

Title Page

Title: Forecasting the Diabetes Burden Across USA–Mexico Border States Through 2030: A Multi-Indicator, Multi-Model Analysis

Authors: Kaustubh Wagh,^{1*}, Alexander Kirpich,¹ Ike Okosun,¹ Sunmi Lee,² Sverre-Erik Mamelund,^{3*} Gerardo Chowell,^{1,2*}

Affiliations:

1 Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, GA, 30303, USA

2 Department of Applied Mathematics, Kyung Hee University, Yongin 17104, Korea

3 Centre for Research on Pandemics & Society (PANSOC), Oslo Metropolitan University, Oslo, 0130, Norway

* Corresponding authors:

Kaustubh Wagh kwagh2@student.gsu.edu;

Sverre-Erik Mamelund masv@oslomet.no;

Gerardo Chowell gchowell@gsu.edu

Word Count:

Abstract:

Main text:

Abstract

Objectives: The USA-Mexico border area faces an elevated diabetes burden, with prevalence 2-3 times higher than national averages in both countries. Yet, no forward-looking projections exist beyond 2021, limiting healthcare system planning. We provide the first multi-indicator, multi-model forecasts of diabetes burden through 2030 across border states. **Methods:** Diabetes mortality, prevalence, and disability-adjusted life years (DALYs) data were extracted from the Global Burden of Disease Study 2021 for all ten border states in both countries (1990-2021). Eight different time-series models including ARIMA, generalized additive models, generalized linear models, Prophet, and *n*-sub-epidemic framework were used to project burden from 2022-2030. Models were evaluated and compared using weighted interval scores, with forecasts stratified by age, sex, and geography. **Results:** Multi-model ensembles project 15-17 million people with diabetes, 23,000-26,000 annual deaths, and 1.6-1.8 million Disability-Adjusted Life Years (DALYs) due to diabetes in USA-Mexico border states by 2030. Mexican border states face accelerating mortality among adults aged 20-59 years (15-58% increases), while USA states show prevalence increases of 35-52% in youth and 47-49% in older adults. Males experienced consistently higher projected burden than females. ARIMA demonstrated superior performance for mortality and prevalence forecasting, while ensemble methods performed better for DALYs projections. **Conclusions:** Diabetes burden is projected to increase 30-50% across border states by 2030, driven by distinct regional patterns requiring divergent public health responses. Mexican states require community health worker programs, subsidized medications, and mobile clinics for mortality prevention, while US states must expand school-based screening and nutrition program capacity to address rising youth prevalence. These projections provide

actionable intelligence to support coordinated cross-border efforts in diabetes prevention and management.

Keywords: Diabetes forecasting, USA-Mexico border

Strengths and Limitations of This Study

Strengths

- Multi-model ensemble approach (eight models) with rigorous comparative evaluation using standardized metrics, providing transparency on performance variability across outcomes and demographics.
- Comprehensive stratification by age group (five categories), sex, and geography enables demographically targeted healthcare planning, addressing a gap in prior border health forecasting studies.
- Explicit quantification of forecast uncertainty through 95% prediction intervals and formal calibration assessment, identifying outcome-specific and geographic instances of substantial uncertainty.

Limitation

- Forecasts use statistically modeled disease burden estimates (GBD 2021) rather than primary surveillance data, potentially introducing measurement error and limiting precision in capturing true state-level variation.
- Time-series models assume historical trends continue unchanged, unable to account for policy interventions, technological innovations, or epidemiological shocks. Forecasts are most reliable in short term.

1. Introduction

The USA-Mexico border area faces a high diabetes burden, with a 2-3 times higher prevalence compared to the national average in both Mexico and the United States (1). Ten USA-Mexico border states—four in the USA (California, Arizona, New Mexico, and Texas) and six in Mexico (Baja California, Sonora, Chihuahua, Coahuila, Nuevo León, and Tamaulipas)—form a binational corridor home to over 100 million people (2,3). The border populations experience a diabetes prevalence of approximately 15.7% compared to 8.2% nationally for Mexico and 5.2% for the United States (1). With mortality rates reaching nearly 50% higher than the national averages, diabetes represents a critical and persistent public health challenge in the region (1,4,5). Border states share distinct epidemiological characteristics: higher concentrations of Hispanic/Latino populations, high poverty rates, and lower health insurance coverage compared to the national averages (6,7), alongside limited healthcare infrastructure, genetic susceptibility (e.g., TCF7L2 polymorphisms among Hispanic/Latino populations), and environmental determinants such as limited access to nutritious food, and the built environment for physical activity (8–10). This unique profile, characterized by cross-border mobility and binational health challenges, demands cross-state and country-level analysis to inform regional policy and resource allocation. Accurate and timely projections of future diabetes burden are therefore essential for state health system planning and binational coordination.

Despite extensive research on diabetes at national levels (11–13), critical gaps persist in understanding border states' specific trajectory. To date, most diabetes forecasting studies are focused on national or global projections (14–16), with only a few examining state-level patterns in border regions. Existing forecasting studies rely on single indicators such as prevalence or mortality, limiting comprehensive burden assessment.(17,18). Moreover, few studies provide

age-sex stratification for state-level interventions or compare multiple forecasting methods to quantify prediction uncertainty. While national analyses documents rising diabetes burden (17), these estimates often obscure within and between-country variation, particularly the elevated burden in border states. Our previous work documented consistently elevated historical trends in deaths, prevalence, and disability-adjusted life years (DALYs) across border states from 1990-2021, revealing consistently higher burden compared to national averages (19). However, no projections are available for these specific border states beyond 2021, leaving state policymakers without any future estimates needed to prepare healthcare systems for the coming decades. This work aims to address that and to our knowledge, this is the first state-level, multi-indicator, multi-model diabetes burden forecasting study for the USA–Mexico border region that integrates diabetes mortality, prevalence, and DALYs. This study advances prior work by comparing multiple modeling frameworks and to produce age- and sex-stratified forecasts with quantified uncertainty.

Our study employs state-level data from all ten border states (six Mexico and four USA) as the appropriate geographic unit for several reasons. First, state-level age–sex–stratified diabetes data are consistently available across decades from the Global Burden of Disease Study (20,21), ensuring comparability across countries. Second, state boundaries represent defined functional and administrative units through which health policy is formulated, healthcare systems are organized, and budgets are allocated, making state-level projections directly actionable (22). Third, the epidemiological influence of the border is broad and extends throughout border states due to population mobility, economic integration, and cultural continuity that transcends proximity to the international boundary (23,24). It is also worth emphasizing that border states exhibit distinct diabetes risk profiles compared to interior states in both countries. Fourth, state-

level analysis provides sufficient data for robust age-sex stratified forecasting while enabling meaningful binational comparisons (25). Finally, sister-state relationships and regional health commissions operate at the state level, making state-level burden estimates critical for binational coordination and future collaborations. Thus, state-level analysis aims to provide optimal balance between administrative relevance, data quality, and forecasting reliability for informing border health policy.

We sought to forecast diabetes burden in all ten USA-Mexico border states from 2022 to 2030 using three complementary indicators: deaths (fatal burden), prevalence (healthcare demand), and DALYs (comprehensive disease impact combining premature mortality and disability) (26). Specific objectives include: generating projections for Mexico and USA border states to identify which region face the greatest future burden in terms of numbers; producing age-sex stratified forecasts to inform demographically targeted interventions; comparing eight time-series forecasting models (ARIMA, generalized additive models, generalized linear models, Prophet, and *n*-subepidemic models) to assess prediction uncertainty and identify optimal approaches for forecasting (27,28); evaluating models calibration through coverage analysis to ensure projection reliability; and comparing projected burden between U.S. and Mexican border states to inform binational coordination. These forecasts provide evidence-based estimates to support planning, binational health coordination, and targeted diabetes prevention policies.

2. Methods

2.1. Data Collection and Preparation

We obtained data for this study from the Global Burden of Disease Study 2021 (GBD2021), a publicly accessible database maintained by the Institute for Health Metrics and Evaluation (IHME) (29). We extracted diabetes-related mortality (both type 1 and type 2), prevalence, and

disability-adjusted life years (DALYs) for all USA-Mexico border states from 1990 to 2021 (yearly). Border states included California, Arizona, New Mexico, and Texas on the US side, and Baja California, Sonora, Chihuahua, Coahuila, Nuevo León, and Tamaulipas on the Mexican side. We stratified all estimates by age group and sex to enable related demographic comparisons. We aggregated data for six Mexico states (referred to as MX), four USA states (referred to as US) and all ten USA-Mexico border states (referred to as USMX). The GBD 2021 data incorporate uniform and consistent methods for data collection, processing, and statistical modeling to ensure comparability across locations and time periods (30). Publicly available, de-identified aggregate data were used; institutional review board approval was not required. All data use agreements and terms of service were followed.

We used all 32 years of (yearly) historical data and set the calibration period to 32 years and the forecasting horizon between 2022 and 2030 (yearly forecasting for 9 years). This nine-year horizon aligns with the 2030 Sustainable Development Goal (SDG) targets (31), ensuring policy-relevant projections for the current global health agenda. Data was preprocessed using Microsoft Excel pivot tables. All forecasting analyses were conducted using RStudio, Posit Software, Boston, MA, USA (32) and MATLAB, MathWorks Inc., Natick, MA, USA (33). Specifically, we used the StatModPredict dashboard (developed in RStudio v. 4.5.0) to generate forecasts from the ARIMA, GAM, GLM, and Prophet models (34). For the ensemble n-sub-epidemic modeling framework, we employed the SubEpiPredict toolbox (SubEpiPredict toolbox) implemented in MATLAB (R2025) (35). Both toolboxes support reproducible workflows, automated parameter selection, and standardized output generation.

2.2. Forecasting Models

2.2.1. ARIMA (Autoregressive Integrated Moving Average)

ARIMA is a time series model that builds forecasts by identifying patterns in past diabetes burdens (deaths, prevalence, DALYs) data (36). It combines three components: trends over time (autoregression), smoothing of random fluctuations (moving average), and differencing to stabilize the data to ensure necessary stationarity. ARIMA works the best for short-term forecasting when historical trends are stable, linear and non-seasonal (28), while it is less effective when periodic patterns shift suddenly due to interventions or emerging risk factors (28).

2.2.2. GAM (Generalized Additive Model)

GAMs extend traditional regression models that allow for flexible, nonlinear relationships between time and diabetes burdens (deaths, prevalence, DALYs) (37). Instead of fitting a straight line, GAMs use smooth curves to model long-term trends in diabetes burdens (38) which make such models well suited for capturing gradual changes in diabetes burdens that may result from evolving lifestyle behaviors, treatment access, or demographic shifts.

