, and evaluate access at finer spatial resolution—each of which would build directly on the siting-driven equity formulation developed here.

## **4. Methods**

### **4.1. Data sources**

This study integrated publicly available data describing mammography screening demand, existing facility supply, travel geography, race and ethnicity, rurality, and social vulnerability across Texas. Data were harmonized at the census-tract, county, or facility level. The study area included 254 counties and 6,896 census tracts.

#### **Population and demographic data**

Baseline mammography demand was derived from the U.S. Census Bureau 2018–2022 American Community Survey (ACS) five-year estimates. Census-tract population counts corresponding to women aged 40–74 years were retrieved through the Census Data API and aggregated to counties, yielding 5,581,261 screening-eligible women statewide.<sup>18,19</sup> Demand at each geographic node was defined as one annual screening examination per eligible woman and was not adjusted for current screening participation.

County race and ethnicity distributions were derived from ACS using eight mutually exclusive categories: Hispanic of any race and non-Hispanic White, Black, Native American, Asian, Pacific Islander, Other race, and Two or more races.<sup>27</sup> County-specific group proportions were applied to the screening-eligible population to estimate the number of women from each group residing in each county. Census-tract and county boundaries, geographic identifiers, tract coordinates, and mapping geometries were obtained from the 2022 Census TIGER/Line and cartographic boundary files.<sup>22</sup> ACS estimates were retrieved programmatically through the Census Data API.

#### **Existing mammography facilities and screening volume**

The existing supply network was based on the U.S. Food and Drug Administration Mammography Facility Database, which identifies facilities certified under the Mammography Quality Standards Act (MQSA). The analytic extract used in this study contained 615 facilities.<sup>16</sup> Observed annual screening volume was obtained from the 2024 Texas Health Care Information Collection outpatient public-use data. Mammography records were classified as screening or diagnostic examinations using standard

HPCPS/CPT definitions, and diagnostic-only examinations were excluded from the screening-throughput measure; the detailed coding algorithm is provided in the supplementary materials (Section S1).<sup>17</sup> CDC PLACES county-level mammography prevalence estimates from the 2022 release were used for external geographic validation and related screening-utilization diagnostics.<sup>20</sup>

### Travel geography

Two distance representations were used. Road-network driving times, computed with the Open-Source Routing Machine<sup>28,29</sup> over OpenStreetMap road-network data between population-weighted county origins and facilities, defined accessibility (a 60-minute threshold) and were used for the optimization and for the coverage, potential-access, and equity results reported in the text and in Tables 1–2.

Counties were represented by population-weighted origins derived from tract-level population counts and coordinates. A straight-line (haversine) tract-to-facility distance was used only where sub-county resolution is required—candidate-site generation, the fine-resolution accessibility maps, the tract-level characterization of the selected sites, and the sensitivity-analysis figures—because these require an access value at every census tract; because straight-line distance understates travel where road networks are sparse or indirect, these tract-level surfaces are somewhat optimistic relative to the county drive-time values reported in the text.

### Rurality and social vulnerability

County rurality was obtained from the U.S. Department of Agriculture Economic Research Service 2023 Rural–Urban Continuum Codes.<sup>23,24</sup> Census-tract social vulnerability was obtained from the 2022 CDC/ATSDR Social Vulnerability Index.<sup>24</sup> RUCC and SVI were used only to characterize the locations selected by the optimization model and were not included in the objective function or constraints.

## 4.2. Facility Capacity Estimation

Because facility-level capacity is not publicly reported, we estimated an annual screening capacity for each MQSA facility from observed throughput where available and from facility-class unit priors otherwise. THCIC volumes<sup>17</sup> were linked to MQSA facilities<sup>16</sup> by name-based matching constrained by metropolitan area and health system; observed screening volume  $v_j$  was assigned to 133 facilities, and the remaining 482 received accreditation-class unit priors (matching statistics and priors in Section S2 of supplementary materials). For a measured facility  $j$ , the credited number of mammography units and the resulting annual screening capacity  $\kappa_j$  were

$$u_j(\rho) = \min \left\{ u_{\max}, \left\lceil \frac{v_j}{C\rho} \right\rceil \right\}, j \in J^M, \quad \kappa_j(\rho) = C \cdot u_j(\rho), j \in J^M \cup J^U$$

where  $C$  is the annual screening capacity of one mammography unit,  $u_{\max}$  the maximum number of units credited to a facility,  $J^M$  and  $J^U$  the set of measured and unmeasured facilities, respectively, and  $\rho$  the assumed per-unit utilization used to infer the unit count; lower  $\rho$  implies lower utilization and therefore more installed units to explain the same observed volume. The utilization parameter enters only through the inferred unit count and is not subsequently multiplied into the capacity. Capacity was estimated under the conservative utilization assumption ( $\rho = 1.0$ ); the full set of utilization states used to construct the stochastic scenarios is defined below (see Section S4.1). Unmeasured facilities were assigned their class prior, held constant across utilization states. The resulting statewide capacity estimate and its validation against independent CDC PLACES<sup>20</sup> benchmarks are reported in the Results.

