Supplemental Online Content

Lundeen EA, Burke-Conte Z, Rein DB, et al. Prevalence of diabetic retinopathy in the US in 2021. Published online June 15, 2023. *JAMA Ophthalmol*. doi:10.1001/jamaophthalmol.2023.2289

eAppendix 1. Description of Source Data Processing

eAppendix 2. Description of Statistical Models

eAppendix 3. Imputation of Missing Data

eAppendix 4. ICD-10 Diagnosis Codes

eAppendix 5. Description of Validation and Verification

eTable. Diabetic Retinopathy (DR) and Vision-Threatening Diabetic Retinopathy (VTDR) Prevalence Rates (%) Among Those with Diabetes by US State in 2021

eFigure 1. Age-, Sex/Gender-, and Race and Ethnicity-Standardized Diabetic Retinopathy Prevalence Among Those with Diabetes by US State in 2021, with 95% Uncertainty Intervals

eFigure 2. Age-, Sex/Gender-, and Race and Ethnicity-Standardized Vision-Threatening Diabetic Retinopathy Prevalence Among Those with Diabetes by US State in 2021, with 95% Uncertainty Intervals

eReferences

This supplemental material has been provided by the authors to give readers additional information about their work.

eAppendix 1. Description of Source Data Processing

National Health and Nutrition Examination Survey (NHANES)

We used the 2017–March 2020 cycle of NHANES in the estimation of diabetes prevalence, and the 2005–2008 cycles of NHANES in the estimation of diabetic retinopathy (DR) and vision-threatening diabetic retinopathy (VTDR) prevalence among those with diabetes.

We defined diabetes, in all NHANES cycles, as a participant having a hemoglobin A1c ≥6.5%, taking insulin, or reporting ever having been told by a doctor or health professional that they have diabetes. Among 2005–2008 NHANES participants with diabetes 40 years or older, we used data from reading center-graded retinal fundus photos to determine DR and VTDR status. We defined DR as any retinopathy in the presence of diabetes, including nonproliferative retinopathy (mild, moderate, or severe), proliferative retinopathy, or macular edema. We defined VTDR as having, in the presence of diabetes, severe nonproliferative retinopathy, proliferative retinopathy, panretinal photocoagulation scars, or macular edema. For individuals with insufficient data, we mapped their DR and VTDR status to a missing value; 20.1% and 20.4% of 2005–2008 participants with diabetes 40 years and older had missing values for DR and VTDR status, respectively. We imputed missing categorical indicators of DR and VTDR (eAppendix 3).

NHANES reported these self-reported race and ethnicity categories: non-Hispanic White, non-Hispanic Black, Mexican American, other Hispanic, and all other non-Hispanic races and ethnicities. Our final composite indicator of race and ethnicity combined the

NHANES categories of “Mexican American” and “Other Hispanic,” due to no other data source reporting specific Hispanic origin. Therefore, we estimate prevalence in each of four categories: non-Hispanic White, non-Hispanic Black, Hispanic, and all other non-Hispanic races and ethnicities. Hereafter, we refer to this indicator as “race” and to these four categories as “White,” “Black,” “Hispanic,” and “Other races.”

NHANES data report a variable named “gender,” but this reflects a survey question, “Is {NAME} male or female?,” which only allowed binary responses, was only asked by the interviewer if they hadn’t already assumed the gender of the respondent, and could be interpreted as asking about biological sex. As such, this variable is best understood as a non-differentiated sex/gender measure, a proxy measure for both sex and gender that does not directly measure either.[1] Hereafter, we refer to this variable as “sex/gender” to reflect this limitation. Due to the binary categories available in NHANES, we only estimate prevalence in “male” and “female” sex/gender categories. We are unable to make estimates for gender minority populations, who are known to have generally worse health outcomes and be more likely to lack healthcare access.[2,3]

The NHANES sample design changed in 2007-2008 to include oversampling of all Hispanic groups rather than oversampling of Mexican Americans only. Our use of exam weights adjusts for the oversampling in each cycle, but due to this change, sample size for the Hispanic category is smaller in the 2005–2006 cycle, which leads to increased uncertainty.

NHANES data were available at the individual participant level, allowing calculation of prevalence cross-stratified by age, sex/gender, and race. However, cross-stratified sample sizes were small, an issue addressed by our use of statistical models (eAppendix 2).

National-level Medicare and MarketScan claims

Claims data were mapped to diagnoses of diabetes, DR, and VTDR in accordance with the ICD-10 codes listed in eAppendix 4. At the national level, claims data were stratified by single year of age, but were not stratified by race. We used Medicare data for adults above 67 years, since the overwhelming majority of US adults are covered by Medicare after that age. Due to the high coverage of Medicaid among minors and Medicare among seniors, and the bias this could introduce to commercial insurance claims, we only included MarketScan data in age groups between 20 and 65 years, exclusive. Medicare data used algorithmic adjustments for missing responses.

County-level Medicare and Medicaid claims

Claims data were mapped to diagnoses of diabetes, DR, and VTDR in accordance with the ICD-10 codes listed in eAppendix 4. At the county level, claims data were stratified by age, non-differentiated sex/gender, and a detailed race and ethnicity category (either non-Hispanic White, non-Hispanic Black, non-Hispanic Asian, non-Hispanic American Indian or Alaska Native, Hispanic, or all other races and ethnicities), but not jointly stratified by these three factors; each value reported was stratified by only *one* of age, non-differentiated sex/gender, or race and ethnicity. Medicare and Medicaid data used algorithmic adjustments for missing responses.

Population-based study data

When population-based studies (PBS) reported results stratified by race and ethnicity, we mapped the reported categories to the four categories of our composite "race" indicator, which is described in the NHANES section above. All PBS that reported race used participants' self-reported race to inform this variable.

No PBS reported results stratified by both sex *and* gender. PBS that reported results by either sex *or* gender did not always state which was recorded, or whether the values were self-reported. PBS that used self-reported values did not specify what question was asked, making it difficult to assess whether sex and gender were differentiated. All PBS that reported sex/gender only reported binary categories of male and female. Data limitations force us to assume that all PBS sex/gender measures are equivalent to the non-differentiated sex/gender measure in NHANES, which is a limitation of this analysis. Our results may not reflect prevalence in gender minority populations due to this inconsistency, as well as other limitations to sex/gender measures mentioned previously.

We used the most granular data available from each PBS data source. When results from a study were available stratified by more than one of age, race, and sex/gender but were not available fully cross-stratified by those factors, we included each single-factor stratification of the data separately. For example, the SEARCH results reported prevalence by sex/gender and by racial group, but not cross-stratified by both. Therefore, both the prevalence rates by sex/gender (for all racial groups combined) and the prevalence rates by race (for all sex/gender groups combined) were included.

We used each study's reported age range, if available. When a maximum age was not specified for a study, we used a value of 90 years. The studies of DR in youth did not report minimum or maximum ages; instead of a range, we considered the prevalence to be observed at the point estimate of mean age in those studies.

Not all PBS reported VTDR. We calculated estimated VTDR prevalences for these studies by applying a ratio found in national 2019 Medicare claims. For each study, we used the set of reported diabetic eye conditions that was most similar to our VTDR definition, without necessarily being mutually exclusive. For example, the process in the case of the Los Angeles Latino Eye Study (LALES) was as follows:

(1) The set of diabetic eye conditions reported by LALES that was most similar to VTDR was severe non-proliferative DR, proliferative DR, and diabetic macular edema. As noted above, this is not a mutually exclusive set of conditions.

(2) We calculated the ratio between the VTDR prevalence and the sum of the prevalences of those three conditions in the national 2019 Medicare claims.

(3) We found the sum of the prevalences of those three conditions within each stratum reported by LALES.

(4) We multiplied those sums by the calculated ratio to estimate the VTDR value for each LALES stratum.

We used exactly the same procedure for the Chinese American Eye Study (CHES) and Diabetic Retinopathy Inpatient Study (DRIPS). In the case of Proyecto Vision and Eye Research (VER), which did not report severe non-proliferative DR separately, the process

was identical except that we used a different set of conditions: all DR and diabetic macular edema.

A list of all data sources used in the estimation of national DR and VTDR prevalence, including NHANES, claims, and PBS, can be found in eAppendix 2.

eAppendix 2. Description of Statistical Models

National DR and VTDR prevalence

We represent our prevalence models with a stochastic component and a systematic component, where the stochastic component is a negative binomial model of count data:

$$Y_i \sim \text{NegativeBinomial}\left(N_{\text{eff}}^i, \pi_i, \eta\right),$$

where i indexes the specific measurement; Y_i is the prevalence count of those with DR, or in a separate model VTDR, in measurement i; N_{eff}^i is the effective sample size from which the count was taken (and so $p_i = \frac{Y_i}{N_{\text{eff}}^i}$ is the prevalence rate typically reported in a population-based study [PBS]); π_i is the prevalence rate predicted by the model; and η is the over-dispersion parameter of the negative binomial distribution (assumed to be the same for all measurements). We used DisMod-MR 1.1.1,[4] which uses the Python PyMC 2 package[5] to implement this Bayesian computation, and follows the formulation of the negative binomial model provided by PyMC 2, where $Pr(Y = x|\mu, \alpha) = \frac{\Gamma(x+\alpha)}{x!\Gamma(\alpha)}\left(\frac{\alpha}{(\mu+\alpha)}\right)^{\alpha}\left(\frac{\mu}{(\mu+\alpha)}\right)^{x}$, which in terms of the equation above has $x = p_i \cdot N_{\text{eff}}^i$, $\mu = \pi_i$, and $\alpha = \eta$. We fit the models with 100,000 iterations of Markov chain Monte Carlo using an Adaptive Metropolis step method.

In the systematic component of the model, we included fixed effects for non-differentiated sex/gender ($\alpha_{\text{sex/gender}}$), age ($\beta_k$ for $k = 0, \ldots, K$), race (γ_j for j = White, Black, Hispanic, and Other races), and data source (δ_j for j = Proyecto VER, MESA, LALES, SEARCH, TODAY, CHES, DRIPS, TODAY2, MarketScan, and Medicare; we coded NHANES as

the reference category). Our formulation includes a piece-wise linear spline model on age, with spline knots $(\text{knot}_1, \text{knot}_2, \dots)$ indexed by k for $k = 1, \dots, K$ as follows:

$$\pi_i = \exp\{\alpha_{\text{sex/gender}}[\text{sex/gender}_i = \text{female}]\} \times \left(\int_{x=\text{age}_0^i}^{\text{age}_1^i} \left(\beta_0 + \sum_{k=1}^{K} \beta_k \, (x - \text{knot}_k)^+ \right) dx \right) \times \exp\left\{ \sum_{j=1}^{4} \gamma_j \, [\text{race}_i = j] + \sum_{j=1}^{10} \delta_j \, [\text{source}_i = j] \right\},$$

where sex/gender_i is the non-differentiated sex/gender measure of measurement i; age_0^i and age_1^i are the start and end of the age group measured in measurement i; race_i is the race group measured in measurement i; source_i is the data source for measurement i; and notation $[\text{variable} = \text{value}]$ represents an indicator function which takes value 1.0 if the variable is equal to the value and 0.0 otherwise and notation $(\text{value})^+$ represents the value in the parenthesis if it is positive, and takes value zero otherwise. We used $K = 8$ spline knots at ages (0, 10, 20, 40, 60, 70, 80, 100). Our model has an age-standardizing likelihood to account for the heterogeneous reporting of age groups in PBS data.