2.2.3. GLM (Generalized Linear Model)

GLMs are classic statistical models used to relate a discrete response variable (like mortality count) to one or more predictors (such as time). They assume a specific linear relationship between the predictors and the outcome, transformed by a so-called link function, and use a distribution suited for the type of data used in the analysis (e.g., Poisson or normal) (39). GLMs have straightforward formulations and easily interpretable and effective when trends are relatively smooth.

2.2.4. Prophet (Developed by Meta/Facebook)

Prophet is a decomposable time series model that breaks down diabetes burdens' (deaths, prevalence, DALYs) trends into three components: an overall trend, seasonal effects (e.g., annual patterns), and abrupt changes due to major events or interventions (40). It is available as a package in R and specifically designed to handle missing data and outliers well (41). Prophet is especially useful for longer-term forecasting where trends may shift, making it helpful for identifying inflection points in chronic disease burden.

2.2.5. n-Sub-Epidemic Model

This model interprets diabetes burdens' (deaths, prevalence, DALYs) trends as a series of overlapping "sub-epidemics", each representing a surge or phase in diabetes deaths (42,43). Each sub-epidemic has its own timing and intensity, allowing the overall model to capture complex mortality dynamics such as plateaus or multiple peaks often seen in real-world chronic disease progression. It is particularly powerful in reflecting regional differences or the cumulative effects of repeated disruptions.

2.2.6. Ensemble Sub-Epidemic Models (Unweighted and Weighted)

These models combine forecasts from the top-performing sub-epidemic models. The unweighted ensemble gives equal importance to each model, while the weighted version gives more influence to models with better fit (44). This ensemble approach improves robustness by balancing strengths across models and helps produce more reliable forecasts with quantified uncertainty. Under the n-sub-epidemic modeling framework, we developed several models, including the n-subepidemic ensemble (of top 2 ranked models) unweighted [NSE-ensemble(2)(UW)], the n-subepidemic ensemble (of top 2 ranked models) weighted [NSE-ensemble(2)(W)], and the top two individually ranked models [NSE-ranked(1) and NSE-

ranked(2)]. The ensemble models [NSE-ensemble(2)(UW), and NSE-ensemble(2)(W)] are superior to individual ranked models [NSE-ranked(1), and NSE-ranked(2)] as they improve forecasting performance by systematically incorporating the predictive accuracy of individual models (44). Although the utilized *n*-sub-epidemic framework was initially validated primarily on infectious-disease time series, its flexible structure allows application to chronic disease burden data, where overlapping sub-trends may reflect multiple epidemiological drivers or intervention phases. However, for transparency and to show the individual models that informed our ensemble approaches, we also presented the results of NSE-ranked (1) and NSE-ranked (2) models.

2.3. Model Evaluation

Models were evaluated using a set of performance metrics that measure both point and probabilistic forecast accuracy simultaneously (44). In particular, the utilized models were evaluated based on Weighted Interval Score (WIS) and mean absolute error metrics, which quantify the deviation between predicted and the observed values for each given point in time and aggregated across time points. We used data from 2013 to 2021 (9 years) to measure the performance of models. In this context, smaller values of those metrics indicate better model performance. The reliability of uncertainty estimates was evaluated based on the proportion of observations falling in the predicted intervals, with 95% P.I. coverage where higher values ideally estimate better calibrated uncertainty (45). It is important to note that WIS provides a balance between sharp and out-of-range predictions due to penalizing prediction intervals that are too wide and observations that fall outside predicted range, thus sustaining an overall balance between forecast sharpness and calibration (42).

2.4. Quantifying Forecast

The forecasts for each model were numerically quantified. i.e. the percentage change formula was used to calculate the relative change (%) in the number of diabetes-related burdens (i.e. deaths, prevalence, DALYs) between 2021 and 2030. The exact formula was $(\text{Forecast (2030)} - \text{Observed (2021)}) / \text{Observed (2021)} \times 100$. Here, forecast (2030) refers to the forecasted number of diabetes-related burdens in 2030 by a given model while observed (2021) refers to the observed number of diabetes-related burdens in 2021, which is the last year of available data. The overall diabetes-related burdens in 2030 were calculated by first calculating the mean forecast across all models.

Forecasts from multiple models were combined using weighted and unweighted averaging approaches. For weighted averaging, model weights were derived from the Weighted Interval Score (WIS), a proper scoring rule that evaluates both point prediction accuracy and prediction interval calibration. Weights were computed as the inverse values of WIS i.e. $1/\text{WIS}$ and normalized to sum to one within each geography-category stratum. Weighted forecasts were calculated by multiplying each model's prediction by its normalized weight and summing across all models. For unweighted averaging, all models contributed equally (weight = $1/n$, where n is the number of models). Prediction intervals (95% PI) were derived by aggregating the lower and upper bounds from individual models using the corresponding weighting scheme. Both methods were applied to generate forecasts separately for deaths, prevalence, and DALYs stratified by geography, sex, and age group.

3. Results

3.1 Performance Metrics

3.1.1 Deaths

Mortality forecasting showed uniformly strong performance with minimal model differentiation. ARIMA, GAM, and GLM achieved comparable accuracy (MAE 2.3-3.1, WIS 2.0-3.0) across all demographic strata and geographic regions (Fig 1A, Supplementary Fig S1A). Ensemble methods showed no systematic improvement. Error magnitudes remained consistent across Mexican border states (MAE 2.5-3.0), US border states (MAE 2.8-3.3), and the combined USMX region. Model rankings were stable across age and sex strata.

3.1.2 Prevalence

Prevalence forecasting revealed marked model performance differences among models with substantially higher errors than mortality outcomes. ARIMA achieved the lowest MAE values, ranging from 3.2-4.1 in Mexican border states to 4.5-5.0 in US border states, with WIS values following similar patterns (Fig 1B, Supplementary Fig S1B). GAM showed comparable performance. Prophet and ensemble approaches (NSE-Ensemble-Ranked, NSE-Ensemble-UW/W) consistently ranked lowest, with MAE values 15-40% higher than ARIMA across most age-sex combinations. The performance gap between ARIMA and ensemble methods was largest in Mexican border states (MAE difference 1.0-1.5 units) and smallest in US border states (MAE difference 0.5-0.8 units).

3.1.3 DALYs

DALY forecasts showed error magnitudes intermediate between deaths and prevalence. ARIMA and GAM maintained lower errors (MAE 2.7-4.3, WIS 2.4-4.2) compared to Prophet and ensemble variants (MAE 3.5-5.2, WIS 3.8-5.0) (Fig 1C, Supplementary Fig S1C). The youngest

age groups (<20 years) achieved MAE values 40-50% lower than older populations (60+ years) across all models. Male and female strata produced nearly identical error magnitudes and model rankings within each geographic region for all three outcomes.

3.2 Calibration Metrics

3.2.1 Deaths

Death forecasts demonstrated the steepest age-dependent performance gradient across all outcomes (Figure 2A, Supplementary Fig S2A). Younger populations achieved lower WIS score and coverage of 95% PI closer to 95% (i.e. better calibration), with models showing WIS values of -0.7 to 0.6 log₁₀ units and coverage of 84-100% for ages <20 years. However, GLM critically failed in middle-aged males (40-59 years), achieving only 46% coverage compared to >92% for ensemble methods. Regional disparities emerged in older populations, where Mexico consistently maintained superior coverage (>92%) across all models, while US border states required ensemble approaches to achieve adequate calibration.

3.2.2 Prevalence

Prevalence forecasts exhibited the widest performance variability and most severe model failures (Figure 2B, Supplementary Fig S2B). NSE-Ranked (2) catastrophically under-covered predictions in US border states for both sexes (7.69% coverage), despite acceptable performance elsewhere. This outcome showed pronounced geographic stratification: Mexico border states maintained robust calibration in younger populations (WIS: 0.5-2.9 log₁₀ units, coverage >85%), whereas US border states struggled across all age groups, with older populations (>60 years) requiring ensemble methods to achieve minimally adequate coverage.

3.2.3 DALYs

DALY forecasts revealed ensemble superiority across demographic strata (Figure 2C, Supplementary Fig S2C). NSE-Ensemble(2)(W) maintained stable performance in USMX both-sex populations (WIS: 2.5-3.3 log₁₀ units, coverage >92%), while GAM showed age-dependent variability—excellent calibration in youth (<20 years: WIS 1.4, coverage 100%) but substantial degradation in older groups. Sex-stratified analyses showed males experienced moderately elevated WIS in young adults (20-39 years: 3.2-3.7 versus 2.8-3.4 log₁₀ units for females). Mexico and US border states diverged primarily in middle-aged populations, where GLM coverage dropped to 46-54% in US regions.

3.3 Forecasting

3.3.1 Deaths

Diabetes mortality forecasts from 2021-2030 varied substantially by geography (Fig 3A). MX showed the highest increases in younger adults (20-39 years: GAM 57.8%, ARIMA 14.7%) and middle-aged adults (40-59 years: GAM 40.5%, ARIMA 33.4%). US demonstrated decreases in younger populations (20-39 years: GAM -35.0%, ARIMA -10.0%) but mixed patterns in older ages. USMX showed consistent increases across all age groups (ARIMA 10-21%, GAM 10-49%). Males experienced higher mortality increases than females across all geographies (MX males: ARIMA 28.4% vs females 15.3%; US males: 21.6% vs females 3.2%). NSE models predicted mortality reductions in older US adults (60-79 years: NSE-ranked (2) -21.6%).

Forecasted mortality projections varied by weighting approach (Fig 4). WIS-weighted forecasts projected 23,156 deaths (95% PI: 20,525-25,845) across the combined USA-Mexico border region, while unweighted forecasts estimated 26,419 deaths (95% PI: 22,472-30,528). Geographic distribution showed Mexico forecasted at 13,943 deaths (95% PI: 11,995-15,923)

under weighted and 15,959 deaths (95% PI: 13,578-18,458) under unweighted approaches, with the United States at 12,632 (95% PI: 10,263-15,269) and 13,421 deaths (95% PI: 10,772-16,309), respectively. Males were consistently projected to experience higher mortality than females across both methods: 14,066 (95% PI: 12,556-15,631) versus 12,050 deaths (95% PI: 10,622-13,520) for weighted, and 15,368 (95% PI: 13,155-17,716) versus 14,012 deaths (95% PI: 11,896-16,253) for unweighted forecasts. Age-specific projections demonstrated strong concordance between methods, with the 60-79 age group expected to bear the highest burden under both weighted (13,012 deaths) and unweighted (14,848 deaths) approaches, followed by the 80+ and 40-59 age groups. Model-specific projected deaths by 2030 are shown in Table 1, with forecasted trajectories of deaths in Mexico and USA border states are presented in Supplementary Figs S3.1 and S3.2, respectively.

