**Title:** Simulation of Alcohol Control Policies for Health Equity (SIMAH): Study Design and First Results

**Authors:** Charlotte Probst, Charlotte Buckley, Aurélie M. Lasserre, William C. Kerr, Nina Mulia, Klajdi Puka, Robin C. Purshouse, Yu Ye, Jürgen Rehm

**Correspondence Address:** Institute for Mental Health Policy Research, Centre for Addiction and Mental Health (CAMH), 33 Ursula-Franklin Street, Toronto, Ontario, M5S 2S1, Canada (e-Mail: charlotte.probst@camh.ca)

**Affiliations:** Institute for Mental Health Policy Research, Centre for Addiction and Mental Health (CAMH), 33 Ursula Franklin Street, Toronto, Ontario, Canada, M5S 2S1 (Charlotte Probst, Aurélie M. Lasserre, Klajdi Puka, Jürgen Rehm);
Heidelberg Institute for Global Health, Heidelberg University, Im Neuenheimer Feld 324, 69120 Heidelberg, Germany (Charlotte Probst);
Department of Psychiatry, University of Toronto, 250 College Street, Toronto, Ontario, Canada, M5T 1R8 (Charlotte Probst, Jürgen Rehm);
Institute of Clinical Psychology and Psychotherapy, Technische Universität Dresden, Chemnitzer Str. 46, 01187 Dresden, Germany (Jürgen Rehm);
Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, Ontario, Canada, M5T 3M7 (Jürgen Rehm);
Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, 33 Ursula-Franklin Street, Toronto, Ontario, Canada, M5S 2S1 (Charlotte Probst, Jürgen Rehm);
Center for Interdisciplinary Addiction Research, Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany (Jürgen Rehm);
Department of International Health Projects, Institute for Leadership and Health Management, I.M. Sechenov First Moscow State Medical University, Leontyevsky Pereulok 9, 125009 Moscow, Russian Federation (Jürgen Rehm);
Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK (Charlotte Buckley, Robin C. Purshouse);
Alcohol Research Group, Public Health Institute, 6001 Shellmound St, Suite 450, Emeryville, California 94608, US (William C. Kerr, Nina Mulia, Yu Ye).

## ABSTRACT

Since about 2010, life expectancy at birth in the United States (US) has stagnated and begun to decline with concurrent increases in the socioeconomic divide in life expectancy. The Simulation of Alcohol Control Policies for Health Equity (SIMAH) project uses a novel microsimulation approach to investigate the extent to which alcohol use, socioeconomic status (SES), and race/ethnicity contribute to unequal developments in US life expectancy and how alcohol control interventions could reduce such inequalities. Representative, secondary data from several sources will be integrated into one coherent, dynamic microsimulation to model life course changes in SES and alcohol use and cause-specific mortality attributable to alcohol use by SES, race/ethnicity, age, and sex. Markov models will be used to inform transition intensities between levels of SES and drinking patterns. The model will be used to compare a baseline to multiple counterfactual intervention scenarios. The preliminary results indicate that the crucial microsimulation component provides a good fit to observed demographic changes in the population, providing a robust baseline model for further simulation work. By demonstrating the feasibility of this novel approach, the SIMAH project promises to offer superior integration of relevant empirical evidence to inform public health policy for a more equitable future.

The Simulation of Alcohol Control Policies for Health Equity (SIMAH) project uses a novel microsimulation approach to investigate the extent to which alcohol use, socioeconomic status (SES), and race/ethnicity contribute to unequal developments in United States (US) life expectancy and how alcohol control interventions could reduce such inequalities. Microsimulation approaches have been successfully used to analyze population impacts of public health interventions. [1-4] However, the application of microsimulation techniques to matters of public health is only recently picking up speed.[5,6] This project will be the first, to the best of our knowledge, to use microsimulation to i) model alcohol-attributable mortality on the population level and ii) estimate policy impacts.

Since World War II, mortality rates in the US have been generally decreasing and, overall, individuals in later generations could expect to live longer and healthier lives than their predecessors.[7] However, in recent years, life expectancy at birth has stagnated and begun to decline in the US, even before the COVID-19 pandemic.[8,9]

One potential explanation for these recent trends are increases in premature mortality rates among specific demographic subgroups described by SES, race/ethnicity, age, and sex.[8] In particular, within some age groups of non-Hispanic White individuals, American Indian/Alaskan Native individuals, and individuals with low SES, all-cause mortality rates have increased by more than 1% annually over the past 15 years.[10,11]

Among the causes of death that are contributing most towards reductions in life expectancy since 2010 are poisoning, suicide, motor vehicle-related and other unintentional injuries, liver disease and cirrhosis, and diabetes mellitus[12] – all of which are causally linked to alcohol use.[13] Apart from being one of the most important risk factors for premature mortality,[14] alcohol use also contributes to socioeconomic inequalities in mortality.[15-18] The alcohol-attributable mortality risk follows a continuous dose-response relationship with SES[17] and the mortality gap between low and high SES is up to twice as large for alcohol-attributable compared to all-cause mortality.[18] Identifying and addressing alcohol-related health disparities is an important health equity issue. Alcohol-attributable disease burden and mortality are largely preventable and there are several effective population- and individual-level alcohol control interventions that can reduce the harmful use of alcohol.[19-21] The most cost-effective interventions at the population level are the regulation of prices through taxation or minimum unit prices (a floor price level for the retail sale of a beverage per unit of ethanol),[22] and the restriction of commercial and public availability of

alcohol.[23-25] On the individual level, screening and brief interventions are effective in decreasing the prevalence of harmful alcohol use.[26,27] While the overall effectiveness of such alcohol control interventions is firmly established, little attention has been paid to the equitable distribution of their health benefits and their ability to reduce health inequalities.[28-30]

### Objectives

Objective I of the SIMAH project is to investigate the extent to which alcohol use impacted mortality rates underlying the recent stagnation and declines in life expectancy and the increasing inequalities in life expectancy using a detailed microsimulation model of alcohol-attributable mortality in the US on the national level, as well as for 15 selected states. Objective II is to inform policy design by modeling future mortality reductions in different SES and race/ethnicity groups for alcohol control intervention strategies on a 10-year intervention planning horizon. This manuscript outlines the SIMAH study design, and presents initial results on generating a synthetic population that is representative of the adult US population in the year 2000 and simulating dynamics in the population over time under Objective I. No preliminary results will be presented for Objective II which will be addressed in the later years of the project.