## 4.3. Spatial Accessibility and Candidate Site Generation

Potential access was measured with the enhanced two-step floating catchment area (E2SFCA) method with exponential distance decay. Let  $P_i$  denote the screening-eligible population in tract  $i$ ,  $\kappa_j$  the screening capacity of facility  $j$ ,  $d_{ij}$  the distance between  $i$  and  $j$ , and  $\beta$  the distance-decay rate. In the first step, a decay-weighted supply-to-demand ratio is computed for each facility; in the second step, location access is the sum of decayed facility ratios within the catchment:

$$R_j = \frac{\kappa_j}{\sum_{i: d_{ij} \leq d_0} P_i e^{-\beta d_{ij}}},$$

$$A_i = \sum_{j: d_{ij} \leq d_0} R_j e^{-\beta d_{ij}}.$$

$A_i$  is interpreted as effective annual screening slots per eligible woman. The optimization evaluates this construct at the county level on OSRM drive times (60-minute accessibility, with a 45-minute effective time for same-county pairs whose centroid route exceeds 60 minutes); the maps and descriptive group comparisons use a finer tract-level, straight-line implementation of the same construct.

Candidate sites define a finite, need-informed set of locations in which a new facility could be placed; the optimization subsequently selects which to open. Candidate generation applied the tract-level, straight-line implementation of the E2SFCA<sup>5,6</sup> construct ( $\beta = 0.025 \text{ km}^{-1}$ , 60-km catchment) under the conservative capacity specification, classifying tracts with access below 0.80 effective screening slots per eligible woman as access-gap tracts. Within each county containing gaps, gap tracts were partitioned by population-weighted k-means and each cluster was snapped to its most-populous tract centroid; the number of candidates per county scaled with the county's population-weighted access shortfall (divided by a 7,500-examination design capacity), capped at six per county. Provisional sites were retained only if at least 2,500 eligible women lived within 60 km. Full procedural detail is provided in Supplementary Material (S3).

#### 4.4. Two-stage stochastic optimization model

We formulated facility siting as a two-stage stochastic mixed-integer program with a lexicographic objective. Full notation is given in Tables S2 and S3. The first-stage binary decision  $y_j = 1$  opens candidate facility  $j$  before demand and existing-facility capacity are known. In each scenario  $s$ , second-stage variables  $x_{ij}^s \geq 0$  allocate women in county  $i$  to accessible facility  $j$ , and  $u_i^s \geq 0$  records unmet demand. Uncertainty is represented by nine scenarios combining three demand multipliers (0.9, 1.0, 1.1; weights 0.25, 0.50, 0.25) and three capacity-utilization states ( $\rho = 1.0, 0.8, 0.6$ ; weights 0.50, 0.30, 0.20); scenario probabilities  $\pi_s$  are the products of the marginal weights, and  $D_i^s$  and  $\kappa_j^s$  denote the scenario demand and capacity. A county–facility pair  $(i, j)$  is accessible,  $(i, j) \in \mathcal{P}$ , under the 60-minute rule, with decay weight

$$e_{ij} = \exp(-\beta d_{ij}) \mathbb{1}[(i, j) \in \mathcal{P}].$$

Each candidate facility is a dedicated center with a fixed annual screening capacity of 7,500 examinations, constant across scenarios.