3.3.2 Prevalence

Universal prevalence increases characterized most forecasts, but geographic patterns varied considerably (Fig 3B). US showed the steepest rises in youth (<20 years: 35-52%) and older adults (60-79 years: 47-49%), while MX exhibited modest youth increases (11-13%) but substantial middle-aged burden (34% for 40-59 years). USMX demonstrated 22-45% increases across most ages. All geographies showed similar increases among those 80+ years (30-37%). Sex differences remained minimal across all regions, typically under 6 percentage points, with males showing marginally higher prevalence growth than females.

Prevalence forecasts showed substantial differences between weighting methodologies (Fig 4). WIS-weighted forecasts predicted 14.9 million cases (95% PI: 13.0-17.0 million) across the USA-Mexico border region, whereas unweighted forecasts projected higher prevalence at 17.4 million cases (95% PI: 14.8-20.2 million). Geographic patterns remained consistent across

methods, with the United States accounting for the majority: 12.5 million cases (95% PI: 10.4-14.8 million) weighted versus 14.9 million cases (95% PI: 12.5-17.5 million) unweighted, compared to Mexico's 2.5 million (95% PI: 2.1-2.9 million) weighted and 2.9 million (95% PI: 2.4-3.3 million) unweighted. Sex-stratified projections indicated males with slightly higher prevalence than females under both approaches: 7.8 million (95% PI: 6.7-8.9 million) versus 7.1 million (95% PI: 6.2-8.2 million) for weighted, and 9.1 million (95% PI: 7.7-10.5 million) versus 8.3 million (95% PI: 7.1-9.6 million) for unweighted forecasts. The 60-79 age group was consistently projected to demonstrate the highest disease burden across both methodologies, followed by the 40-59 age group. Model-specific projected prevalence by 2030 is shown in Table 2, with forecasted trajectories of prevalence in Mexico and USA border states are presented in Supplementary Figs S4.1 and S4.2, respectively.

3.3.3 DALYs

Geographic heterogeneity in DALYs projections was pronounced (Fig 3C). USMX exhibited the highest burden increases among youth and young adults (46% for both age groups), while MX showed near-zero or negative changes in these populations. US demonstrated substantial DALYs increases in the oldest adults (80+ years: 33-46%), contrasting with lower increases in MX (7-10%) for the same age group. Middle-aged adults showed wide geographic variation, with USMX predicting 27-43% increases versus 3-6% in MX. Males experienced 4-10 percentage point greater DALYs increases than females in USMX and US.

Disability-adjusted life year projections exhibited notable variation between weighting schemes (Fig 4). WIS-weighted forecasts estimated 1.6 million DALYs (95% PI: 1.4-1.8 million) across the USA-Mexico border region, while unweighted forecasts projected 1.8 million DALYs (95% PI: 1.5-2.1 million). Geographic contributions differed by method, with the United States

projected at 1.1 million DALYs (95% PI: 949,000-1.3 million) weighted versus 1.3 million (95% PI: 1.1-1.5 million) unweighted, and Mexico at 531,000 (95% PI: 451,000-617,000) weighted versus 607,000 (95% PI: 516,000-704,000) unweighted. Sex-specific projections consistently showed males accumulating higher DALYs than females: 857,000 (95% PI: 739,000-982,000) versus 716,000 (95% PI: 617,000-822,000) for weighted, and 990,000 (95% PI: 841,000-1.1 million) versus 833,000 (95% PI: 708,000-966,000) for unweighted forecasts. Both methodologies revealed similar age patterns, with peak burden projected in younger populations (20-39 and <20 age groups), reflecting substantial years of life lost, while older age groups (60-79 and 80+) showed comparatively lower DALY estimates. Model-specific projected deaths by 2030 are shown in Table 3, with forecasted trajectories of DALYs in Mexico and USA border states are presented in Supplementary Figs S5.1 and S5.2, respectively.

Across all outcomes, ARIMA and GAM consistently produced the most reliable forecasts, demonstrating strong point prediction accuracy and similarities across demographic groups. In contrast, weighted ensemble approaches, particularly WIS-weighted models, showed much larger uncertainty in calibration and coverage performance. Together, these results highlight that no single model dominated across all indicators, emphasizing the importance of combining complementing statistical methods and model ensembles to balance accuracy and reliability in diabetes burden forecasting.

4. Discussion

This study provides the first comprehensive state-level forecast of the diabetes burden across all 10 USA-Mexico border states through 2030. Our multi-model ensemble approach projects approximately 23,000-26,000 diabetes deaths, 15-17 million people with diabetes, and 1.6-1.8

million DALYs annually by 2030 across the ten USA-Mexico border states, with highlighted disparities between Mexican and US border states which reflect distinct epidemiological trajectories. Overall, our findings reveal the urgent need for coordinated, binational health strategies to address the escalating diabetes burden in this region.

Our forecasts reveal three critical epidemiological patterns shaping the future diabetes burden across the border region. First, Mexican border states face accelerating diabetes mortality, particularly among younger and middle-aged adults (20-59 years), with projected increases of 15-58% by 2030. This contrasts sharply with US border states, where mortality is projected to decline in younger populations but increase among older adults. This divergence likely reflects Mexico's ongoing epidemiological transition, characterized by rising obesity prevalence, delayed diabetes diagnosis, and limited access to glycemic control medications (12,46). The USA patterns suggests successful prevention efforts in younger cohorts but accumulated disease burden in aging populations with longer diabetes duration (47).

Second, prevalence projections uniformly indicate substantial increases across all border states, but with markedly steeper growth in US border states (35-52% in youth, 47-49% in older adults) compared to their Mexican counterpart. This pattern may reflect the expected improved survival with diabetes in the US through better treatment access, paradoxically increasing prevalence despite stabilizing incidence (48). The pronounced youth prevalence increase in US border states align with rising childhood obesity and type 2 diabetes onset at younger ages, a phenomenon extensively documented in Hispanic border communities (49,50).

Third, DALY projections highlight large increases in years of life lost among youth in USA border states, while Mexican states show stable or slightly declining DALYs in younger groups. This geographic heterogeneity with USMX projecting 46% increases in youth while Mexican

states show near-zero changes suggests differential impacts of early-onset diabetes and contrasting healthcare effectiveness on different sides of the border (51).

Our findings extend and complement previous studies in several important ways. National-level projections by Lin et al. (2018) estimated US diabetes prevalence reaching 54.9 million by 2030 (11), while our border state forecasts (12.5-14.9 million) thus representing 23-27% of this expected national burden despite border states comprising only 15% of the US population, confirming the disproportionate regional impact. Similarly, Meza et al. (2015) projected Mexican national diabetes prevalence at 13.7 million by 2030 (52), with our border state expected estimates (2.5-2.9 million) representing 18-21% of the national total while border states contain approximately 20% of Mexico's population, suggesting burden comparable to population distribution in Mexico while still elevated in absolute rate values. What is crucial is that our study addresses a few methodological gaps in previous forecasting research since most prior studies employed single models (i.e. typically ARIMA or exponential smoothing) without quantifying prediction uncertainty through detailed model comparisons (53,54). Our ensemble approach revealed substantial differences in model's performance, particularly for prevalence forecasting where ARIMA outperformed ensemble methods by 15-40%. These findings challenge assumptions that ensemble averaging universally improves forecast accuracy (38) and suggests outcome-specific model selection strategies may be necessary and should not be discarded. The catastrophic under-coverage that was observed with certain models in US border states [7.69% for NSE-ranked (2)] underscores the importance of rigorous calibration assessment beyond point estimate accuracy. While ARIMA and GAM offered strong interpretability and transparent trend estimation, ensemble models provided greater flexibility and improved uncertainty calibration. Recognizing that the optimal model may vary by outcome type (i.e.,

mortality, prevalence, or DALYs) represents an important methodological contribution of this study, supporting an outcome-dependent approach to model selection.

Our age and sex stratified forecasts also address a critical limitation in existing border health research. Previous border diabetes studies provided descriptive epidemiology (55) or intervention evaluations (56) but lacked forward-looking projections with the desired level of demographic stratification details necessary for healthcare capacity planning. The pronounced sex differences identified in the analysis including males experiencing 4-10 percentage point greater increases in mortality and DALYs align with documented gender disparities in diabetes self-management, healthcare utilization, and cardiovascular complications among Hispanic populations (51).

Since GBD 2023 data were released after our initial analysis, we compared our projections with these independent estimates as a validation check. Our multi-model ensemble forecast 11.6 million (95% PI: 10.7–12.5 million) people with diabetes in the USA-Mexico Border (USMX) region for 2023, which closely aligned with the GBD estimate of 11.4 million (95% uncertainty interval: 10.6–12.3 million). Similarly, our forecast of 13–15 thousand diabetes-related deaths in six Mexico border states (MX) in 2023 was consistent with the GBD estimate of 13–16 thousand deaths.

4.1 Public Health and Healthcare System Implications

The diabetes projections for the USA-Mexico border highlighted in this study reveal a pressing public health issue that requires urgent attention. First, it is essential to develop a coordinated binational health policy and allocate resources effectively. Prioritizing initiatives that promote diabetes health by supporting weight management, encouraging regular physical activity, and promoting a balanced diet rich in fiber, fruits, vegetables, and whole grains are warranted.

Second, Mexican border states require immediate strengthening of primary care infrastructure for diabetes screening and early treatment. The projected mortality increase among working-age adults (20-59 years) is expected to generate substantial economic losses through premature death and disability (57). Targeted interventions should include community health worker programs, which have demonstrated effectiveness in Hispanic border communities (58), and improved access to metformin, statins, and antihypertensive medications through public pharmacy networks. Third, US border states must expand healthcare capacity to be ready to accommodate 35-52% prevalence increases in youth populations. This requires school-based screening programs, pediatric endocrinology workforce expansion, and integration of behavioral health services to address psychosocial barriers to diabetes self-management in Hispanic youth (59). The projected burden which is concentrated in the 60-79 age group necessitates geriatric care infrastructure investments and age-appropriate disease management programs. Four, binational coordination mechanisms, including the USA-Mexico Border Health Commission and sister-state partnerships, should consider utilizing these state-specific projections to align policy priorities and share evidence-based interventions. In particular, the divergent epidemiological trajectories we identified suggest opportunities for bidirectional learning, with Mexico potentially adopting similar policies and benefiting from US prevention successes in younger populations while the US could adopt Mexico's community-based care models that maintain a lower DALY burden in older adults despite resource constraints (55). By taking the above actions, we can significantly enhance diabetes health and outcomes for communities on both sides of the border.