In contrast to our model for vision loss and blindness,[6] we did not include interaction terms between sex/gender, age, and race effects in this model, due to a lack of sufficient data.

We used a Bayesian framework for inference with weakly informative priors for model parameters to assist in regularization, which primarily allowed the data to inform the model estimates. Prior distributions for $\alpha_{\text{sex/gender}}, \beta_k, \gamma_j$, and δ_j were all set to independent normal distributions with mean 0.0 and standard deviation 1.0. β_0 was set to

0.0, because there is no prevalence of DR or VTDR at birth. All prevalence values were restricted to the range [0,1].

We included all data in the likelihood with the assumption that they applied to a national-level estimate, which we assumed to be constant over time.

Data sources used in the estimation of national DR and VTDR prevalence

Data source	Date of publication	Date of data collection	Age groups	Stratified by sex/gender	Stratified by race	Reported VTDR	Fundus photographs vs OCT	Individual- vs. Summary-level data
Proyecto Vision and Eye Research (VER)[7]	2001	April 1997–September 1999	40–90 years (one age group)	No	Yes, all Hispanic	No, used ratio from Medicare claims[a]	Fundus photographs	Summary-level
Multi-Ethnic Study of Atherosclerosis (MESA)[8]	2006	2002–2004	45–85 years (one age group)	No	Yes	Yes	Fundus photographs	Summary-level
Los Angeles Latino Eye Study (LALES)[9]	2004	February 2000–May 2003	40–90 years, by 10-year age groups	Yes; self-reported with unknown question; called "gender" by study authors	Yes, all Hispanic	No, used ratio from Medicare claims[a]	Fundus photographs	Summary-level
SEARCH for Diabetes in Youth[10,11]	2012, 2023	2009–2019	22.63 (single age group at mean age)	Yes; self-reported; non-differentiated sex/gender	Yes	Yes	Fundus photographs	Summary-level
Treatment Options for Type 2 Diabetes in Adolescents and Youth (TODAY)[12]	2013	2010–2011	13.7 (single age group at mean	No	No	Yes	Fundus photographs	Summary-level

			age)					
Chinese American Eye Study (CHES)[13]	2016	February 2010–October 2013	50–90 years, by 10-year age groups	Yes; self-reported with unknown question; called "gender" by study authors	Yes, all Other	No, used ratio from Medicare claims[a]	Fundus photographs	Summary-level
Diabetic Retinopathy Inpatient Study (DRIPS)[14]	2016	September 2011–August 2012	18–90 years (one age group)	No	No	No, used ratio from Medicare claims[a]	Fundus photographs	Summary-level
Follow-up to TODAY study (TODAY2)[15,16]	2021	2017–2018	25.4 (single age group at mean age)	No	No	Yes	Fundus photographs	Summary-level
IBM® MarketScan® claims[17]	N/A	2016	21–64 years, by single-year age	Yes; as coded on claim; non-differentiated sex/gender	No	Yes	Both	Individual-level
Medicare Part B fee-for-service claims[18]	N/A	2019	68+ years, by single-year age	No	No	Yes	Both	Individual-level
NHANES 2005–2008[19,20]	N/A	2005–2008	40+ years, by 5-year age groups[b]	Yes; partially self-reported; non-differentiated sex/gender	Yes	Yes	Fundus photographs	Individual-level

[a] As described in the "Population-based study data" section of eAppendix 1.

[b] NHANES data provide age in months for each participant. We created 5-year age groups between 40 and the maximum age not top-coded, which was 85 in the 2005–2006 cycle and 80 in the 2007–2008 cycle. We created an additional group in each cycle of participants with top-coded age. We set the age of this group to the mean of top-coded ages in that cycle, which is reported in the NHANES documentation.[19,20]

County-level random effects

To inform county-level random effects on prevalence, we used a robust linear mixed-effects model with 5% trimming[21] fit to the county-level claims rates derived from 2018 Medicaid and 2017–2019 Medicare fee-for-service claims data. We developed this model to estimate the county-to-county variation in prevalence standardized for age, non-differentiated sex/gender, and race in the absence of large-sample cross-stratified data for directly standardizing the rates (such data were available at the state level in our prior work on visual acuity loss and blindness, via the public-use microdata sample of the American Community Survey).

Claims data have small sample size in some age group, sex/gender, or race strata when also stratified by county. Additionally, when the number of people treated in a stratum is 1 or 2, the claims rate is suppressed. To address these problems, we estimated the claims rate in each county stratum by using the binomial observations in the data to update a prior beta distribution, with mean μ equal to the average county's claims rate for that stratum, data source, and year, and effective sample size ν equal to 10 ($\alpha = \mu\nu$, $\beta = (1 - \mu)\nu$). This prevented claims rates observed in small samples from unduly influencing county random effects. In the case of suppressed claims rates due to 1 or 2 people treated, we updated the prior with a truncated beta-binomial distribution, with α and β parameters equal to those of the prior, n equal to the stratum sample size, and successes k truncated to the interval [1,2].

Claims data strata with sample size 1–10 were suppressed entirely, so these data were excluded from our analysis. In cases where there were no unsuppressed data in a county,

our random effect was zero, presuming the county to have the same standardized prevalence as the average county.

Our model took the form of predicting the log of the claims rate ratio for each stratum available in the claims data, where the ratio is equal to the estimated county-level claims rate divided by the national-level claims rate. We included fixed effects for sex/gender ($\alpha_{\text{sex/gender}}$ for sex/gender = female, male; we coded 'all' as the reference category), age group (β_{age} for age = all, <18, 18–39, 40–64, ≥85; we coded 65–84 as the reference category), race and ethnicity (γ_j for j = non-Hispanic White, non-Hispanic Black, non-Hispanic Asian, non-Hispanic American Indian or Alaska Native, Hispanic, and all other races and ethnicities; we coded 'all' as the reference category), data source ($\delta_{Medicaid}$; we coded Medicare as the reference category), year (ζ_j for j = 2018, 2019; we coded 2017 as the reference category), and ophthalmologists per capita in 2018 according to the 2019–2020 Area Health Resources File[22] (η):

$$\begin{aligned}\log\frac{\text{CR}_{i,j}}{\text{national CR}_i} \sim \quad & \beta_0 + \sum_{k=1}^{2} \alpha_k\,[\text{sex/gender}_i = k] + \sum_{k=1}^{5} \beta_k\,[\text{age\_group}_i = k] + \\ & \sum_{k=1}^{6} \gamma_k\,[\text{race\_eth}_i = k] + \delta_{\text{Medicaid}}[\text{data\_source}_i = \text{Medicaid}] + \\ & \sum_{k=1}^{2} \zeta_k\,[\text{year}_i = k] + \\ & \eta \times \text{ophth\_per\_capita}_i + \text{interaction\_terms} + u_j + \epsilon_i,\end{aligned}$$

where i indexes the available county-level claims rate values, stratified by age group, by sex/gender, or by race and ethnicity, j ranges across counties, and u_j is the random effect for which we want to estimate a mean and standard deviation.

We included interaction terms to control for additional confounders that could impact inclusion in each source of claims data:

$$\begin{aligned}\text{interaction\_terms} = \quad & [\text{data\_source}_i = \text{Medicare}] \times \{\eta \times \text{medicare\_advantage\_pct}_i + \\ & [\text{age\_group\_start}_i < 65] \sum_{k=1}^{5} \theta_k \, \text{SSDI\_pct}_i [\text{age\_group}_i = k]\} + \\ & [\text{data\_source}_i = \text{Medicaid}] \times \sum_{k=1}^{51} \iota_k \, [\text{state}_j = k]\end{aligned}$$

where i indexes the available county-level claims rate values, stratified by age group, by sex/gender, or by race and ethnicity, j ranges across counties, η is the coefficient of percentage of Medicare beneficiaries enrolled in Medicare Advantage in June of 2018,[23] age_group_start is the lower bound of the age group of value i (which is 0 for <18 and all ages), θ_k is the coefficient of the state-level percentage of the working-age population 18 to 65 years covered by Medicare through Social Security Disability Insurance (SSDI) interacted with age group k, ι_k is the coefficient of US state (or District of Columbia) k when the data source is Medicaid, and state_j is the US state (or District of Columbia) for county j.

County-level DR and VTDR prevalence

We combined national DR and VTDR prevalence estimates with county-level random effects to generate DR and VTDR prevalence estimates stratified by sex/gender, age, race, and US county. These were the unique set of estimates that satisfy the following two conditions:

(1) for a specified age group a, sex/gender s, and race r, the prevalence rate ratio in our estimates p between counties c_1 and c_2 is equal to the ratio implied by the log-space random effects u for the same counties:

$$p_{a,s,r,c_1}/p_{a,s,r,c_2} = \exp\{u_{c_1} - u_{c_2}\}$$

(2) for a specified age group a, sex/gender s, and race r, the weighted mean of our prevalence estimates p among those with diabetes across all counties, using weights w from our fully stratified 2021 population and the diabetes prevalence estimates d from the county-level diabetes prevalence model described below, is equal to our national DR or VTDR prevalence estimate P:

$$P_{a,s,r} = \frac{\sum_{c=0}^{3,143} w_{a,s,r,c} * d_{a,s,r,c} * p_{a,s,r,c}}{\sum_{c=0}^{3,143} w_{a,s,r,c} * d_{a,s,r,c}}$$

After generating these county-level prevalence rates, we clipped them to be between 0 and 1 inclusive. We also clipped VTDR rates to never be greater than the corresponding DR rate. These steps had negligible effects on our results.

National diabetes prevalence

We represent our diabetes prevalence model with a stochastic component and a systematic component, where the stochastic component is a Bernoulli model of binary data:

$$Y_i \sim \text{Bernoulli}(\pi_i),$$

where i indexes the specific 2017–March 2020 NHANES respondent, Y_i is a binary response representing presence of diabetes for respondent i, and π_i is the prevalence rate predicted by the model.