Potential access is linear in the opening decisions. Defining the decay-weighted competing demand

$$\text{den}_j^s = \sum_{i:(i,j) \in \mathcal{P}} D_i^s e_{ij},$$

and each facility's access contribution

$$g_{ij}^s = \frac{\kappa_j^s e_{ij}}{\text{den}_j^s},$$

(both scenario constants), county potential access and group-mean access are

$$\begin{aligned} \hat{A}_i^s &= \sum_{j \in J^E} g_{ij}^s + \sum_{j \in J^C} g_{ij}^s y_j, \\ \hat{A}_k^s &= \frac{1}{R_k^s} \sum_{i \in I} D_i^s \phi_{ik} \hat{A}_i^s. \end{aligned}$$

where  $J^E$  and  $J^C$  are the existing and candidate facilities,  $\phi_{ik}$  the share of county  $i$ 's eligible women in group  $k$ .  $\hat{A}_k^s$  is the population-weighted mean potential access of group  $k$  and the normalizing term  $R_k^s = \sum_{i \in I} D_i^s \phi_{ik}$  is the total eligible-woman population of group  $k$  under scenario  $s$ . Because  $\text{den}_j^s$  and  $g_{ij}^s$  do not depend on the opening or allocation decisions, county potential access  $\hat{A}_i^s$ —and hence each group mean  $\hat{A}_k^s$ —is linear in the opening variables  $y$ .

The population-weighted between-group variance (BGV), with group weights  $w_k$  equal to each group's statewide share of eligible women and mean  $\bar{A}^s = \sum_{k \in K} w_k \hat{A}_k^s$  is

$$\sum_{k \in K} w_k \left( \hat{A}_k^s - \bar{A}^s \right)^2.$$

Phase 1 maximizes expected coverage, equivalently minimizing expected unmet demand:

$$U^* = \min_{y,x,u} \sum_{s \in S} \pi_s \sum_{i \in I} u_i^s.$$

subject to, for every scenario  $s$ , demand balance and capacity, together with the budget and per-county cap:

$$\begin{aligned} \sum_{j: (i,j) \in \mathcal{P}} x_{ij}^s + u_i^s &= D_i^s, \quad \forall i \in I, s \in S. \\ \sum_{i: (i,j) \in \mathcal{P}} x_{ij}^s &\leq \kappa_j^s, \quad \forall j \in J^E, \quad \sum_{i: (i,j) \in \mathcal{P}} x_{ij}^s \leq \kappa_j^s y_j, \quad \forall j \in J^C, s \in S. \\ \sum_{j \in J^C} y_j &\leq B, \quad \sum_{j \in J^C} y_j \leq \text{cap}_c, \forall c, y_j \in \{0,1\}, \quad x_{ij}^s \geq 0, \quad u_i^s \geq 0. \end{aligned}$$

Phase 2 then minimizes the expected population-weighted BGV of potential access subject to all Phase-1 constraints, the access definitions and a coverage lock that holds expected unmet demand within a small tolerance  $\epsilon$  of the Phase-1 optimum:

$$\min \sum_{s \in S} \pi_s \sum_{k \in K} w_k \left( \hat{A}_k^s - \bar{A}^s \right)^2, \text{ s.t. } \sum_{s \in S} \pi_s \sum_{i \in I} u_i^s \leq U^* + \epsilon, \quad \epsilon = 0.01U^* + 1.$$

The coverage lock prevents an apparent equity gain achieved by sacrificing coverage. Because each group-access measure  $\hat{A}_k^s$  is linear in  $y$ , the Phase-2 objective is a convex quadratic, and the extensive form is therefore a convex mixed-integer quadratic program (MIQP). The MIQP was implemented in Python and solved with Gurobi Optimizer 12.0.3; Phase 1 reached a 0.07% optimality gap (effectively optimal) and Phase 2 solved to optimality. Model dimensions and full solver settings are given in Supplementary material (Section S4.2).

#### 4.5. Value of information

On the Phase-1 coverage objective, we computed the value of the stochastic solution (VSS) and the expected value of perfect information (EVPI). Let  $Q(y, s)$  be the minimum unmet demand under scenario  $s$  for a fixed siting plan  $y$ . The stochastic optimum is  $SP = \min_y \sum_{s \in S} \pi_s Q(y, s)$ ; the expected result of the deterministic mean-value plan  $\hat{y}$  is  $EEV = \sum_{s \in S} \pi_s Q(\hat{y}, s)$ ; and the wait-and-see value, allowing the siting to adapt to each scenario, is  $WS = \sum_{s \in S} \pi_s \min_y Q(y, s)$ . Then

$$VSS = EEV - SP, \quad EVPI = SP - WS.$$

$VSS$  is the expected reduction in unmet demand from explicitly modeling uncertainty rather than using a mean-value plan, and  $EVPI$  is the maximum expected reduction attainable from perfect foresight of the realized scenario.