4.2 Limitations

Several limitations warrant consideration. First, forecasts assume historical trends continue without accounting for potential policy interventions, technological innovations (e.g. improved medications and treatment), or epidemiological shocks that could alter trajectories (60). Second, the utilized GBD estimates rely on statistical modeling rather than comprehensive surveillance systems in some border states, potentially introducing measurement error (26). Third, the GBD model estimates themselves are not derived from primary surveillance data at the state level, which may limit precision in capturing substate variations. Fourth, our models do not incorporate diabetes type differentiation (i.e. Type 1, and Type 2), fertility and migration patterns, or cross-border healthcare utilization factors affects border regions (61). Fourth, forecasting performance varied substantially by outcome (data type) and geography (state), with prevalence projections in US border states showing wider prediction intervals than mortality forecasts. Finally, our study period ended in 2021, precluding assessment of potential pandemic-related disruptions to diabetes care and screening that may influence future burden (62).

5. Conclusion

Diabetes burden across USA-Mexico border states will increase 30-50% by 2030, with divergent epidemiological trajectories requiring distinct binational responses. Mexican border states demand immediate mortality prevention through expanded treatment access, while US border states must prepare healthcare infrastructure for increases in youth prevalence. These evidence-based projections provide actionable intelligence for resource allocation, healthcare system strengthening, and binational policy coordination to mitigate the escalating diabetes crisis throughout the joint border region. Beyond diabetes, the utilized forecasting framework can be

applied to other chronic diseases and cross-border settings to support evidence-based planning and progress toward global health targets.

Acknowledgments:

The authors would like to thank Hamid Karami, Georgia State University, for sharing his valuable expertise in MATLAB coding through hemera.rs.gsu.edu (interactive online entry point to advanced research computing) which greatly contributed to the success of this research.

Author Contributions:

KW: design, analysis, interpretation, drafting, and editing; AK: review, and editing; IO: review, and editing; SL: review; SEM: review, and editing; GC: conception, supervision, review, and editing.

Conflict of Interest Statement:

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

Data Availability Statement:

The datasets generated and analyzed during the current study are available publicly as described in methods section.

Source of Funding:

The authors received no funding from an external source.

References

1. PAHO. Along U.S. - Mexico Border, Diabetes Cases Are the “Tip of the Iceberg” - PAHO/WHO | Pan American Health Organization [Internet]. 2010 [cited 2025 Aug 15]. Available from: <https://www.paho.org/en/news/21-10-2010-along-us-mexico-border-diabetes-cases-are-tip-iceberg>
2. US Census Bureau. Census.gov. 2024 [cited 2025 Sept 30]. State Population by Characteristics: 2020-2024. Available from: <https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-detail.html>
3. CONAPO. Population projections, five-year [Internet]. Mexico City. Mexico: General Secretariat of the National Population Council (CONAPO); 2025 [cited 2025 Aug 12]. Available from: <https://www.datos.gob.mx>
4. Fisher-Hoch SP, Rentfro AR, Salinas JJ, Pérez A, Brown HS, Reininger BM, et al. Socioeconomic status and prevalence of obesity and diabetes in a Mexican American community, Cameron County, Texas, 2004-2007. *Prev Chronic Dis*. 2010 May;7(3):A53.
5. Ingram M, Gallegos G, Elenes J. Diabetes Is a Community Issue: The Critical Elements of a Successful Outreach and Education Model on the U.S.-Mexico Border. *Prev Chronic Dis*. 2004 Dec 15;2(1):A15.
6. kffdrishti. Health and Health Care in the U.S.-Mexico Border Region [Internet]. KFF. 2022 [cited 2025 Oct 9]. Available from: <https://www.kff.org/racial-equity-and-health-policy/health-and-health-care-in-the-u-s-mexico-border-region/>
7. Velasco-Mondragon E, Jimenez A, Palladino-Davis AG, Davis D, Escamilla-Cejudo JA. Hispanic health in the USA: a scoping review of the literature. *Public Health Rev*. 2016 Dec 7;37(1):31.
8. Rangel Gómez MG, López Jaramillo AM, Svarch A, Tonda J, Lara J, Anderson EJ, et al. Together for Health: An Initiative to Access Health Services for the Hispanic/Mexican Population Living in the United States. *Front Public Health* [Internet]. 2019 Sept 24 [cited 2025 Oct 9];7. Available from: <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2019.00273/full>
9. Cabral J, Cuevas AG. Health inequities among Latinos/Hispanics: Documentation status as a determinant of health. *J Racial Ethn Health Disparities*. 2020 Oct;7(5):874–9.
10. Palmer ND, Lehtinen AB, Langefeld CD, Campbell JK, Haffner SM, Norris JM, et al. Association of TCF7L2 Gene Polymorphisms with Reduced Acute Insulin Response in Hispanic Americans. *J Clin Endocrinol Metab*. 2008 Jan 1;93(1):304–9.

11. Ong KL, Stafford LK, McLaughlin SA, Boyko EJ, Vollset SE, Smith AE, et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *The Lancet*. 2023 July 15;402(10397):203–34.
12. Basto-Abreu A, Barrientos-Gutiérrez T, Rojas-Martínez R, Aguilar-Salinas CA, López-Olmedo N, De la Cruz-Góngora V, et al. [Prevalence of diabetes and poor glycemic control in Mexico: results from Ensanut 2016.]. *Salud Publica Mex*. 2020;62(1):50–9.
13. Lin X, Xu Y, Pan X, Xu J, Ding Y, Sun X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Sci Rep*. 2020 Sept 8;10(1):14790.
14. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of Type 2 Diabetes – Global Burden of Disease and Forecasted Trends. *J Epidemiol Glob Health*. 2020 Mar;10(1):107–11.
15. Huang X, Wu Y, Ni Y, Xu H, He Y. Global, regional, and national burden of type 2 diabetes mellitus caused by high BMI from 1990 to 2021, and forecasts to 2045: analysis from the global burden of disease study 2021. *Front Public Health* [Internet]. 2025 Jan 23 [cited 2025 Sept 9];13. Available from: <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2025.1515797/full>
16. He KJ, Wang H, Xu J, Gong G, Liu X, Guan H. Global burden of type 2 diabetes mellitus from 1990 to 2021, with projections of prevalence to 2044: a systematic analysis across SDI levels for the global burden of disease study 2021. *Front Endocrinol* [Internet]. 2024 Nov 8 [cited 2025 Sept 2];15. Available from: <https://www.frontiersin.org/journals/endocrinology/articles/10.3389/fendo.2024.1501690/full>
17. IDF. Diabetes Atlas. 2024 [cited 2025 May 6]. Diabetes Atlas. Available from: <https://diabetesatlas.org/>
18. Antonio-Villa NE, Bello-Chavolla OY, Fermín-Martínez CA, Ramírez-García D, Vargas-Vázquez A, Basile-Alvarez MR, et al. Diabetes subgroups and sociodemographic inequalities in Mexico: a cross-sectional analysis of nationally representative surveys from 2016 to 2022. *Lancet Reg Health – Am* [Internet]. 2024 May 1 [cited 2025 Sept 12];33. Available from: [https://www.thelancet.com/journals/lanam/article/PIIS2667-193X\(24\)00059-0/fulltext](https://www.thelancet.com/journals/lanam/article/PIIS2667-193X(24)00059-0/fulltext)
19. Wagh K, Harding JL, Salinas CAA, Kanzler KE, Chowell G. Closing the Gap: A Binational Analysis of Diabetes Mortality and Disability across the USA-Mexico Border Region [Internet]. *medRxiv*; 2025 [cited 2025 Nov 11]. p. 2025.10.28.25337683. Available from: <https://www.medrxiv.org/content/10.1101/2025.10.28.25337683v1>

20. Murray CJ, Ezzati M, Flaxman AD, Lim S, Lozano R, Michaud C, et al. GBD 2010: design, definitions, and metrics. *The Lancet*. 2012 Dec 15;380(9859):2063–6.
21. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019 - *The Lancet* [Internet]. [cited 2025 Oct 9]. Available from: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30752-2/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30752-2/fulltext)
22. Border Health Commission USM. Healthy Border [Internet]. 2024 Mar [cited 2025 Sept 2]. Available from: <https://www.hhs.gov/about/agencies/oga/about-oga/what-we-do/international-relations-division/americas/border-health-commission/healthy-border/index.html>
23. Barquera S, Schillinger D, Aguilar-Salinas CA, Schenker M, Rodríguez LA, Hernández-Alcaraz C, et al. Collaborative research and actions on both sides of the US-Mexico border to counteract type 2 diabetes in people of Mexican origin. *Glob Health*. 2018 Aug 22;14(1):84.
24. Homedes N, Ugalde A. Globalization and Health at the United States–Mexico Border. *Am J Public Health*. 2003 Dec;93(12):2016–22.
25. Seiglie JA, Franco RR, Wirtz VJ, Meigs JB, Mendoza MA, Miranda JJ, et al. Regional and state-level patterns of type 2 diabetes prevalence in Mexico over the last three decades. *Diabetes Res Clin Pract*. 2021 July 1;177:108927.
26. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*. 2020 Oct 17;396(10258):1204–22.
27. Sun J, Hu W, Ye S, Deng D, Chen M. The Description and Prediction of Incidence, Prevalence, Mortality, Disability-Adjusted Life Years Cases, and Corresponding Age-Standardized Rates for Global Diabetes. *J Epidemiol Glob Health*. 2023 July 3;13(3):566–76.
28. Hyndman RJ, Athanasopoulos G. *Forecasting: Principles and Practice*. 3rd ed. edition. Melbourne, Australia: Otexts; 2021. 442 p.
29. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2021 (GBD 2021) Results [Internet]. Seattle, United States: Institute for Health Metrics and Evaluation (IHME); 2024 [cited 2025 Aug 12]. Available from: <https://vizhub.healthdata.org/gbd-results>
30. Schumacher AE, Kyu HH, Aali A, Abbafati C, Abbas J, Abbasgholizadeh R, et al. Global age-sex-specific mortality, life expectancy, and population estimates in 204 countries and territories and 811 subnational locations, 1950–2021, and the impact of the COVID-19 pandemic: a comprehensive demographic analysis for the Global Burden of Disease Study 2021. *The Lancet*. 2024 May 18;403(10440):1989–2056.