In the systematic component of the model, we included fixed effects for sex/gender ($\alpha_{\text{sex/gender}}$), linear, quadratic and cubic terms for age (β_{power} for power = 1, 2, 3) following Lin et al.,[24] race (γ_j for j = White, Black, Hispanic, and Other races), and income above/below 185% of poverty level (δ_{income}) in order to account for the NHANES oversampling based on income at or below 185% of the poverty level.[25] Our formulation uses a logit link function, as follows:

$$\begin{aligned}\text{logit}(\pi_i) = \quad & \alpha_{\text{sex/gender}}[\text{sex/gender}_i = \text{female}] + \sum_{j=1}^{3} \beta_j \, \text{age}_i^j \times \sum_{j=1}^{4} \gamma_j \, [\text{race}_i = j] + \\ & \delta_{\text{income}}[\text{income}_i = \text{above threshold}] + \text{interaction\_terms},\end{aligned}$$

where sex/gender_i is the sex/gender of respondent i; age_i is the age of respondent i; race_i is the race of respondent i; income_i is a dichotomous value of income above or below 185% of the poverty level for respondent i; and notation [variable = value] represents an indicator function which takes value 1.0 if the variable is equal to the value and 0.0 otherwise.

We also included the first-order interaction of sex/gender and age following Lin et al., as well as all first-order interaction terms with income in order to fully account for the NHANES income-based oversampling as noted above:

$$\begin{aligned}\text{interaction\_terms} = \quad & \alpha_{\text{sex/gender}}'[\text{sex/gender}_i = \text{female}] \times \text{age}_i + \\ & \alpha_{\text{sex/gender}}''[\text{sex/gender}_i = \text{female}][\text{income}_i = \text{above threshold}] + \\ & \beta_1''\text{age}_i[\text{income}_i = \text{above threshold}] + \\ & \sum_{j=1}^{4} \gamma_j''[\text{race}_i = j][\text{income}_i = \text{above threshold}],\end{aligned}$$

where $\alpha_{\text{sex/gender}}'$ is the coefficient of sex/gender interacted with (linear) age, $\alpha_{\text{sex/gender}}''$ is the coefficient of sex/gender interacted with income, β_1'' is the coefficient of (linear) age interacted with income, and γ_j'' is the coefficient of race j interacted with income (for j = White, Black, Hispanic, and Other races).

We implemented our diabetes prevalence model using the Bambi Python package.[26] Priors on all parameters were set by the default prior algorithm implemented in that package, which generates weakly informative priors based on the scale of the observed data.

After making predictions for all combinations of sex/gender, 5-year age group, race, and income above/below 185% of the poverty level, we aggregated these predictions to the sex/gender, age group, and race level. Specifically, we performed a weighted average within each sex/gender, age group, and race stratum, where the weights for the above-threshold and below-threshold income groups were determined by the weights of those groups in the 2019 American Community Survey (ACS) Public Use Microdata Sample[27] in the corresponding stratum.

County-level diabetes prevalence

We generated diabetes prevalence estimates stratified by sex/gender, age, race, and US county. These drew on county-level total (diagnosed and undiagnosed) diabetes prevalence estimates previously published by Dwyer-Lindgren et al., which were stratified by sex only and reported for the year 2012.[28] We mapped this sex stratification to our composite non-differentiated sex/gender measure; it is unclear given the complex nature of the Dwyer-Lindgren et al. analysis whether the reported sex variable was well-differentiated from gender. We used the same values within each age group and race combination for a given sex/gender, on the assumption that age group and race do not modify the effect of county. We used age-standardized values because values standardized for both age and race were not reported; this disaggregation applies county variation associated with different race distributions twice, which is a limitation of our analysis. We generated the unique set of estimates that satisfy the following two conditions:

(1) for a specified age group a, sex/gender s, and race r, the prevalence rate ratio in our estimates p between counties c_1 and c_2 is equal to the prevalence rate ratio between the same counties in the previously published estimates L for the same sex/gender:

$$p_{a,s,r,c_1}/p_{a,s,r,c_2} = L_{s,c_1}/L_{s,c_2}$$

(2) for a specified age group a, sex/gender s, and race r, the weighted mean of our prevalence estimates p across all counties, using weights w from our fully stratified 2021 population, is equal to the national diabetes prevalence estimate P after the income aggregation in the previous step:

$$P_{a,s,r} = \frac{\sum_{c=0}^{3,143} w_{a,s,r,c} * p_{a,s,r,c}}{\sum_{c=0}^{3,143} w_{a,s,r,c}}$$

After generating these county-level prevalence rates, we clipped them to be between 0 and 1 inclusive. This step had a negligible effect on our results.

eAppendix 3. Imputation of Missing Data

Missing data in meta-analysis can jeopardize inferences from the study's examination data, and possibly bias the estimates. The vision examination data from the 2005–2008 cycles of NHANES is a key input in our meta-analytic estimates of the prevalence of diabetic retinopathy (DR) and vision-threatening diabetic retinopathy (VTDR) in the US; however, when we mapped these data to dichotomous outcomes, we were unable to determine a DR and VTDR value for 20.1% and 20.4%, respectively, of participants with diabetes aged ≥40 years.

Simple approaches, such as complete case analysis (also called list-wise deletion) and available-case analysis, may produce biased results. For both DR and VTDR, the mean age among individuals with diabetes over 40 years with insufficient eye exam data to determine a value was 68.3 years, substantially higher than the mean age among individuals with diabetes over 40 years with sufficient data (63.2 years). Since DR and VTDR prevalence rates increase with age until stabilizing in very old ages, our imputation method must take age into account.

Imputation requires choices about methods and variable selection. While in the past, variable selection has been performed with simple heuristics,[29] in the face of evidence that overfitting is a practical issue in imputation,[30] a more sophisticated method seems desirable. Using information criteria (IC) has been proposed,[31] but this may be difficult to implement across methods and model types.

To choose an imputation approach, we conducted a five-fold cross-validation exercise from machine learning.[32] We withheld data from each fold in turn, allowed our imputation methods to impute it as if it were missing, then compared the imputed values with the true values. We removed not only the dichotomous outcome variable from each row in our validation set, but also some subset of auxiliary variables according to a missingness pattern sampled at random from a truly missing row.

As a benchmark, we compared imputation performance to a complete-case analysis. Our analytical model stratified NHANES data by age, sex/gender, and race/ethnicity before processing, so a complete-case analysis is equivalent to filling missing values with the mean of the matching age, sex/gender, and race/ethnicity strata. We cross-validated this equivalent method to determine whether imputation was beneficial relative to complete-case analysis.

We considered two imputation algorithms: the bootstrap-based expectation maximization algorithm provided by the Amelia II R package[33] and multiple imputation by chained equations (MICE) as implemented by the statsmodels Python package.[34]

MICE constructs a regression model for each column with missing values, using the variables from the other columns as the predictors. If multiple columns have missing values, the MICE procedure iterates through the columns, fitting each with the previously imputed values for the other columns and adding random parameter perturbations in order to explore the space. We tested both bootstrap resampling and Gaussian perturbation methods.

To perform imputation, we ran the method (MICE or Amelia) 100 times and took the average of the imputed values as a probability prediction of the outcome of interest (DR or VTDR). For MICE, we used 10 cycles in each run.

In order to narrow the search space of auxiliary variable combinations, we selected the exam and survey variables most plausibly linked to DR: measured hemoglobin A1c (referred to here as lbxgh), self-reported diagnosis of diabetes or prediabetes (diq010), self-reported taking of insulin (diq050), self-reported diagnosis of DR (diq080), presence of macular edema according to examination (macular_edema), presence of any retinopathy according to examination (any_retinopathy) in the case of DR or presence of severe retinopathy according to examination (severe_retinopathy) in the case of VTDR, and the self-reported number of years the patient has had diagnosed diabetes (years_with_diabetes). Before applying the imputation methods, we "pre-imputed" some values using logical relationships between these variables, such as the fact that DR is defined as either macular edema or any retinopathy.

For each imputation method, we:

(1) Cross-validated imputation using only a continuous age variable and imputation using only a dummy-coded categorical age variable grouped in 5-year bins. We selected the simpler continuous variable when its mean area under the ROC curve (AUC) across the five folds was better than the best mean AUC minus one standard error.

(2) Paired the selected age variable with some or none of the sex/gender and (categorical) race/ethnicity variables, once again running cross-validation on each possibility and

selecting the simplest within one standard error of the best AUC. We defined "simplest" as having the fewest variables included, counting dummy variables.

(3) Paired the selected set of demographic variables with every combination of the exam and survey variables listed above, except that we did not allow macular_edema to be used without any_retinopathy or vice versa in DR, and we did not allow macular_edema to be used without severe_retinopathy or vice versa in VTDR. This is because each pair is mutually exclusive and defines the respective condition. Once again, we selected the simplest variable combination within one standard error of the best AUC, except that in this step there was a tie between equally complex combinations in imputing DR. We resolved this tie with domain knowledge that lbxgh (A1c) is more clearly linked to DR than diq050 (taking insulin); this formula also had the higher AUC.

This process, repeated for DR and VTDR, yielded the imputation methods and variable sets used in our model. For both DR and VTDR, the imputation method selected was Amelia. The variable sets were as follows:

(1) DR ~ sex + race_eth + lbxgh + macular_edema + any_retinopathy + years_with_diabetes

(2) VTDR ~ race_eth + lbxgh + diq080

The tables below contain AUC and standard error (SE) for the best-performing methods and variable combinations tested for imputing DR and VTDR, respectively.