#### Acknowledgements

The authors thank the Texas Health Care Information Collection (THCIC) and the Texas Department of State Health Services for providing access to the Texas Outpatient Event Public Use Data File. The analyses, interpretations, and conclusions presented in this work are those of the authors and do not necessarily reflect the views of THCIC or the Texas Department of State Health Services.

Driving times were computed using the Open Source Routing Machine with map data © OpenStreetMap contributors, available under the Open Database License.

#### Author contributions

M.M. conceived the study, designed the optimization framework and methodology, and administered the project. M.M. and C.W. implemented the optimization and analysis code, conducted the investigation, performed the formal analysis, and prepared the figures. M.M., C.W. and C.B. curated the data. M.M. validated the results and supervised the work. All authors contributed to the interpretation of the results. M.M. wrote the original draft. M.M., C.W., C.B. and S.S. reviewed and edited the manuscript. All authors read and approved the final version.

**Data availability**

The datasets analyzed during the current study are derived from publicly available sources: the U.S. Census Bureau American Community Survey 2018–2022 five-year estimates and the 2022 TIGER/Line and cartographic boundary files; the U.S. Food and Drug Administration Mammography Facility Database; the CDC PLACES 2022 release; the U.S. Department of Agriculture Economic Research Service 2023 Rural–Urban Continuum Codes; the 2022 CDC/ATSDR Social Vulnerability Index; and OpenStreetMap road-network data, from which driving times were computed using the Open Source Routing Machine. Each source is cited in the Methods.

Observed mammography examination volumes were obtained from the Texas Outpatient Event Public Use Data File (1Q2024–4Q2024), distributed by the Texas Health Care Information Collection (THCIC), Texas Department of State Health Services (DSHS). Under the terms of the THCIC Data Use Agreement the authors may not release, copy, or otherwise grant access to these data to any other person or entity without the written approval of DSHS, and therefore cannot share them directly. The data are, however, obtainable by any researcher on the same terms under which they were acquired for this study, by submitting the Outpatient Event Data Order Form and executing the accompanying Data Use Agreement with DSHS; fees vary by requester type and are waived for Texas state agencies and state universities.

The derived datasets generated during the current study (the estimated facility-level annual screening capacities, the generated candidate-site set, the county–facility drive-time matrix, and the optimization results) are available from the corresponding author on reasonable request.

**Code availability**

The optimization model and the analysis code that support the findings of this study are available from the corresponding author on reasonable request.

**Funding**

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

**Competing interests**

The authors declare no competing interests.

**Supplementary Information**

Supplementary Material accompanies this paper, comprising Sections S1–S4 and Tables S1–S3.