31. SDG Target 3.4 | Noncommunicable diseases and mental health: By 2030, reduce by one third premature mortality from non-communicable diseases through prevention and treatment and promote mental health and well-being [Internet]. [cited 2025 Nov 17]. Available from: <https://www.who.int/data/gho/data/themes/topics/indicator-groups/indicator-group-details/GHO/sdg-target-3.4-noncommunicable-diseases-and-mental-health>
32. Download RStudio [Internet]. Posit. [cited 2025 Nov 17]. Available from: <https://posit.co/downloads/>
33. The MathWorks Inc. MATLAB R2025b [Internet]. Natick, Massachusetts; 2025. Available from: <https://www.mathworks.com>
34. Bleichrodt A, Phan A, Luo R, Kirpich A, Chowell G. StatModPredict: A user-friendly R-Shiny interface for fitting and forecasting with statistical models. PLOS ONE. 2025 Aug 7;20(8):e0329791.
35. Chowell G, Dahal S, Bleichrodt A, Tariq A, Hyman JM, Luo R. *SubEpiPredict*: A tutorial-based primer and toolbox for fitting and forecasting growth trajectories using the ensemble n -sub-epidemic modeling framework. Infect Dis Model. 2024 June 1;9(2):411–36.
36. Wiley.com [Internet]. [cited 2025 Oct 9]. Time Series Analysis: Forecasting and Control, 5th Edition | Wiley. Available from: <https://www.wiley.com/en-us/Time+Series+Analysis%3A+Forecasting+and+Control%2C+5th+Edition-p-9781118675021>
37. Wood SN. Generalized Additive Models: An Introduction with R, Second Edition. 2nd ed. New York: Chapman and Hall/CRC; 2017. 496 p.
38. Simpson GL. Modelling Palaeoecological Time Series Using Generalised Additive Models. Front Ecol Evol [Internet]. 2018 Oct 26 [cited 2025 Oct 9];6. Available from: <https://www.frontiersin.org/journals/ecology-and-evolution/articles/10.3389/fevo.2018.00149/full>
39. Foundations of Linear and Generalized Linear Models | Wiley [Internet]. [cited 2025 Oct 9]. Available from: <https://www.wiley.com/en-us/Foundations+of+Linear+and+Generalized+Linear+Models-p-9781118730034>
40. Taylor SJ, and Letham B. Forecasting at Scale. Am Stat. 2018 Jan 2;72(1):37–45.
41. Facebook Open Source. Prophet: Forecasting at Scale. [cited 2025 Oct 9]. Prophet. Available from: <http://facebook.github.io/prophet/>
42. Chowell G, Luo R, Sun K, Roosa K, Tariq A, Viboud C. Real-time forecasting of epidemic trajectories using computational dynamic ensembles. Epidemics. 2020 Mar 1;30:100379.

43. Mathematical and Statistical Modeling for Emerging and Re-emerging Infectious Diseases | SpringerLink [Internet]. [cited 2025 Oct 9]. Available from: <https://link.springer.com/book/10.1007/978-3-319-40413-4>
44. An ensemble n-sub-epidemic modeling framework for short-term forecasting epidemic trajectories: Application to the COVID-19 pandemic in the USA | PLOS Computational Biology [Internet]. [cited 2025 Oct 9]. Available from: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1010602>
45. Bleichrodt A, Luo R, Kirpich A, Chowell G. Retrospective evaluation of short-term forecast performance of ensemble sub-epidemic frameworks and other time-series models: The 2022-2023 mpox outbreak across multiple geographical scales, July 14th, 2022, through February 26th, 2023. *MedRxiv Prepr Serv Health Sci*. 2023 Oct 17;2023.05.15.23289989.
46. Rtveladze K, Marsh T, Barquera S, Romero LMS, Levy D, Melendez G, et al. Obesity prevalence in Mexico: impact on health and economic burden. *Public Health Nutr*. 2014 Jan;17(1):233–9.
47. Gregg EW, Cheng YJ, Srinivasan M, Lin J, Geiss LS, Albright AL, et al. Trends in cause-specific mortality among adults with and without diagnosed diabetes in the USA: an epidemiological analysis of linked national survey and vital statistics data. *The Lancet*. 2018 June 16;391(10138):2430–40.
48. Geiss LS, Wang J, Cheng YJ, Thompson TJ, Barker L, Li Y, et al. Prevalence and Incidence Trends for Diagnosed Diabetes Among Adults Aged 20 to 79 Years, United States, 1980-2012. *JAMA*. 2014 Sept 24;312(12):1218–26.
49. Dabelea D, Mayer-Davis EJ, Saydah S, Imperatore G, Linder B, Divers J, et al. Prevalence of Type 1 and Type 2 Diabetes Among Children and Adolescents From 2001 to 2009. *JAMA*. 2014 May 7;311(17):1778–86.
50. Fisher-Hoch SP. Missed Opportunities for Diagnosis and Treatment of Diabetes, Hypertension, and Hypercholesterolemia in a Mexican American Population, Cameron County Hispanic Cohort, 2003–2008. *Prev Chronic Dis* [Internet]. 2012 [cited 2025 Oct 9];9. Available from: https://www.cdc.gov/pcd/issues/2012/11_0298.htm
51. Caballero AE. Understanding the Hispanic/Latino Patient. *Am J Med*. 2011 Oct 1;124(10):S10–5.
52. Meza R, Barrientos-Gutierrez T, Rojas-Martinez R, Reynoso-Noverón N, Palacio-Mejia LS, Lazcano-Ponce E, et al. Burden of Type 2 Diabetes in Mexico: Past, Current and Future Prevalence and Incidence Rates. *Prev Med*. 2015 Dec;81:445–50.

53. Boyle JP, Thompson TJ, Gregg EW, Barker LE, Williamson DF. Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. *Popul Health Metr*. 2010 Oct 22;8:29.
54. Boutayeb A, Twizell EH, Achouayb K, Chetouani A. A mathematical model for the burden of diabetes and its complications. *Biomed Eng Online*. 2004 June 28;3(1):20.
55. Stoddard P, He G, Vijayaraghavan M, Schillinger D. Disparities in undiagnosed diabetes among United States-Mexico border populations. *Rev Panam Salud Publica Pan Am J Public Health*. 2010 Sept;28(3):198–206.
56. Culturally Competent Diabetes Self-Management Education for Mexican Americans | Diabetes Care | American Diabetes Association [Internet]. [cited 2025 Oct 9]. Available from: <https://diabetesjournals.org/care/article/25/2/259/23360/Culturally-Competent-Diabetes-Self-Management>
57. American Diabetes Association. Economic Costs of Diabetes in the U.S. in 2017. *Diabetes Care*. 2018 Mar 22;41(5):917–28.
58. Prezio EA, Cheng D, Balasubramanian BA, Shuval K, Kendzor DE, Culica D. Community Diabetes Education (CoDE) for uninsured Mexican Americans: A randomized controlled trial of a culturally tailored diabetes education and management program led by a community health worker. *Diabetes Res Clin Pract*. 2013 Apr 1;100(1):19–28.
59. Shaibi GQ, Konopken Y, Hoppin E, Keller CS, Ortega R, Castro FG. Effects of a Culturally Grounded Community-Based Diabetes Prevention Program for Obese Latino Adolescents. *Diabetes Educ*. 2012 July 1;38(4):504–12.
60. Zhou B, Lu Y, Hajifathalian K, Bentham J, Cesare MD, Danaei G, et al. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants. *The Lancet*. 2016 Apr 9;387(10027):1513–30.
61. Byrd TL, Law JG. Cross-border utilization of health care services by United States residents living near the Mexican border. *Rev Panam Salud Pública*. 2009 Aug;26(2):95–100.
62. Carr MJ, Wright AK, Leelarathna L, Thabit H, Milne N, Kanumilli N, et al. Impact of COVID-19 on diagnoses, monitoring, and mortality in people with type 2 diabetes in the UK. *Lancet Diabetes Endocrinol*. 2021 July;9(7):413–5.

Figures and Tables

Figure 1. Weighted Interval Score (WIS) performance across forecasting models for diabetes burden indicators.

Panel A shows deaths, Panel B shows prevalence, and Panel C shows disability-adjusted life years (DALYs). Values are displayed on a log10 scale, with lower values indicating better prediction interval accuracy. Heatmaps compare performance across geographic regions (MX, US, USMX), demographic categories (sex and age groups), and eight forecasting models including classical time series models and n-sub-epidemic (NSE) approaches.

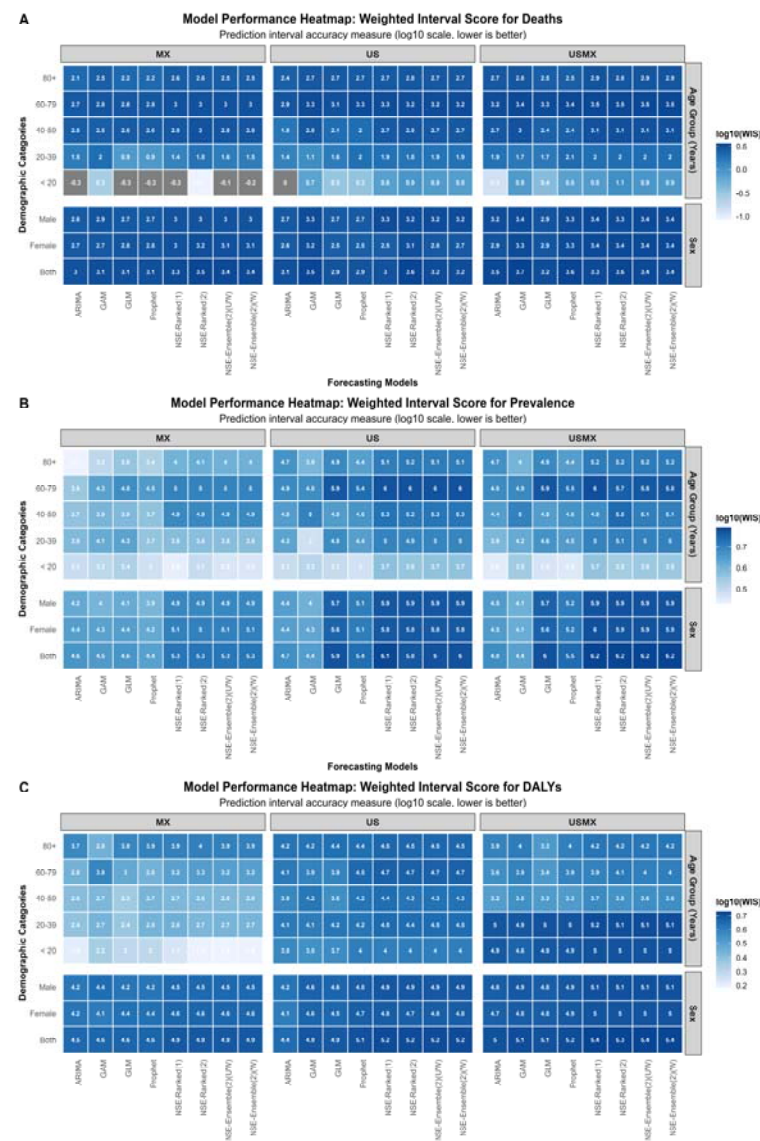


Figure 2: Weighted Interval Score (WIS) calibration across forecasting models for diabetes burden indicators.