Best Twenty Imputation Methods for Imputing Diabetic Retinopathy, as Compared by Area Under the Curve (AUC)

Variable set	AUC	SE (percentage points)	# of variables	Within one SE
sex + race_eth + lbxgh[a] + diq050[c] + macular_edema[e] + any_retinopathy[f] + years_with_diabetes[g]	80.04%	2.22pp	10	Yes
sex + race_eth + lbxgh[a] + macular_edema[e] + any_retinopathy[f] + years_with_diabetes[g]	79.29%	2.25pp	9	Yes
sex + race_eth + lbxgh[a] + diq050[c] + diq080[d] + macular_edema[e] + any_retinopathy[f] + years_with_diabetes[g]	79.18%	2.22pp	11	Yes
sex + race_eth + lbxgh[a] + diq010[b] + diq050[c] + diq080[d] + macular_edema[e] + any_retinopathy[f] + years_with_diabetes[g]	79.07%	1.80pp	14	Yes
sex + race_eth + lbxgh[a] + diq010[b] + diq050[c] + macular_edema[e] + any_retinopathy[f] + years_with_diabetes[g]	78.98%	1.84pp	13	Yes
sex + race_eth + lbxgh[a] + diq010[b] + diq050[c] + diq080[d] + years_with_diabetes[g]	78.64%	1.48pp	12	Yes
sex + race_eth + lbxgh[a] + diq080[d] + macular_edema[e] + any_retinopathy[f] + years_with_diabetes[g]	78.60%	1.90pp	10	Yes
sex + race_eth + lbxgh[a] + diq010[b] + diq080[d] + macular_edema[e] + any_retinopathy[f] + years_with_diabetes[g]	78.44%	1.70pp	13	Yes
sex + race_eth + diq050[c] + macular_edema[e] + any_retinopathy[f] + years_with_diabetes[g]	78.03%	1.67pp	9	Yes
sex + race_eth + lbxgh[a] + diq010[b] + diq050[c] + years_with_diabetes[g]	77.99%	1.49pp	11	Yes
sex + race_eth + lbxgh[a] + diq050[c] + years_with_diabetes[g]	77.82%	2.01pp	8	No
sex + lbxgh[a] + diq050[c] + years_with_diabetes[g]	77.77%	2.32pp	4	No
sex + race_eth + lbxgh[a] + diq010[b] + macular_edema[e] + any_retinopathy[f] +	77.73%	1.49pp	12	No

years_with_diabetes[g]				
sex + lbxgh[a] + diq010[b] + diq050[c] + diq080[d] + years_with_diabetes[g]	77.69%	1.91pp	8	No
sex + lbxgh[a] + diq050[c] + diq080[d] + years_with_diabetes[g]	77.66%	2.07pp	5	No
sex + lbxgh[a] + diq010[b] + diq050[c] + years_with_diabetes[g]	77.58%	1.86pp	7	No
sex + lbxgh[a] + diq050[c] + years_with_diabetes[g]	77.45%	2.21pp	4	No
sex + race_eth + lbxgh[a] + diq050[c] + diq080[d] + years_with_diabetes[g]	77.41%	2.00pp	9	No
sex + lbxgh[a] + diq050[c] + diq080[d] + years_with_diabetes[g]	77.33%	2.19pp	5	No
sex + lbxgh[a] + diq080[d] + years_with_diabetes[g]	77.29%	1.50pp	4	No

[a] Measured hemoglobin A1c.

[b] Self-reported diagnosis of diabetes or pre-diabetes.

[c] Self-reported taking of insulin.

[d] Self-reported diagnosis of DR.

[e] Presence of macular edema according to examination.

[f] Presence of any retinopathy according to examination.

[g] Self-reported number of years the patient has had diagnosed diabetes.

Best Twenty Imputation Methods for Imputing Vision-Threatening Diabetic Retinopathy, as Compared by Area Under the Curve (AUC)

Variable set	AUC	SE (percentage points)	# of variables	Within one SE
race_eth + lbxgh[a] + diq050[c] + diq080[d] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	82.29%	2.64pp	10	Yes
race_eth + lbxgh[a] + diq010[b] + diq050[c] + diq080[d] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	81.79%	1.94pp	13	Yes
race_eth + lbxgh[a] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	81.39%	2.90pp	8	Yes
race_eth + lbxgh[a] + diq010[b] + diq080[d] + years_with_diabetes[g]	81.30%	2.84pp	10	Yes
race_eth + lbxgh[a] + diq010[b] + diq050[c] + diq080[d] + macular_edema[e] + severe_retinopathy[f]	81.19%	1.93pp	12	Yes
race_eth + lbxgh[a] + diq080[d] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	81.18%	2.39pp	9	Yes
race_eth + lbxgh[a] + diq010[b] + diq080[d] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	80.91%	2.44pp	12	Yes
race_eth + lbxgh[a] + diq010[b] + diq080[d] + macular_edema[e] + severe_retinopathy[f]	80.77%	1.36pp	11	Yes
race_eth + lbxgh[a] + diq080[d] + macular_edema[e] + severe_retinopathy[f]	80.58%	1.40pp	8	Yes
race_eth + diq080[d] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	80.50%	1.36pp	8	Yes
race_eth + lbxgh[a] + diq080[d]	80.40%	1.91pp	6	Yes
race_eth + lbxgh[a] + diq050[c] + diq080[d] + years_with_diabetes[g]	80.29%	3.22pp	8	Yes
race_eth + diq050[c] + macular_edema[e] + severe_retinopathy[f] +	80.21%	2.12pp	8	Yes

years_with_diabetes[g]				
race_eth + lbxgh[a] + diq050[c] + diq080[d] + years_with_diabetes[g]	79.99%	2.80pp	8	Yes
race_eth + lbxgh[a] + diq010[b] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	79.99%	3.26pp	11	Yes
age + lbxgh[a] + diq010[b] + diq050[c] + diq080[d] + years_with_diabetes[g]	79.76%	3.70pp	8	Yes
race_eth + lbxgh[a] + diq050[c] + diq080[d] + macular_edema[e] + severe_retinopathy[f]	79.75%	2.46pp	9	Yes
race_eth + lbxgh[a] + diq050[c] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	79.73%	2.67pp	9	Yes
race_eth + lbxgh[a] + diq080[d] + years_with_diabetes[g]	79.70%	3.52pp	7	Yes
race_eth + lbxgh[a] + diq010[b] + diq050[c] + macular_edema[e] + severe_retinopathy[f] + years_with_diabetes[g]	79.62%	4.30pp	12	No

[a] Measured hemoglobin A1c.

[b] Self-reported diagnosis of diabetes or pre-diabetes.

[c] Self-reported taking of insulin.

[d] Self-reported diagnosis of DR.

[e] Presence of macular edema according to examination.

[f] Presence of severe retinopathy according to examination.

[g] Self-reported number of years the patient has had diagnosed diabetes.

eAppendix 4. ICD-10 Diagnosis Codes

ICD-10 Diagnosis Codes Used to Define Diabetes in Medicare, Medicaid, and MarketScan Claims Data

To define people with diabetes in the Medicare, Medicaid, and MarketScan claims data, we used the Chronic Conditions Data Warehouse diabetes algorithm, which can be found at: https://www2.ccwdata.org/web/guest/condition-categories. People were coded with a diabetes diagnosis if they had ≥1 inpatient or ≥2 different day outpatient diagnosis codes (International Classification of Diseases [ICD], Tenth Revision) on claims during the calendar year or 1-year look-back period.

Diabetes ICD-10 Codes included: E08.00, E08.01, E08.10, E08.11, E08.21, E08.22, E08.29, E08.311, E08.319, E08.321, E08.3211, E08.3212, E08.3213, E08.3219, E08.329, E08.3291, E08.3292, E08.3293, E08.3299, E08.331, E08.3311, E08.3312, E08.3313, E08.3319, E08.339, E08.3391, E08.3392, E08.3393, E08.3399, E08.341, E08.3411, E08.3412, E08.3413, E08.3419, E08.349, E08.3491, E08.3492, E08.3493, E08.3499, E08.351, E08.3511, E08.3512, E08.3513, E08.3519, E08.3521, E08.3522, E08.3523, E08.3529, E08.3531, E08.3532, E08.3533, E08.3539, E08.3541, E08.3542, E08.3543, E08.3549, E08.3551, E08.3552, E08.3553, E08.3559, E08.359, E08.3591, E08.3592, E08.3593, E08.3599, E08.36, E08.37X1, E08.37X2, E08.37X3, E08.37X9, E08.39, E08.40, E08.41, E08.42, E08.43, E08.44, E08.49, E08.51, E08.52, E08.59, E08.610, E08.618, E08.620, E08.621, E08.622, E08.628, E08.630, E08.638, E08.641, E08.649, E08.65, E08.69, E08.8, E08.9, E09.00, E09.01, E09.10, E09.11, E09.21, E09.22, E09.29, E09.311, E09.319, E09.321,

E09.3211, E09.3212, E09.3213, E09.3219, E09.329, E09.3291, E09.3292, E09.3293, E09.3299, E09.331, E09.3311, E09.3312, E09.3313, E09.3319, E09.339, E09.3391, E09.3392, E09.3393, E09.3399, E09.341, E09.3411, E09.3412, E09.3413, E09.3419, E09.349, E09.3491, E09.3492, E09.3493, E09.3499, E09.351, E09.3511, E09.3512, E09.3513, E09.3519, E09.3521, E09.3522, E09.3523, E09.3529, E09.3531, E09.3532, E09.3533, E09.3539, E09.3541, E09.3542, E09.3543, E09.3549, E09.3551, E09.3552, E09.3553, E09.3559, E09.359, E09.3591, E09.3592, E09.3593, E09.3599, E09.36, E09.37X1, E09.37X2, E09.37X3, E09.37X9, E09.39, E09.40, E09.41, E09.42, E09.43, E09.44, E09.49, E09.51, E09.52, E09.59, E09.610, E09.618, E09.620, E09.621, E09.622, E09.628, E09.630, E09.638, E09.641, E09.649, E09.65, E09.69, E09.8, E09.9, E10.10, E10.11, E10.21, E10.22, E10.29, E10.311, E10.319, E10.321, E10.3211, E10.3212, E10.3213, E10.3219, E10.329, E10.3291, E10.3292, E10.3293, E10.3299, E10.331, E10.3311, E10.3312, E10.3313, E10.3319, E10.339, E10.3391, E10.3392, E10.3393, E10.3399, E10.341, E10.3411, E10.3412, E10.3413, E10.3419, E10.349, E10.3491, E10.3492, E10.3493, E10.3499, E10.351, E10.3511, E10.3512, E10.3513, E10.3519, E10.359, E10.36, E10.37X1, E10.37X2, E10.37X3, E10.37X9, E10.39, E10.40, E10.41, E10.42, E10.43, E10.44, E10.49, E10.51, E10.52, E10.59, E10.610, E10.618, E10.620, E10.621, E10.622, E10.628, E10.630, E10.638, E10.641, E10.649, E10.65, E10.69, E10.8, E10.9, E11.00, E11.01, E11.10, E11.11, E11.21, E11.22, E11.29, E11.311, E11.319, E11.321, E11.3211, E11.3212, E11.3213, E11.3219, E11.329, E11.3291, E11.3292, E11.3293, E11.3299, E11.331, E11.3311, E11.3312, E11.3313, E11.3319, E11.339, E11.3391, E11.3392, E11.3393, E11.3399, E11.341, E11.3411, E11.3412, E11.3413, E11.3419, E11.349, E11.3491, E11.3492, E11.3493, E11.3499, E11.351, E11.3511, E11.3512, E11.3513, E11.3519, E11.3521, E11.3522,