## References

1. National Cancer Institute. SEER Cancer Stat Facts: Female Breast Cancer. Preprint at <https://seer.cancer.gov/statfacts/html/breast.html>.
2. US Preventive Services Task Force. Screening for Breast Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* **331**, 1918–1930 (2024).
3. Wiese, D., Islami, F. & Henry, K. A. Changes in Geographic Accessibility to Mammography by State and Rural-Urban Status, United States, 2006–2022. *J. Natl. Cancer Inst.* **115**, 337–340 (2023).
4. Hashtarkhani, S. *et al.* Analyzing Geospatial and Socioeconomic Disparities in Breast Cancer Screening Among Populations in the United States: Machine Learning Approach. *JMIR Cancer* **11**, e59882 (2025).
5. Lian, M., Struthers, J. & Schootman, M. Comparing GIS-Based Measures in Access to Mammography and Their Validity in Predicting Neighborhood Risk of Late-Stage Breast Cancer. *PLoS One* **7**, e43000 (2012).
6. Ngui, A. N. & Apparicio, P. Optimizing the Two-Step Floating Catchment Area Method for Measuring Spatial Accessibility to Medical Clinics in Montreal. *BMC Health Serv. Res.* **11**, 166 (2011).
7. Luo, W. & Qi, Y. An Enhanced Two-Step Floating Catchment Area (E2SFCA) Method for Measuring Spatial Accessibility to Primary Care Physicians. *Health Place* **15**, 1100–1107 (2009).
8. Hughes, A. E., Lee, S. C., Eberth, J. M., Berry, E. & Pruitt, S. L. Do Mobile Units Contribute to Spatial Accessibility to Mammography for Uninsured Women? *Prev. Med. (Baltim.)* **138**, 106156 (2020).
9. Campbell, J. E. *et al.* Optimizing Mobile Mammography Deployment in Oklahoma 2024: A Two-Step Floating Catchment Area Approach. *Geospat. Health* **21**, 1426 (2026).
10. Stuart, R. M. *et al.* How to Do (or Not to Do) Health Resource Allocations Using Constrained Mathematical Optimization. *Health Policy Plan.* **38**, 122–128 (2023).
11. Barbosa, R., Souza, M. J. F. & Miranda Junior, G. Mammography Unit Location: Reconciling Maximum Coverage and Budgetary Constraints. in *Proceedings of the 25th International Conference on Enterprise Information Systems* vol. 1 187–194 (SciTePress, 2023).
12. Keya, K. N., Islam, R., Pan, S., Stockwell, I. & Foulds, J. Equitable Allocation of Healthcare Resources with Fair Survival Models. in *Proceedings of the 2021 SIAM International Conference on Data Mining* 190–198 (Society for Industrial and Applied Mathematics, 2021). doi:10.1137/1.9781611976700.22.
13. Zhong, Z., Sengupta, R., Paynabar, K. & Waller, L. A. Multi-Objective Allocation of COVID-19 Testing Centers: Improving Coverage and Equity in Access. *IIEE Trans. Healthc. Syst. Eng.* **15**, 46–72 (2025).
14. Hsiao, T.-W., Liao, C.-Y., Waller, L. A. & Paynabar, K. Balancing Access, Precision, and Equity in Adaptive Test Site Allocation with an Application to COVID-19 in Atlanta, Georgia. *Sci. Rep.* **15**, 32775 (2025).
15. Birge, J. R. & Louveaux, F. *Introduction to Stochastic Programming*. (Springer, 2011). doi:10.1007/978-1-4614-0237-4.
16. U.S. Food & Administration, D. Mammography Facility Database. Preprint at <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfMQSA/mqsa.cfm>.
17. Texas Department of State Health Services. Texas Outpatient Event Public Use Data File, 2025. Preprint at <https://www.dshs.texas.gov/chs/thcic/downloadandpurchasedata/outpatientPUDF>.
18. U.S. Census Bureau. Census Data API: 2022 American Community Survey 5-Year Detailed Tables. Preprint at <https://api.census.gov/data/2022/acs/acs5.html> (2022).
19. U.S. Census Bureau. 2022 American Community Survey 5-Year Estimates, Detailed Table B01001: Sex by Age. Preprint at <https://data.census.gov/table/ACS5Y2022.B01001> (2022).

20. Centers for Disease Control & Prevention. PLACES: Local Data for Better Health, County Data 2022 Release. Preprint at <https://data.cdc.gov/500-Cities-Places/PLACES-Local-Data-for-Better-Health-County-Data-20/duw2-7jbt> (2022).
21. U.S. Food & Administration, D. MQSA National Statistics. Preprint at (2026).
22. U.S. Census Bureau. 2022 TIGER/Line Shapefiles. Preprint at <https://www.census.gov/geographies/mapping-files/2022/geo/tiger-line-file.html> (2022).
23. U.S. Department of Agriculture Economic Research Service. 2023 Rural-Urban Continuum Codes. Preprint at <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes> (2024).
24. Centers for Disease Control *et al.* CDC/ATSDR Social Vulnerability Index 2022 Database, Texas. Preprint at <https://www.atsdr.cdc.gov/place-health/php/svi/svi-data-documentation-download.html> (2022).
25. Webster, J. L., Goldstein, N. D., Rowland, J. P., Tuite, C. M. & Siegel, S. D. A Catchment and Location-Allocation Analysis of Mammography Access in Delaware, US: Implications for Disparities in Geographic Access to Breast Cancer Screening. *Breast Cancer Research* **25**, 137 (2023).
26. Ogunwale, W. O., Ajao, O. A. & Ogunbayo, O. E. Mapping the Uneven Landscape: A Geospatial Analysis of Breast Cancer Incidence and Mammography Access Across Illinois (2017–2021). *medRxiv* <https://doi.org/10.1101/2025.06.03.25328939> (2025) doi:10.1101/2025.06.03.25328939.
27. U.S. Census Bureau. 2022 American Community Survey 5-Year Estimates, Detailed Table B03002: Hispanic or Latino Origin by Race. Preprint at <https://data.census.gov/table/ACSDT5Y2022.B03002> (2022).
28. Project OSRM. Open Source Routing Machine. Preprint at <https://project-osrm.org/>.
29. Luxen, D. & Vetter, C. Real-Time Routing with OpenStreetMap Data. in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* 513–516 (ACM, 2011). doi:10.1145/2093973.2094062.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMethods.docx](#)