Panel A shows deaths, Panel B shows prevalence, and Panel C shows disability-adjusted life years (DALYs). Values are displayed on a log10 scale, with lower values indicating better prediction interval accuracy. Heatmaps compare performance across geographic regions (MX, US, USMX), demographic categories (sex and age groups), and seven forecasting models including classical time series models and n-sub-epidemic (NSE) approaches.

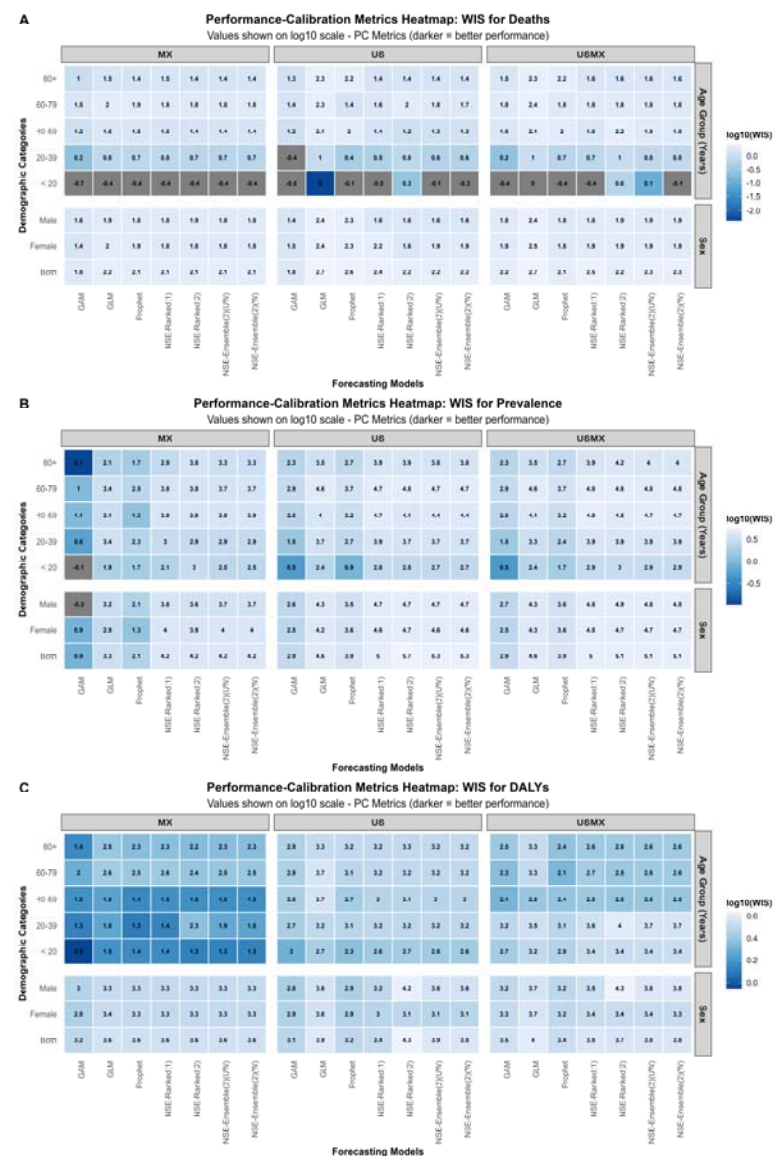


Figure 3: Forecasted percent change in diabetes burden indicators, 2021-2030.

Panel A shows deaths, Panel B shows prevalence, and Panel C shows disability-adjusted life years (DALYs). Values represent percent change from 2021 baseline, with the color scale ranging from blue (decreases) through white (no change) to red (increases). Heatmaps compare model-based forecasts across geographic regions (MX, US, USMX), sex and age groups, and eight forecasting models including classical time series methods (ARIMA, GAM, GLM, Prophet) and neural structure ensemble (NSE) approaches.

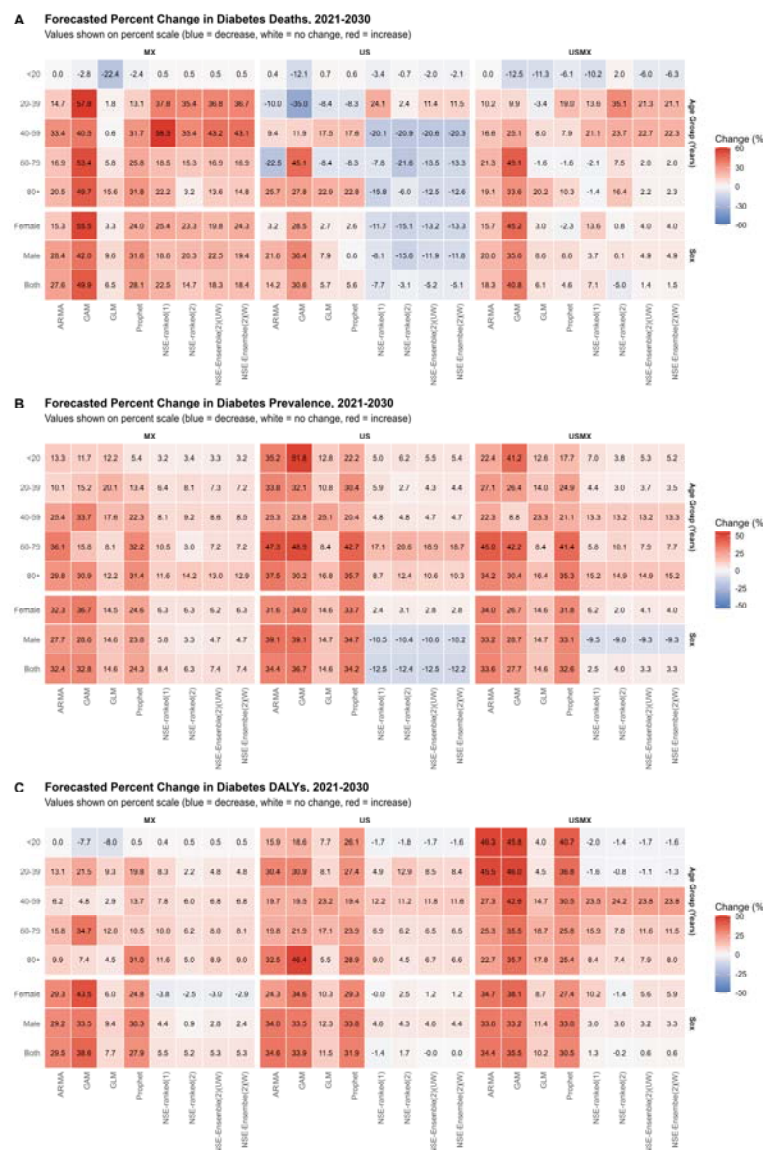


Figure 4: Forecasted Diabetes Burden in the USA-Mexico Border Region, 2030: Comparing WIS-Weighted and Unweighted Model Average.

Point estimates and 95% prediction intervals for diabetes deaths, prevalence, and disability-adjusted life years (DALYs) in 2030. WIS-weighted averages (blue circles) assign greater weight to models with better performance compared to unweighted (pink triangles), which treats all models equally. Each outcome uses independent scales.

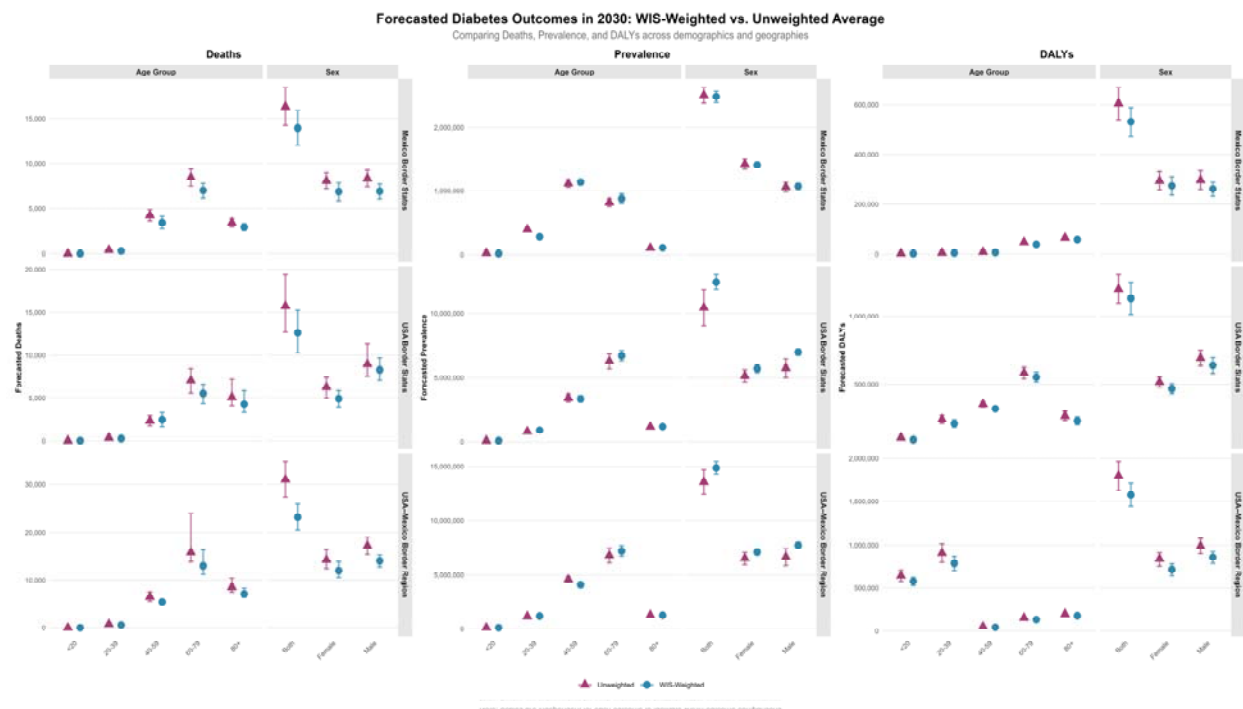


Table 1. Projected Deaths in the USA-Mexico Border Region by 2030 Stratified by Sex and Age Group.

Median projected deaths (in thousands) with 95% prediction intervals across eight forecasting models. Results are presented for three geographic regions (USA-Mexico Border Region, USA Border States, Mexico Border States) stratified by sex and five age groups (<20, 20-39, 40-59, 60-79, 80+ years).