E11.3523, E11.3529, E11.3531, E11.3532, E11.3533, E11.3539, E11.3541, E11.3542, E11.3543, E11.3549, E11.3551, E11.3552, E11.3553, E11.3559, E11.359, E11.3591, E11.3592, E11.3593, E11.3599, E11.36, E11.37X1, E11.37X2, E11.37X3, E11.37X9, E11.39, E11.40, E11.41, E11.42, E11.43, E11.44, E11.49, E11.51, E11.52, E11.59, E11.610, E11.618, E11.620, E11.621, E11.622, E11.628, E11.630, E11.638, E11.641, E11.649, E11.65, E11.69, E11.8, E11.9, E13.00, E13.01, E13.10, E13.11, E13.21, E13.22, E13.29, E13.311, E13.319, E13.321, E13.3211, E13.3212, E13.3213, E13.3219, E13.329, E13.3291, E13.3292, E13.3293, E13.3299, E13.331, E13.3311, E13.3312, E13.3313, E13.3319, E13.339, E13.3391, E13.3392, E13.3393, E13.3399, E13.341, E13.3411, E13.3412, E13.3413, E13.3419, E13.349, E13.3491, E13.3492, E13.3493, E13.3499, E13.351, E13.3511, E13.3512, E13.3513, E13.3519, E13.3521, E13.3522, E13.3523, E13.3529, E13.3531, E13.3532, E13.3533, E13.3539, E13.3541, E13.3542, E13.3543, E13.3549, E13.3551, E13.3552, E13.3553, E13.3559, E13.359, E13.36, E13.39, E13.40, E13.41, E13.42, E13.43, E13.44, E13.49, E13.51, E13.52, E13.59, E13.610, E13.618, E13.620, E13.621, E13.622, E13.628, E13.630, E13.638, E13.641, E13.649, E13.65, E13.69, E13.8, E13.9

ICD-10 Diagnosis Codes Used to Define Vision-Threatening Diabetic Retinopathy in Medicare, Medicaid, and MarketScan Claims Data

Code	Description
E08.341	Severe non-proliferative diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.341	Severe non-proliferative diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.341	Severe non-proliferative diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.341	Severe non-proliferative diabetic retinopathy with macular edema (type 2 diabetes mellitus)
E13.341	Severe non-proliferative diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.349	Severe non-proliferative diabetic retinopathy without macular edema (diabetes mellitus due to underlying condition)
E09.349	Severe non-proliferative diabetic retinopathy without macular edema (drug/chemical-induced diabetes mellitus)
E10.349	Severe non-proliferative diabetic retinopathy without macular edema (type 1 diabetes mellitus)
E11.349	Severe non-proliferative diabetic retinopathy without macular edema (type 2 diabetes mellitus)
E13.349	Severe non-proliferative diabetic retinopathy without macular edema (other specified diabetes mellitus)
E08.351	Proliferative diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.351	Proliferative diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.351	Proliferative diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.351	Proliferative diabetic retinopathy with macular edema (type 2 diabetes mellitus)
E13.351	Proliferative diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (diabetes mellitus due to underlying condition)

E09.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (drug/chemical induced diabetes mellitus)
E10.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (type 1 diabetes mellitus)
E11.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (type 2 diabetes mellitus)
E13.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (other specified diabetes mellitus)
E08.353	Proliferative diabetic retinopathy with traction retinal detachment not involving macula (diabetes mellitus due to underlying condition)
E09.353	Proliferative diabetic retinopathy with traction retinal detachment not involving macula (drug/chemical induced diabetes mellitus)
E10.353	Proliferative diabetic retinopathy with traction retinal detachment not involving the macula (type 1 diabetes mellitus)
E11.353	Proliferative diabetic retinopathy with traction retinal detachment not involving the macula (type 2 diabetes mellitus)
E13.353	Proliferative diabetic retinopathy with traction retinal detachment not involving the macula (other specified diabetes mellitus)
E08.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (diabetes mellitus due to underlying condition)
E09.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (drug/chemical induced diabetes mellitus)
E10.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (type 1 diabetes mellitus)
E11.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (type 2 diabetes mellitus)
E13.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (other specified diabetes mellitus)
E08.355	Stable proliferative diabetic retinopathy (diabetes mellitus due to underlying condition)
E09.355	Stable proliferative diabetic retinopathy (drug or chemical induced diabetes mellitus)
E10.355	Stable proliferative diabetic retinopathy (type 1 diabetes mellitus)
E11.355	Stable proliferative diabetic retinopathy (type 2 diabetes mellitus)
E13.355	Stable proliferative diabetic retinopathy (other specified diabetes mellitus)

E08.359	Proliferative diabetic retinopathy without macular edema (diabetes mellitus due to underlying condition)
E09.359	Proliferative diabetic retinopathy without macular edema (drug/chemical-induced diabetes mellitus)
E10.359	Proliferative diabetic retinopathy without macular edema (type 1 diabetes mellitus)
E11.359	Proliferative diabetic retinopathy without macular edema (type 2 diabetes mellitus)
E13.359	Proliferative diabetic retinopathy without macular edema (other specified diabetes mellitus)
E08.321	Mild non-proliferative diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.321	Mild non-proliferative diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.321	Mild non-proliferative diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.321	Mild non-proliferative diabetic retinopathy with macular edema (type 2 diabetes mellitus)
E13.321	Mild non-proliferative diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.331	Moderate non-proliferative diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.331	Moderate non-proliferative diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.331	Moderate non-proliferative diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.331	Moderate non-proliferative diabetic retinopathy with macular edema (type 2 diabetes mellitus)
E13.331	Moderate non-proliferative diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.311	Unspecified diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.311	Unspecified diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.311	Unspecified diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.311	Unspecified diabetic retinopathy with macular edema (type 2 diabetes mellitus)

E13.311	Unspecified diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.37X	Diabetic macular edema, resolved following treatment (diabetes mellitus due to underlying condition)
E09.37X	Diabetic macular edema, resolved following treatment (drug or chemical induced diabetes mellitus)
E10.37X	Diabetic macular edema, resolved following treatment (type 1 diabetes mellitus)
E11.37X	Diabetic macular edema, resolved following treatment (type 2 diabetes mellitus)
E13.37X	Diabetic macular edema, resolved following treatment (other specified diabetes mellitus)

ICD-10 Diagnosis Codes Used to Define Any Diabetic Retinopathy in Medicare, Medicaid, and MarketScan Claims Data

Code	Description
E08.341	Severe non-proliferative diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.341	Severe non-proliferative diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.341	Severe non-proliferative diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.341	Severe non-proliferative diabetic retinopathy with macular edema (type 2 diabetes mellitus)
E13.341	Severe non-proliferative diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.349	Severe non-proliferative diabetic retinopathy without macular edema (diabetes mellitus due to underlying condition)
E09.349	Severe non-proliferative diabetic retinopathy without macular edema (drug/chemical-induced diabetes mellitus)
E10.349	Severe non-proliferative diabetic retinopathy without macular edema (type 1 diabetes mellitus)
E11.349	Severe non-proliferative diabetic retinopathy without macular edema (type 2 diabetes mellitus)
E13.349	Severe non-proliferative diabetic retinopathy without macular edema (other specified diabetes mellitus)
E08.351	Proliferative diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.351	Proliferative diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.351	Proliferative diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.351	Proliferative diabetic retinopathy with macular edema (type 2 diabetes mellitus)
E13.351	Proliferative diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (diabetes mellitus due to underlying condition)

E09.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (drug/chemical induced diabetes mellitus)
E10.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (type 1 diabetes mellitus)
E11.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (type 2 diabetes mellitus)
E13.352	Proliferative diabetic retinopathy with traction retinal detachment involving the macula (other specified diabetes mellitus)
E08.353	Proliferative diabetic retinopathy with traction retinal detachment not involving macula (diabetes mellitus due to underlying condition)
E09.353	Proliferative diabetic retinopathy with traction retinal detachment not involving macula (drug/chemical induced diabetes mellitus)
E10.353	Proliferative diabetic retinopathy with traction retinal detachment not involving the macula (type 1 diabetes mellitus)
E11.353	Proliferative diabetic retinopathy with traction retinal detachment not involving the macula (type 2 diabetes mellitus)
E13.353	Proliferative diabetic retinopathy with traction retinal detachment not involving the macula (other specified diabetes mellitus)
E08.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (diabetes mellitus due to underlying condition)
E09.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (drug/chemical induced diabetes mellitus)
E10.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (type 1 diabetes mellitus)
E11.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (type 2 diabetes mellitus)
E13.354	Proliferative diabetic retinopathy with combined traction retinal detachment & rhegmatogenous retinal detachment (other specified diabetes mellitus)
E08.355	Stable proliferative diabetic retinopathy (diabetes mellitus due to underlying condition)
E09.355	Stable proliferative diabetic retinopathy (drug or chemical induced diabetes mellitus)
E10.355	Stable proliferative diabetic retinopathy (type 1 diabetes mellitus)
E11.355	Stable proliferative diabetic retinopathy (type 2 diabetes mellitus)
E13.355	Stable proliferative diabetic retinopathy (other specified diabetes mellitus)

E08.359	Proliferative diabetic retinopathy without macular edema (diabetes mellitus due to underlying condition)
E09.359	Proliferative diabetic retinopathy without macular edema (drug/chemical-induced diabetes mellitus)
E10.359	Proliferative diabetic retinopathy without macular edema (type 1 diabetes mellitus)
E11.359	Proliferative diabetic retinopathy without macular edema (type 2 diabetes mellitus)
E13.359	Proliferative diabetic retinopathy without macular edema (other specified diabetes mellitus)
E08.321	Mild non-proliferative diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.321	Mild non-proliferative diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.321	Mild non-proliferative diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.321	Mild non-proliferative diabetic retinopathy with macular edema (type 2 diabetes mellitus)
E13.321	Mild non-proliferative diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.331	Moderate non-proliferative diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.331	Moderate non-proliferative diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.331	Moderate non-proliferative diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.331	Moderate non-proliferative diabetic retinopathy with macular edema (type 2 diabetes mellitus)
E13.331	Moderate non-proliferative diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.311	Unspecified diabetic retinopathy with macular edema (diabetes mellitus due to underlying condition)
E09.311	Unspecified diabetic retinopathy with macular edema (drug or chemical induced diabetes mellitus)
E10.311	Unspecified diabetic retinopathy with macular edema (type 1 diabetes mellitus)
E11.311	Unspecified diabetic retinopathy with macular edema (type 2 diabetes mellitus)

E13.311	Unspecified diabetic retinopathy with macular edema (other specified diabetes mellitus)
E08.37X	Diabetic macular edema, resolved following treatment (diabetes mellitus due to underlying condition)
E09.37X	Diabetic macular edema, resolved following treatment (drug or chemical induced diabetes mellitus)
E10.37X	Diabetic macular edema, resolved following treatment (type 1 diabetes mellitus)
E11.37X	Diabetic macular edema, resolved following treatment (type 2 diabetes mellitus)
E13.37X	Diabetic macular edema, resolved following treatment (other specified diabetes mellitus)
E08.319	Unspecified diabetic retinopathy without macular edema (diabetes mellitus due to underlying condition)
E09.319	Unspecified diabetic retinopathy without macular edema (drug or chemical induced diabetes mellitus)
E10.319	Unspecified diabetic retinopathy without macular edema (type 1 diabetes mellitus)
E11.319	Unspecified diabetic retinopathy without macular edema (type 2 diabetes mellitus)
E13.319	Unspecified diabetic retinopathy without macular edema (other specified diabetes mellitus)
E08.329	Mild non-proliferative diabetic retinopathy without macular edema (diabetes mellitus due to underlying condition)
E09.329	Mild non-proliferative diabetic retinopathy without macular edema (drug or chemical induced diabetes mellitus)
E10.329	Mild non-proliferative diabetic retinopathy without macular edema (type 1 diabetes mellitus)
E11.329	Mild non-proliferative diabetic retinopathy without macular edema (type 2 diabetes mellitus)
E13.329	Mild non-proliferative diabetic retinopathy without macular edema (other specified diabetes mellitus)
E08.339	Moderate non-proliferative diabetic retinopathy without macular edema (diabetes mellitus due to underlying condition)
E09.339	Moderate non-proliferative diabetic retinopathy without macular edema (drug or chemical induced diabetes mellitus)
E11.339	Moderate non-proliferative diabetic retinopathy without macular edema (type 2 diabetes mellitus)

E10.339	Moderate non-proliferative diabetic retinopathy without macular edema (type 1 diabetes mellitus)
E13.339	Moderate non-proliferative diabetic retinopathy without macular edema (other specified diabetes mellitus)
E08.39	Other diabetic ophthalmic complication (diabetes mellitus due to underlying condition)
E09.39	Other diabetic ophthalmic complication (drug or chemical induced diabetes mellitus)
E10.39	Other diabetic ophthalmic complication (type 1 diabetes mellitus)
E11.39	Other diabetic ophthalmic complication (type 2 diabetes mellitus)
E13.39	Other diabetic ophthalmic complication (other specified diabetes mellitus)

eAppendix 5. Description of Validation and Verification

We assessed the validity of our model results by first estimating the prevalence of DR and VTDR in the NHANES 2005–2008 series of data using complete case analysis. We then evaluated the effect of five modeling steps to determine if the changes were in line with expectations and resulted in a logical final estimate. The five modeling steps were: 1) the use of imputation to adjust for missing data (eAppendix 3); 2) the incorporation of population-based study (PBS) and claims data to inform model estimates; 3) the use of NHANES 2017–March 2020 diabetes estimates, modeled with a logistic regression as described in eAppendix 2; 4) adjusting demographic composition of the NHANES population in 2005–2008 to match that of the US population in 2021 (while keeping the total population count constant); and 5) increasing the target population count in the model to match the US population in 2021.

For both DR and VTDR, adjusting for missing data resulted in a higher prevalence (see tables below). This finding is intuitively valid as the inability to grade a photo is likely related to older age and disease severity.[35,36] Adding PBS and claims data resulted in a lower prevalence of both DR and VTDR. Compared to the NHANES analysis with imputation, including PBS and claims data reduced the estimated point prevalence of DR from 29.71% to 27.07%, and VTDR from 6.24% to 5.14%, for people with diabetes 40 years or older. Using diabetes prevalence rates estimated from 2017–March 2020 NHANES data increased DR and VTDR counts in all subgroups, due to the increase in diabetes prevalence since the NHANES 2005–2008 data were collected. Changing the population's demographic composition from the NHANES 2005–2008 population weights to the 2021 US population

increased estimates of both DR and VTDR. Similarly, increasing the overall population count from 2005–2008 to 2021 increased the number of DR and VTDR cases. These changes are expected as the demographic composition of the US in 2021 is considerably older than the sample of non-institutionalized adults in the NHANES sample of 2005–2008 as the Baby Boomer generation (those born from 1946 to 1964) has aged into older age groups and the NHANES sample does not contain institutionalized individuals. Further, the full US population in 2021 (331.9 million) is approximately 9% larger than the population in 2005–2008 (304.1 million). Based on these results, we conclude that our model appropriately adjusted for missing data in NHANES, appropriately incorporated PBS and claims data, appropriately incorporated more recent diabetes prevalence rates, and correctly adjusted for population differences between the 2021 US population and the 2005–2008 NHANES sample.

The values in the final row in each table have very small differences from our final results. This is because these are the values before distribution to the county level and clipping to logical bounds, as described in eAppendix 2.

Impact of Modeling Steps on the Prevalence of Diabetic Retinopathy as Compared to Estimation using Complete Case Analysis in NHANES (2005–2008)

Model Step	DR count	DR rate	DR count (diabetes)	DR rate (diabetes)	DR count (diabetes 40+ years)	DR rate (diabetes 40+ years)	DR count (diabetes 80+ years)	DR rate (diabetes 80+ years)
Complete case analysis using NHANES population weights	–	–	–	–	6.82M	28.67%	0.58M	39.63%
Imputation for missing data using NHANES population weights	–	–	–	–	7.07M	29.71%	0.75M	34.31%
Adding PBS and claims data	5.90M	2.01%	5.90M	23.03%	5.30M	27.07%	0.44M	23.37%
Updating to NHANES 2017–March 2020 diabetes	6.90M	2.35%	6.90M	25.61%	6.33M	26.65%	0.53M	23.87%

prevalence								
Changing from NHANES population structure to 2021 population structure (including group quarters), but keeping total population constant	8.52M	2.89%	8.52M	26.43%	7.93M	27.38%	0.67M	24.79%
Scaling up to 2021 population including group quarters	9.61M	2.89%	9.61M	26.43%	8.94M	27.38%	0.75M	24.79%

Impact of Modeling Steps on the Prevalence of Vision-Threatening Diabetic Retinopathy as Compared to Estimation using Complete Case Analysis in NHANES (2005–2008)

Model Step	VTDR count	VTDR rate	VTDR count (diabetes)	VTDR rate (diabetes)	VTDR count (diabetes 40+ years)	VTDR rate (diabetes 40+ years)	VTDR count (diabetes 80+ years)	VTDR rate (diabetes 80+ years)
Complete case analysis using NHANES population weights	–	–	–	–	1.34M	5.62%	0.08M	5.14%
Imputation for missing data using NHANES population weights	–	–	–	–	1.49M	6.24%	0.11M	5.18%
Adding PBS and claims data	1.09M	0.37%	1.09M	4.27%	1.01M	5.14%	0.06M	3.19%
Updating to NHANES 2017–March 2020 diabetes prevalence	1.27M	0.43%	1.27M	4.70%	1.16M	4.88%	0.07M	3.12%
Changing from NHANES population structure to 2021 population structure (including group quarters), but keeping total population constant	1.63M	0.55%	1.63M	5.07%	1.52M	5.25%	0.09M	3.48%
Scaling up to 2021 population including group quarters	1.84M	0.55%	1.84M	5.07%	1.71M	5.25%	0.11M	3.48%

eTable. Diabetic Retinopathy (DR) and Vision-Threatening Diabetic Retinopathy (VTDR)

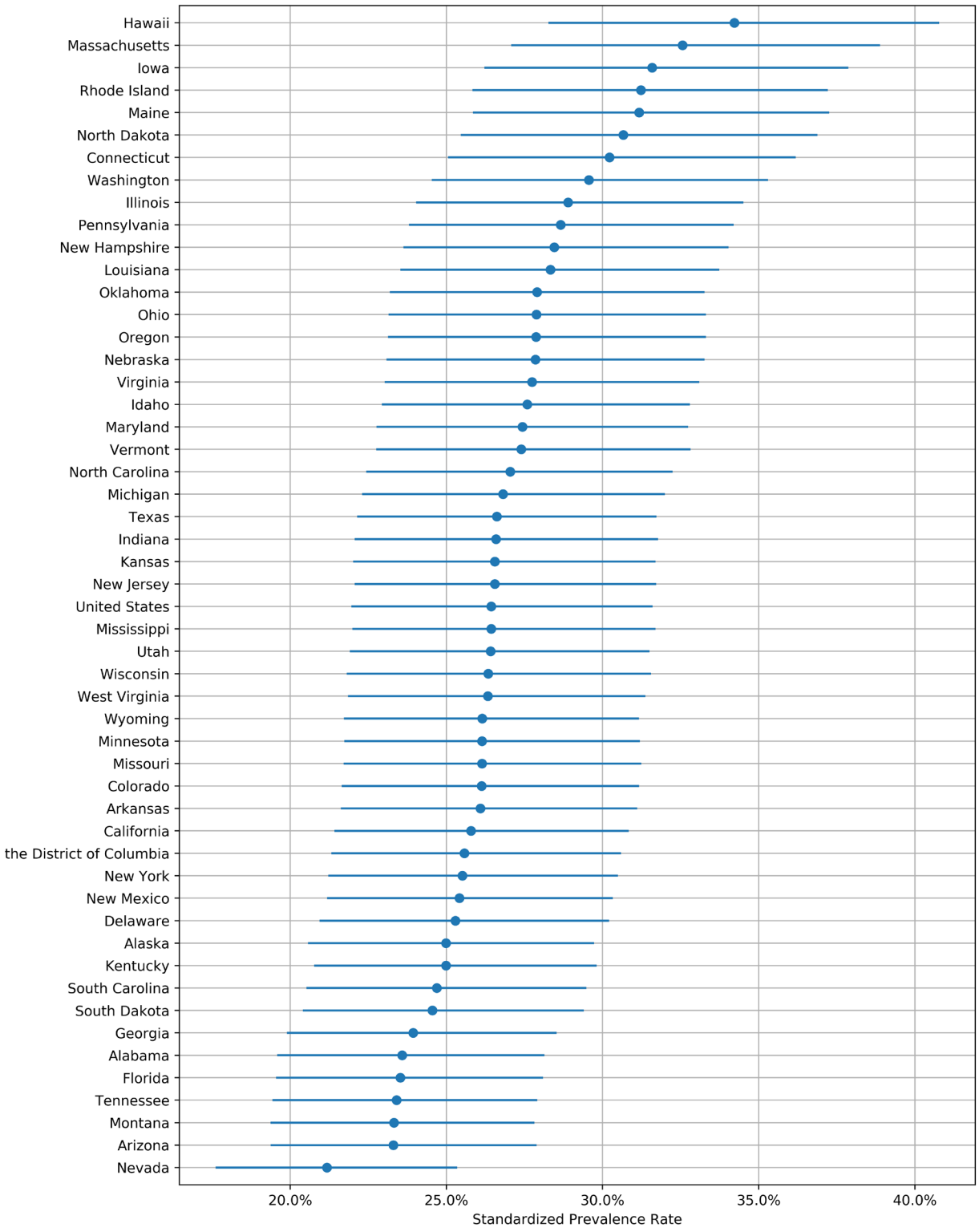
Prevalence Rates (%) Among Those with Diabetes by US State in 2021