Deaths (2030) — Median and 95% PI								
Category	ARIMA	GAM	GLM	Prophet	NSE-ranked(1)	NSE-ranked(2)	NSE-Ensemble(2)(UW)	NSE-Ensemble(2)(W)
Sex								
USA-Mexico Border Region								
Both	33,523 (28,562–38,484)	39,895 (32,365–47,425)	30,061 (28,780–31,342)	29,639 (27,172–31,771)	30,348 (27,927–33,556)	26,912 (24,263–29,630)	28,733 (24,748–33,044)	28,755 (24,829–33,013)
Female	14,924 (12,266–17,583)	18,723 (14,708–22,738)	13,281 (12,556–14,005)	12,599 (11,418–13,906)	14,652 (12,309–16,859)	13,001 (12,185–13,850)	13,416 (12,214–16,489)	13,416 (12,221–16,502)
Male	18,539 (16,167–20,911)	20,947 (17,244–24,650)	16,780 (16,189–17,371)	16,778 (15,657–17,887)	16,010 (14,517–17,595)	16,394 (14,698–17,865)	16,205 (14,589–17,782)	16,202 (14,580–17,759)
USA Border States								
Both	17,229 (11,620–22,837)	19,708 (16,399–23,017)	15,946 (14,467–17,426)	15,940 (13,072–19,000)	13,934 (11,220–19,136)	14,617 (11,936–17,031)	14,302 (11,508–18,318)	14,327 (11,468–18,346)
Female	6,579 (5,204–7,953)	8,190 (6,946–9,433)	6,544 (5,707–7,381)	6,538 (4,849–8,091)	5,629 (4,658–6,634)	5,410 (4,113–6,581)	5,535 (4,243–6,597)	5,527 (4,295–6,584)
Male	10,602 (8,381–12,843)	11,891 (9,116–14,666)	9,402 (8,746–10,058)	8,771 (7,781–9,779)	8,014 (6,898–11,486)	7,361 (6,345–10,262)	7,681 (6,498–10,696)	7,688 (6,492–10,526)
Mexico Border States								
Both	16,903 (13,764–20,043)	19,864 (14,259–25,469)	14,115 (13,430–14,799)	16,975 (16,274–17,656)	16,228 (14,457–18,036)	15,192 (13,792–16,479)	15,678 (13,954–17,612)	15,683 (13,958–17,603)
Female	7,517 (6,484–8,550)	10,139 (7,107–13,171)	6,737 (6,392–7,082)	8,086 (7,729–8,467)	8,181 (7,645–8,710)	8,042 (7,505–8,586)	7,812 (7,201–8,489)	8,107 (7,559–8,669)
Male	8,637 (7,109–10,165)	9,557 (6,776–12,338)	7,378 (7,029–7,726)	8,852 (8,513–9,174)	7,979 (7,476–8,455)	8,095 (7,526–8,670)	8,243 (7,554–8,932)	8,033 (7,490–8,584)
Age groups (Years)								
USA-Mexico Border Region								
<20	45 (37–53)	39 (26–52)	40 (38–42)	42 (39–45)	40 (36–45)	46 (38–54)	42 (37–53)	42 (37–53)
20–39	764 (644–883)	762 (525–999)	669 (636–703)	824 (781–868)	787 (694–903)	936 (786–1,165)	841 (709–1,123)	839 (707–1,109)
40–59	6,446 (5,450–7,443)	6,914 (4,972–8,857)	5,969 (5,739–6,199)	5,966 (5,539–6,396)	6,693 (5,789–8,211)	6,837 (5,765–7,672)	6,787 (5,775–7,924)	6,763 (5,783–7,888)
60–79	17,571 (14,328–20,814)	21,599 (17,879–25,319)	14,262 (13,460–15,064)	14,262 (12,672–15,715)	14,185 (12,792–16,229)	15,575 (13,588–17,888)	14,775 (12,949–24,904)	14,772 (12,962–25,585)
80+	9,033 (7,737–10,330)	10,137 (8,322–11,952)	9,121 (8,667–9,576)	8,367 (7,875–8,871)	7,484 (6,854–8,290)	8,830 (7,106–11,944)	7,759 (6,928–11,418)	7,763 (6,935–11,096)
USA Border States								
<20	22 (18–26)	19 (11–28)	22 (20–25)	22 (17–27)	21 (18–24)	22 (19–25)	22 (18–25)	22 (18–25)
20–39	324 (279–368)	234 (126–342)	330 (307–353)	330 (285–374)	447 (380–555)	368 (316–413)	401 (327–521)	401 (329–525)
40–59	2,645 (1,172–4,117)	2,704 (1,665–3,744)	2,839 (2,553–3,125)	2,843 (2,253–3,372)	1,932 (1,549–2,309)	1,912 (1,591–2,237)	1,919 (1,567–2,281)	1,927 (1,590–2,267)
60–79	5,829 (4,334–7,324)	10,914 (9,030–12,798)	6,888 (6,166–7,611)	6,894 (5,438–8,262)	6,934 (5,880–8,028)	5,895 (4,401–7,405)	6,505 (4,659–7,882)	6,521 (4,660–7,892)
80+	6,002 (3,610–8,395)	6,102 (5,051–7,153)	5,867 (5,252–6,482)	5,863 (4,651–7,110)	4,020 (3,611–4,444)	4,486 (3,459–9,156)	4,178 (3,562–7,518)	4,171 (3,547–7,429)
Mexico Border States								
<20	23 (16–30)	22 (14–30)	18 (16–20)	22 (20–25)	23 (21–25)	23 (21–25)	23 (21–25)	23 (21–25)
20–39	382 (319–446)	526 (305–746)	339 (323–355)	377 (356–396)	459 (414–497)	451 (398–487)	455 (403–494)	455 (405–493)
40–59	4,151 (3,184–5,117)	4,373 (2,773–5,973)	3,130 (2,913–3,346)	4,098 (3,876–4,309)	4,927 (4,393–5,276)	4,212 (3,912–4,477)	4,458 (3,947–5,241)	4,453 (3,953–5,235)
60–79	8,142 (7,193–9,092)	10,686 (7,794–13,579)	7,373 (7,075–7,671)	8,762 (8,436–9,094)	8,257 (7,467–8,991)	8,033 (7,279–8,745)	8,147 (7,343–8,891)	8,148 (7,346–8,928)
80+	3,391 (2,842–3,941)	4,213 (3,014–5,412)	3,255 (3,066–3,443)	3,710 (3,533–3,894)	3,438 (3,199–3,685)	2,904 (2,677–3,200)	3,197 (2,715–3,648)	3,229 (2,711–3,647)
Median and 95% prediction interval shown as Median (LB–UB).								

Table 2. Projected Disease Prevalence in the USA-Mexico Border Region by 2030 Stratified by Sex and Age Group.

Median projected prevalence (in thousands) with 95% prediction intervals across eight forecasting models. Results are presented for three geographic regions (USA-Mexico Border Region, USA Border States, Mexico Border States) stratified by sex and five age groups (<20, 20-39, 40-59, 60-79, 80+ years).

Prevalence (2030) — Median and 95% PI								
Category	ARIMA	GAM	GLM	Prophet	NSE-ranked(1)	NSE-ranked(2)	NSE-Ensemble(2)(UW)	NSE-Ensemble(2)(W)
Sex								
USA-Mexico Border Region								
Both	15,717 (14,399–17,036)	15,027 (14,753–15,301)	13,487 (13,014–13,959)	15,598 (15,057–16,130)	12,056 (10,502–13,560)	12,241 (10,565–13,969)	12,151 (10,564–13,797)	12,157 (10,548–13,793)
Female	7,577 (6,892–8,263)	7,169 (7,036–7,302)	6,480 (6,269–6,691)	7,456 (7,216–7,690)	6,005 (5,181–6,802)	5,767 (4,959–6,575)	5,887 (5,030–6,731)	5,884 (5,029–6,725)
Male	8,134 (7,485–8,783)	7,860 (7,717–8,003)	7,007 (6,745–7,269)	8,131 (7,852–8,435)	5,527 (4,317–6,708)	5,557 (4,268–6,864)	5,539 (4,297–6,752)	5,543 (4,284–6,800)
USA Border States								
Both	12,927 (11,760–14,095)	13,143 (12,941–13,344)	11,025 (10,585–11,466)	12,908 (12,301–13,534)	8,417 (6,241–10,516)	8,420 (6,212–10,725)	8,418 (6,233–10,602)	8,440 (6,133–10,589)
Female	5,841 (5,272–6,410)	5,945 (5,843–6,047)	5,084 (4,889–5,279)	5,934 (5,672–6,181)	4,543 (3,844–5,190)	4,577 (3,918–5,249)	4,562 (3,882–5,223)	4,563 (3,889–5,222)
Male	7,206 (6,566–7,846)	7,202 (7,140–7,265)	5,941 (5,696–6,187)	6,975 (6,677–7,273)	4,633 (3,481–5,787)	4,638 (3,457–5,780)	4,631 (3,467–5,782)	4,650 (3,492–5,767)
Mexico Border States								
Both	2,844 (2,622–3,067)	2,854 (2,767–2,940)	2,461 (2,426–2,497)	2,670 (2,615–2,718)	2,329 (2,181–2,475)	2,284 (2,139–2,428)	2,306 (2,153–2,461)	2,307 (2,153–2,456)
Female	1,613 (1,501–1,724)	1,666 (1,617–1,715)	1,396 (1,378–1,414)	1,518 (1,496–1,541)	1,295 (1,188–1,398)	1,296 (1,190–1,407)	1,295 (1,188–1,404)	1,296 (1,188–1,404)
Male	1,187 (1,071–1,302)	1,195 (1,149–1,241)	1,065 (1,047–1,084)	1,150 (1,112–1,188)	983 (883–1,080)	960 (869–1,055)	973 (875–1,071)	973 (875–1,074)
Age groups (Years)								
USA-Mexico Border Region								
<20	126 (119–134)	146 (144–147)	116 (116–117)	121 (116–126)	110 (102–117)	107 (100–114)	109 (101–116)	109 (101–116)
20–39	1,320 (1,235–1,405)	1,312 (1,300–1,325)	1,184 (1,168–1,200)	1,297 (1,250–1,342)	1,084 (985–1,177)	1,069 (969–1,168)	1,077 (977–1,174)	1,075 (973–1,172)
40–59	4,827 (4,373–5,282)	4,295 (4,136–4,454)	4,867 (4,797–4,936)	4,779 (4,511–5,055)	4,471 (4,174–4,745)	4,467 (4,098–4,849)	4,467 (4,128–4,797)	4,470 (4,141–4,805)
60–79	8,156 (7,374–8,938)	7,996 (7,761–8,230)	6,096 (5,722–6,470)	7,954 (7,617–8,324)	5,952 (5,036–6,837)	6,190 (5,228–7,097)	6,069 (5,135–6,973)	6,059 (5,090–6,997)
80+	1,412 (1,268–1,557)	1,372 (1,320–1,424)	1,224 (1,183–1,265)	1,424 (1,362–1,499)	1,212 (1,071–1,351)	1,209 (1,061–1,353)	1,209 (1,063–1,350)	1,212 (1,064–1,351)
USA Border States								
<20	102 (95–110)	115 (114–116)	85 (84–87)	93 (90–95)	80 (74–85)	81 (74–87)	80 (74–86)	80 (74–86)
20–39	910 (836–984)	898 (889–908)	754 (726–782)	886 (831–939)	720 (638–801)	698 (606–787)	709 (618–796)	710 (618–794)
40–59	3,743 (3,278–4,209)	3,697 (3,619–3,775)	3,739 (3,683–3,794)	3,597 (3,316–3,881)	3,131 (2,749–3,507)	3,132 (2,723–3,536)	3,129 (2,722–3,510)	3,128 (2,731–3,521)
60–79	7,234 (6,481–7,987)	7,312 (7,186–7,437)	5,325 (4,979–5,671)	7,011 (6,663–7,374)	5,754 (4,944–6,530)	5,923 (5,081–6,678)	5,840 (4,986–6,619)	5,829 (5,003–6,626)
80+	1,321 (1,210–1,433)	1,252 (1,199–1,305)	1,123 (1,085–1,160)	1,304 (1,244–1,376)	1,045 (926–1,169)	1,081 (944–1,203)	1,064 (932–1,192)	1,060 (935–1,191)
Mexico Border States								
<20	31 (27–35)	31 (30–31)	31 (30–32)	29 (25–32)	28 (27–29)	28 (27–29)	28 (27–29)	28 (27–29)
20–39	395 (337–453)	413 (408–418)	430 (415–445)	406 (377–440)	382 (364–398)	388 (370–405)	385 (367–403)	384 (367–402)
40–59	1,202 (1,119–1,285)	1,282 (1,212–1,351)	1,128 (1,110–1,146)	1,172 (1,154–1,191)	1,036 (970–1,102)	1,047 (963–1,128)	1,041 (966–1,119)	1,041 (968–1,119)
60–79	970 (873–1,067)	826 (808–843)	771 (741–800)	942 (897–991)	788 (727–849)	735 (660–811)	764 (671–841)	764 (671–841)
80+	118 (106–129)	119 (116–121)	102 (98–106)	119 (113–125)	101 (92–111)	104 (94–112)	102 (93–111)	102 (92–111)