State	DR (crude)	DR (standardized)	VTDR (crude)	VTDR (standardized)
USA	26.43 (21.95, 31.60)	26.43 (21.95, 31.60)	5.06 (3.90, 6.57)	5.06 (3.90, 6.57)
AK	22.41 (17.74, 28.20)	24.99 (20.57, 29.73)	4.18 (2.90, 5.97)	4.91 (3.78, 6.39)
AL	24.80 (20.35, 29.93)	23.58 (19.59, 28.14)	4.63 (3.41, 6.18)	4.33 (3.34, 5.62)
AR	26.34 (21.34, 32.28)	26.09 (21.62, 31.11)	4.60 (3.39, 6.13)	4.80 (3.72, 6.19)
AZ	22.83 (18.78, 27.61)	23.30 (19.37, 27.89)	4.36 (3.34, 5.72)	4.55 (3.50, 5.91)
CA	25.33 (21.04, 30.28)	25.79 (21.41, 30.84)	5.68 (4.30, 7.30)	5.34 (4.11, 6.97)
CO	25.44 (20.54, 30.92)	26.13 (21.65, 31.17)	4.69 (3.55, 6.29)	5.12 (3.93, 6.68)
CT	30.00 (24.36, 36.46)	30.22 (25.06, 36.19)	5.23 (3.94, 7.00)	5.53 (4.25, 7.20)
DC	28.25 (23.43, 33.91)	25.57 (21.31, 30.59)	7.25 (5.51, 9.62)	5.60 (4.31, 7.25)
DE	26.08 (21.37, 31.65)	25.29 (20.94, 30.21)	5.40 (4.02, 7.08)	5.30 (4.08, 6.92)
FL	24.37 (20.36, 28.82)	23.52 (19.55, 28.09)	4.67 (3.66, 5.97)	4.39 (3.38, 5.72)
GA	25.55 (21.34, 30.37)	23.93 (19.89, 28.53)	5.44 (4.13, 7.21)	4.74 (3.66, 6.14)
HI	28.99 (20.59, 40.02)	34.22 (28.27, 40.78)	5.51 (3.13, 9.26)	6.02 (4.62, 7.83)
IA	29.84 (23.28, 37.95)	31.59 (26.21, 37.87)	4.06 (2.77, 5.76)	5.08 (3.92, 6.63)
ID	25.92 (20.40, 32.56)	27.59 (22.94, 32.80)	3.74 (2.62, 5.27)	4.58 (3.52, 5.94)
IL	28.96 (23.90, 34.54)	28.89 (24.03, 34.51)	5.63 (4.34, 7.33)	5.64 (4.34, 7.31)
IN	26.02 (20.68, 32.43)	26.59 (22.06, 31.78)	4.17 (3.03, 5.68)	4.75 (3.65, 6.19)
KS	25.50 (20.22, 31.60)	26.55 (22.01, 31.70)	4.03 (2.99, 5.51)	4.68 (3.62, 6.08)
KY	24.06 (18.90, 30.46)	24.99 (20.76, 29.81)	3.80 (2.70, 5.34)	4.54 (3.49, 5.89)
LA	29.84 (24.74, 35.76)	28.33 (23.53, 33.74)	6.16 (4.62, 8.19)	5.56 (4.30, 7.22)

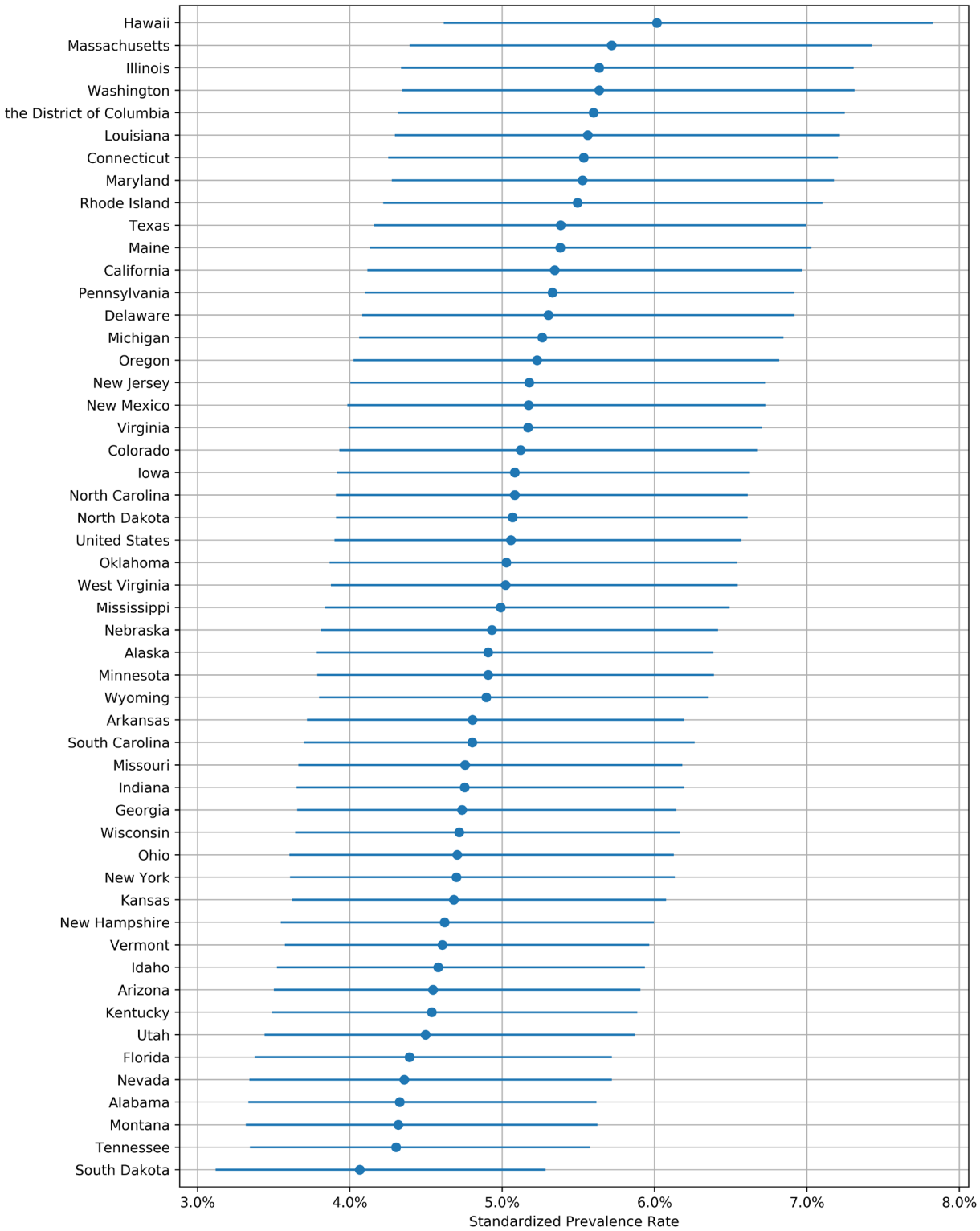
MA	31.31 (25.14, 38.30)	32.56 (27.08, 38.89)	5.11 (3.80, 6.93)	5.72 (4.39, 7.43)
MD	29.02 (24.25, 34.50)	27.43 (22.76, 32.74)	6.44 (4.92, 8.46)	5.53 (4.28, 7.18)
ME	29.71 (22.61, 38.23)	31.17 (25.85, 37.26)	4.13 (2.72, 6.18)	5.38 (4.13, 7.03)
MI	26.70 (21.55, 32.90)	26.81 (22.30, 31.99)	4.88 (3.57, 6.58)	5.26 (4.06, 6.85)
MN	24.97 (19.62, 31.62)	26.14 (21.73, 31.19)	4.05 (2.85, 5.65)	4.91 (3.79, 6.39)
MO	25.84 (20.59, 32.19)	26.14 (21.71, 31.24)	4.22 (3.07, 5.77)	4.76 (3.66, 6.18)
MS	28.99 (23.89, 34.77)	26.43 (21.99, 31.69)	5.85 (4.33, 7.87)	4.99 (3.84, 6.49)
MT	21.83 (16.92, 28.05)	23.32 (19.36, 27.82)	3.39 (2.30, 4.92)	4.32 (3.32, 5.63)
NC	27.83 (22.99, 33.39)	27.05 (22.44, 32.24)	5.25 (3.95, 6.88)	5.08 (3.91, 6.61)
ND	28.49 (22.16, 36.37)	30.66 (25.46, 36.88)	3.98 (2.72, 5.82)	5.07 (3.91, 6.61)
NE	26.43 (20.90, 32.90)	27.85 (23.08, 33.27)	4.12 (2.96, 5.68)	4.93 (3.81, 6.42)
NH	27.00 (20.76, 34.65)	28.46 (23.62, 34.04)	3.60 (2.40, 5.31)	4.62 (3.55, 6.00)
NJ	26.68 (22.28, 31.73)	26.55 (22.07, 31.71)	5.34 (4.13, 6.89)	5.18 (4.00, 6.73)
NM	25.62 (20.95, 30.79)	25.42 (21.18, 30.33)	5.57 (4.18, 7.32)	5.17 (3.98, 6.73)
NV	20.84 (17.45, 24.81)	21.17 (17.61, 25.34)	4.41 (3.39, 5.70)	4.36 (3.34, 5.72)
NY	25.86 (21.78, 30.57)	25.51 (21.22, 30.49)	4.94 (3.84, 6.32)	4.70 (3.61, 6.13)
OH	27.36 (21.82, 34.03)	27.88 (23.15, 33.31)	4.22 (3.06, 5.79)	4.71 (3.60, 6.13)
OK	26.41 (21.28, 32.41)	27.90 (23.19, 33.27)	4.46 (3.24, 5.99)	5.03 (3.87, 6.54)
OR	26.17 (20.66, 32.52)	27.87 (23.13, 33.31)	4.36 (3.15, 6.12)	5.23 (4.02, 6.82)
PA	28.18 (22.60, 34.89)	28.66 (23.80, 34.20)	4.79 (3.54, 6.41)	5.33 (4.10, 6.92)
RI	30.31 (24.14, 37.00)	31.23 (25.83, 37.21)	4.90 (3.64, 6.68)	5.49 (4.22, 7.10)
SC	25.76 (21.16, 31.17)	24.69 (20.52, 29.49)	5.09 (3.79, 6.74)	4.80 (3.70, 6.26)
SD	22.96 (17.84, 29.15)	24.55 (20.40, 29.40)	3.29 (2.27, 4.69)	4.07 (3.12, 5.29)
TN	23.42 (18.94, 28.79)	23.40 (19.43, 27.91)	4.07 (2.99, 5.47)	4.30 (3.34, 5.58)
TX	27.63 (23.13, 32.80)	26.61 (22.14, 31.73)	6.10 (4.76, 7.83)	5.38 (4.16, 7.00)

UT	24.66 (19.64, 30.65)	26.41 (21.91, 31.50)	3.74 (2.72, 5.22)	4.50 (3.44, 5.87)
VA	28.33 (23.56, 33.94)	27.74 (23.03, 33.10)	5.33 (4.04, 6.96)	5.17 (3.99, 6.71)
VT	25.97 (19.77, 33.59)	27.40 (22.76, 32.82)	3.54 (2.31, 5.31)	4.61 (3.57, 5.97)
WA	27.68 (22.05, 34.06)	29.56 (24.53, 35.29)	4.85 (3.52, 6.59)	5.64 (4.34, 7.31)
WI	25.59 (20.08, 32.27)	26.34 (21.81, 31.56)	3.98 (2.82, 5.52)	4.72 (3.64, 6.16)
WV	25.29 (19.52, 32.39)	26.33 (21.86, 31.37)	3.95 (2.65, 5.84)	5.02 (3.88, 6.55)
WY	24.98 (19.53, 31.83)	26.14 (21.72, 31.17)	3.99 (2.78, 5.60)	4.90 (3.80, 6.35)

eFigure 1. Age-, Sex/Gender-, and Race and Ethnicity-Standardized Diabetic Retinopathy Prevalence Among Those with Diabetes by US State in 2021, with 95% Uncertainty Intervals


Hawaii
Massachusetts
Iowa
Rhode Island
Maine
North Dakota
Connecticut
Washington
Illinois
Pennsylvania
New Hampshire
Louisiana
Oklahoma
Ohio
Oregon
Nebraska
Virginia
Idaho
Maryland
Vermont
North Carolina
Michigan
Texas
Indiana
Kansas
New Jersey
United States
Mississippi
Utah
Wisconsin
West Virginia
Wyoming
Minnesota
Missouri
Colorado
Arkansas
California
the District of Columbia
New York
New Mexico
Delaware
Alaska
Kentucky
South Carolina
South Dakota
Georgia
Alabama
Florida
Tennessee
Montana
Arizona
Nevada
20.0%
25.0%
30.0%
35.0%
40.0%
Standardized Prevalence Rate

eFigure 2. Age-, Sex/Gender-, and Race and Ethnicity-Standardized Vision-Threatening Diabetic Retinopathy Prevalence Among Those with Diabetes by US State in 2021, with 95% Uncertainty Intervals


Hawaii
Massachusetts
Illinois
Washington
the District of Columbia
Louisiana
Connecticut
Maryland
Rhode Island
Texas
Maine
California
Pennsylvania
Delaware
Michigan
Oregon
New Jersey
New Mexico
Virginia
Colorado
Iowa
North Carolina
North Dakota
United States
Oklahoma
West Virginia
Mississippi
Nebraska
Alaska
Minnesota
Wyoming
Arkansas
South Carolina
Missouri
Indiana
Georgia
Wisconsin
Ohio
New York
Kansas
New Hampshire
Vermont
Idaho
Arizona
Kentucky
Utah
Florida
Nevada
Alabama
Montana
Tennessee
South Dakota
3.0%
4.0%
5.0%
6.0%
7.0%
8.0%
Standardized Prevalence Rate

eReferences

1. Bauer GR. Sex and gender multidimensionality in epidemiologic research. *American Journal of Epidemiology*. 2023;192(1):122-132.

2. Cicero EC, Reisner SL, Merwin EI, Humphreys JC, Silva SG. The health status of transgender and gender nonbinary adults in the United States. *PloS one*. 2020;15(2):e0228765.

3. Baker KE. Findings from the behavioral risk factor surveillance system on health-related quality of life among US transgender adults, 2014-2017. *JAMA Internal Medicine*. 2019;179(8):1141-1144.

4. Flaxman AD, Vos DT, Murray CJ. *An Integrative Metaregression Framework for Descriptive Epidemiology*. University of Washington Press; 2015.

5. Patil A, Huard D, Fonnesbeck CJ. PyMC: Bayesian stochastic modelling in Python. *Journal of statistical software*. 2010;35(4):1.

6. Flaxman AD, Wittenborn JS, Robalik T, et al. Prevalence of visual acuity loss or blindness in the US: A Bayesian meta-analysis. *JAMA ophthalmology*. Published online 2021.

7. West SK, Klein R, Rodriguez J, et al. Diabetes and diabetic retinopathy in a Mexican-American population: Proyecto VER. *Diabetes care*. 2001;24(7):1204-1209.

8. Wong TY, Klein R, Islam FA, et al. Diabetic retinopathy in a multi-ethnic cohort in the United States. *American journal of ophthalmology*. 2006;141(3):446-455.

9. Varma R, Torres M, Peña F, et al. Prevalence of diabetic retinopathy in adult Latinos: The Los Angeles Latino eye study. *Ophthalmology*. 2004;111(7):1298-1306.

10. Mayer-Davis EJ, Davis C, Saadine J, et al. Diabetic retinopathy in the SEARCH for diabetes in youth cohort: A pilot study. *Diabet Med*. 2012;29(9):1148-1152.

11. Jensen ET, Rigdon J, Rezaei KA, et al. Prevalence, progression and modifiable risk factors for diabetic retinopathy in youth and young adults with youth-onset type 1 and type 2 diabetes: The SEARCH for Diabetes in Youth Study. *Accepted for publication in Diabetes Care*. 2023;XXX(XX):X-XX.

12. Levitsky LL, Danis RP, Drews KL, et al. Retinopathy in youth with type 2 diabetes participating in the TODAY clinical trial. *Diabetes Care*. 2013;36(6):1772-1774.

13. Varma R, Wen G, Jiang X, et al. Prevalence of diabetic retinopathy in adult Chinese American individuals: The Chinese American Eye Study. *JAMA Ophthalmology*. 2016;134(5):563-569.

14. Kovarik JJ, Eller AW, Willard LA, Ding J, Johnston JM, Waxman EL. Prevalence of undiagnosed diabetic retinopathy among inpatients with diabetes: The diabetic retinopathy inpatient study (DRIPS). *BMJ Open Diabetes Research and Care*. 2016;4(1):e000164.

15. Bjornstad Petter, Drews KL, Caprio S, et al. Long-term complications in youth-onset type 2 diabetes. *N Engl J Med*. 2021;385(5):416-426.

16. Gubitosi-Klug R, Libman I, Drews KL, et al. Development and progression of diabetic retinopathy in adolescents and young adults with type 2 diabetes: Results from the TODAY study. *Diabetes Care*. 2021;45(5):1049-1055.

17. Centers for Disease Control and Prevention. Vision and Eye Health Surveillance System: MarketScan. Published online 2022. Accessed September 8, 2022. https://www.cdc.gov/visionhealth/vehss/data/claims/marketscan.html

18. Centers for Disease Control and Prevention. Vision and Eye Health Surveillance System: Medicare. Published online 2022. Accessed September 8, 2022. https://www.cdc.gov/visionhealth/vehss/data/claims/medicare.html

19. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey (NHANES) 2005-2006. Accessed September 6, 2022. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2005

20. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey (NHANES) 2007-2008. Accessed September 6, 2022. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2007

21. Zheng P, Barber R, Sorensen RJ, Murray CJ, Aravkin AY. Trimmed constrained mixed effects models: Formulations and algorithms. *Journal of Computational and Graphical Statistics*. Published online 2021:1-13.

22. Health Services Resources Administration (HRSA). Area Health Resources Files: 2019-2020 County-Level Data. Published online 2020. Accessed July 22, 2022. https://data.hrsa.gov/data/download?data=AHRF

23. Centers for Medicare & Medicaid Services. MA [Medicare Advantage] state/county penetration 2018 06 [June 2018]. Published online 2020. Accessed December 14, 2020. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/MA-State-County-Penetration-Items/MA-State-County-Penetration-2018-06

24. Lin J, Thompson TJ, Cheng YJ, et al. Projection of the future diabetes burden in the United States through 2060. *Population health metrics*. 2018;16(1):1-9.

25. Akinbami LJ, Chen TC, Davy O, et al. National Health and Nutrition Examination Survey, 2017-March 2020 prepandemic file: Sample design, estimation, and analytic guidelines. *Vital and Health statistics Ser 1, Programs and Collection Procedures*. 2022;(190):1-36.

26. Capretto T, Piho C, Kumar R, Westfall J, Yarkoni T, Martin OA. Bambi: A simple interface for fitting Bayesian linear models in Python. *Journal of Statistical Software*. 2022;103(15):1-29. doi:10.18637/jss.v103.i15

27. US Census Bureau. 2019 American Community Survey 1-year Public Use Microdata Sample [CSV data file]. Published online 2020. Accessed July 14, 2021. https://www2.census.gov/programs-surveys/acs/data/pums/2019/1-Year

28. Dwyer-Lindgren L, Mackenbach JP, Van Lenthe FJ, Flaxman AD, Mokdad AH. Diagnosed and undiagnosed diabetes prevalence by county in the US, 1999–2012. *Diabetes care*. 2016;39(9):1556-1562.

29. Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological methods*. 2001;6(4):330.

30. Hardt J, Herke M, Leonhart R. Auxiliary variables in multiple imputation in regression with missing X: A warning against including too many in small sample research. *BMC medical research methodology*. 2012;12(1):1-13.

31. Noghrehchi F, Stoklosa J, Penev S, Warton DI. Selecting the model for multiple imputation of missing data: Just use an IC! *Statistics in Medicine*. 2021;40(10):2467-2497.

32. Kuhn M, Johnson K, others. *Applied Predictive Modeling*. Vol 26. Springer; 2013.

33. Honaker J, King G, Blackwell M. Amelia II: A program for missing data. *Journal of Statistical Software*. 2011;45(7):1-47. https://www.jstatsoft.org/v45/i07/

34. Seabold S, Perktold J. statsmodels: Econometric and statistical modeling with Python. In: *9th Python in Science Conference*.; 2010.

35. Silva PS, Horton MB, Clary D, et al. Identification of diabetic retinopathy and ungradable image rate with ultrawide field imaging in a national teleophthalmology program. *Ophthalmology*. 2016;123(6):1360-1367.

36. Scanlon PH, Foy C, Malhotra R, Aldington SJ. The influence of age, duration of diabetes, cataract, and pupil size on image quality in digital photographic retinal screening. *Diabetes Care*. 2005;28(10):2448-2453.