All values shown in thousands. Median and 95% prediction interval shown as Median (LB–UB).

Table 3. Projected Disability-Adjusted Life Years (DALYs) in the USA-Mexico Border Region by 2030 Stratified by Sex and Age Group.

Median projected DALYs (in thousands) with 95% prediction intervals across eight forecasting models. Results are presented for three geographic regions (USA-Mexico Border Region, USA Border States, Mexico Border States) stratified by sex and five age groups (<20, 20-39, 40-59, 60-79, 80+ years).

DALYs (2030) — Median and 95% PI								
Category	ARIMA	GAM	GLM	Prophet	NSE-ranked(1)	NSE-ranked(2)	NSE-Ensemble(2)(UW)	NSE-Ensemble(2)(W)
Sex								
USA-Mexico Border Region								
Both	2,119,269 (1,826,429–2,412,109)	2,136,550 (1,957,198–2,315,902)	1,738,169 (1,692,171–1,784,168)	2,058,886 (1,960,421–2,161,622)	1,597,982 (1,416,991–1,774,781)	1,573,505 (1,389,159–1,750,837)	1,586,041 (1,398,078–1,763,146)	1,586,478 (1,403,238–1,763,811)
Female	972,204 (822,680–1,121,729)	996,146 (897,319–1,094,973)	784,591 (764,088–805,094)	919,371 (886,113–958,372)	795,001 (743,978–846,623)	711,575 (632,360–790,107)	761,958 (644,753–837,496)	764,198 (648,430–839,033)
Male	1,143,022 (997,159–1,288,884)	1,140,013 (1,052,593–1,227,434)	953,578 (927,269–979,888)	1,138,598 (1,081,980–1,192,700)	880,956 (783,341–973,424)	886,741 (792,848–979,601)	883,400 (788,083–977,222)	883,936 (786,664–977,816)
USA Border States								
Both	1,417,877 (1,197,259–1,638,495)	1,410,368 (1,321,605–1,499,131)	1,174,398 (1,145,558–1,203,239)	1,389,399 (1,300,785–1,478,194)	1,038,414 (947,659–1,129,713)	1,071,242 (964,686–1,176,374)	1,053,283 (953,931–1,162,187)	1,054,002 (955,699–1,165,582)
Female	568,985 (512,869–625,100)	616,338 (587,243–645,433)	505,017 (492,089–517,944)	591,951 (560,496–625,823)	457,685 (414,537–502,059)	469,326 (419,577–514,467)	463,230 (416,281–509,932)	463,522 (416,483–510,052)
Male	798,322 (692,652–903,992)	795,482 (736,335–854,630)	669,382 (652,604–686,160)	797,369 (753,326–845,601)	623,518 (563,570–682,531)	621,467 (560,489–680,476)	623,179 (562,621–680,871)	622,167 (561,229–682,173)
Mexico Border States								
Both	677,837 (582,489–773,184)	725,936 (579,587–872,285)	563,771 (542,198–585,343)	669,809 (647,965–691,291)	552,360 (505,163–601,411)	550,621 (477,130–616,988)	551,354 (488,767–609,179)	551,393 (488,730–608,207)
Female	340,984 (288,806–393,163)	378,406 (299,100–457,712)	279,574 (269,061–290,088)	329,040 (317,663–339,438)	253,733 (216,196–292,667)	257,208 (234,029–289,891)	255,696 (219,221–291,405)	256,021 (219,211–291,744)
Male	335,625 (288,951–382,300)	346,954 (274,550–419,357)	284,196 (272,873–295,519)	338,660 (327,781–350,058)	271,376 (232,183–308,522)	262,211 (223,467–303,487)	266,972 (226,086–306,309)	266,092 (224,969–306,425)
Age groups (Years)								
USA-Mexico Border Region								
<20	810,056 (697,247–922,865)	807,529 (769,930–845,129)	576,192 (547,618–604,767)	779,080 (741,444–818,098)	542,869 (458,421–622,405)	545,975 (471,321–621,411)	544,365 (462,837–621,411)	544,821 (464,944–622,991)
20–39	1,138,605 (929,566–1,347,804)	1,142,407 (1,064,249–1,220,565)	817,472 (779,424–855,519)	1,070,590 (1,028,371–1,115,747)	769,849 (651,910–883,971)	775,986 (641,634–908,579)	774,196 (645,910–897,748)	772,401 (646,444–892,765)
40–59	50,877 (40,741–61,012)	56,972 (44,134–69,810)	45,841 (43,478–48,204)	52,141 (49,784–54,460)	49,345 (46,189–52,012)	49,621 (45,112–52,927)	49,457 (45,565–52,553)	49,461 (45,731–52,516)
60–79	162,418 (144,293–180,542)	175,552 (162,703–188,401)	153,813 (149,007–158,619)	162,958 (148,043–177,222)	150,138 (137,520–161,005)	139,658 (130,092–148,425)	144,547 (131,542–158,242)	144,460 (131,496–158,525)
80+	208,022 (191,329–224,714)	230,128 (211,180–249,075)	199,654 (196,121–203,187)	212,598 (196,402–228,788)	183,723 (171,932–194,856)	182,110 (171,962–192,625)	182,907 (171,943–193,677)	183,007 (172,176–194,025)
USA Border States								
<20	117,785 (107,978–127,593)	120,495 (106,564–134,427)	109,374 (106,846–111,902)	128,156 (124,624–131,609)	99,859 (90,784–109,339)	99,793 (89,990–109,643)	99,837 (90,244–109,460)	99,967 (90,483–109,486)
20–39	272,711 (231,846–313,576)	273,852 (209,317–338,388)	226,026 (216,413–235,639)	266,556 (257,530–275,373)	219,329 (196,266–240,991)	236,186 (208,241–264,414)	226,849 (199,607–260,615)	226,634 (200,038–259,099)
40–59	366,237 (336,698–395,776)	365,659 (323,821–407,497)	376,857 (367,807–385,907)	365,308 (340,881–389,133)	343,152 (327,031–359,005)	340,246 (324,114–356,150)	341,904 (325,340–358,093)	341,529 (325,231–357,456)
60–79	616,788 (578,034–655,541)	627,960 (551,453–704,467)	602,883 (595,341–610,425)	638,122 (595,587–678,370)	550,829 (506,285–591,102)	547,068 (506,731–588,012)	548,640 (506,977–590,109)	548,711 (505,643–590,082)
80+	303,177 (256,163–350,191)	334,798 (270,533–399,064)	241,279 (231,229–251,330)	294,905 (285,060–305,093)	249,365 (215,580–280,414)	239,024 (210,033–269,345)	244,081 (212,419–277,402)	243,852 (212,483–275,803)
Mexico Border States								
<20	3,179 (2,676–3,682)	2,933 (2,034–3,832)	2,924 (2,780–3,068)	3,194 (3,038–3,354)	3,193 (3,034–3,361)	3,196 (3,026–3,358)	3,195 (3,028–3,361)	3,195 (3,030–3,360)
20–39	6,065 (5,475–6,655)	6,514 (5,419–7,608)	5,863 (5,698–6,027)	6,422 (6,036–6,782)	5,810 (5,285–6,325)	5,478 (5,086–5,864)	5,621 (5,128–6,233)	5,618 (5,133–6,237)
40–59	9,074 (8,423–9,724)	8,949 (7,823–10,075)	8,786 (8,631–8,941)	9,711 (8,348–10,058)	8,203 (8,500–9,805)	9,057 (8,510–9,648)	9,121 (8,509–9,745)	9,124 (8,500–9,758)
60–79	49,328 (45,568–53,089)	57,375 (44,660–70,091)	47,701 (46,918–48,485)	47,057 (45,703–48,320)	46,835 (44,073–49,005)	45,206 (42,916–47,388)	45,977 (43,161–48,674)	46,011 (43,154–48,799)
80+	64,866 (45,581–84,151)	63,353 (56,857–69,848)	61,673 (59,226–64,119)	77,313 (74,907–79,876)	65,882 (61,447–70,258)	61,950 (55,691–68,293)	64,253 (56,602–69,844)	64,328 (56,973–69,843)

Median and 95% prediction interval shown as Median (LB–UB).