## METHODS

### Study design and operationalizations

The study design (**Figure 1)** is based on an individual-level microsimulation approach[2,31] to model cause-specific mortality attributable to alcohol use for the years 2000 through 2018 with a forward modeling time-horizon until 2028. The microsimulation approach can systematically integrate disparate data sources into one coherent, dynamic simulation model of the individuals in a population that can be used to compare a baseline to multiple counterfactual scenarios.[6] The individuals in the microsimulation constitute a synthetic baseline population that is representative of the real US population on the national and state level. The advantage of microsimulation models is that they are "able to deal with detail complexity by simulating the life histories of individuals and then estimating the population effect from the sum of the individual effects" (p. 326).[32] The four interventions to be modeled are alcohol taxation, minimum unit pricing, regulation of the availability of alcohol and alcohol screening and brief interventions.

The target population of the study is the adult (18+) general population of the US. For state-level modeling 15 states will be used (California, Colorado, Florida, Indiana, Kentucky, Louisiana, Massachusetts, Michigan, Minnesota, Missouri, New York, Oregon, Pennsylvania, Tennessee, and Texas), covering more than half of the adult US population and all nine US census divisions. Educational attainment will be used as the main measure of SES. Although educational attainment is just one of several possible operationalizations of SES, it is the only one for which national mortality statistics are available. Furthermore, lower education is consistently associated with heavier drinking across racial-gender groups[33] and steep increases in the risk of alcohol-attributable mortality.[17] Educational attainment will be categorized into *high school degree or less*, *some college,* and *college degree or more.* Income, occupation, and employment status will be used in sensitivity analyses as additional indicators of SES. Race/ethnicity will be categorized as *non-Hispanic White*, *non-Hispanic Black*, *Hispanic,* and *Other*. A more detailed grouping will not be possible for the main analyses due to small sample size. However, sensitivity analyses will be performed with separate categories for *American Indian/Alaskan Native* and *Asian/Pacific Islander*.

Alcohol exposure will be categorized into five discrete drinking patterns based on the average grams of absolute alcohol per day (GPD), according to the standards of the World Health Organization:[34] abstainers (past 12 months), category I with up to 20/40 GPD for women/men, category II with 21-40/41-60 GPD for women/men, category III with 41-60/61-100 GPD for women/men, and category IV with >61/>100 GPD for women/men. Consumption by beverage will be represented as the proportion of beer, coolers, wine, and spirits of an individual's total consumption.[28]

Years of potential life lost (YLL) before the age of 75 will be the outcome measure. The following nine cause of death categories will be investigated: *alcohol use disorders* (including alcohol poisoning and other 100% alcohol-attributable causes of death); *motor vehicle accidents*; *other unintentional injuries; suicide*; *liver disease and cirrhosis*; *diabetes mellitus*; *ischemic heart disease*; *ischemic stroke;* and *hypertensive heart disease*. International Classification of Diseases, 10th revision (ICD-10) codes are shown in Web Table 1. Causes of death were selected to include major causes of death for which alcohol use is a risk factor.[13]

**Data sources**

Several data sources will be used and integrated to inform all microsimulation parameters (**Figure 1**). The data sources are summarized in **Table 1**, details are described in the Web Appendix 1.

Population estimates for each subgroup (defined by SES, race/ethnicity, age, and sex), will be based on decennial US Census data, the annual American Community Survey (ACS), and annual Current Population Surveys (March CPS).[35-37] Transitions between levels of educational attainment by subgroup will be informed by data from the Panel Study of Income Dynamics (PSID).[38,39]

Cause-specific mortality estimates in each subgroup will be based on individual death records obtained from the National Vital Statistics System (NVSS).[40]

Individual-level data on frequency and quantity of alcohol use from the Behavioral Risk Factor Surveillance System (BRFSS) will be used to inform alcohol exposure in each subgroup.[41] Data from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC, I and II) will be used to inform transition probabilities between drinking patterns by subgroup.[42,43] All three NESARC surveys (I-III) will be used to impute beverage preferences by subgroup. The aggregate total adult per capita consumption in litres of absolute alcohol[44] will be used to adjust for underreporting of quantity and frequency of alcohol use in population surveys.[45]

It should be noted that information on race and ethnicity are based on self-report in survey and census data (ACS, March CPS, BRFSS, NESARC, and Census). In PSID, race and ethnicity variables are only assessed for household heads and their wives/partners (self-report). The relationship of each participant to the household head will be used to reassign race and ethnicity for all participants that are not the household head or their wife/partner (see Web Appendix 1 for details). Mortality data are based on death records for which race ethnicity are typically filled out by the funeral director, who is asked to consult the decedent's next of kin, but may instead in some cases rely only on observation.[46,47]

Data on alcohol control interventions required for Objective II will come from BRFSS to inform the current state regarding the coverage of screening and brief intervention for harmful alcohol use. State-level data on the current implementation of alcohol taxation, minimum unit pricing, and regulation of alcohol availability will be obtained from recent work by the literature and the Alcohol Policy Information System (APIS).[48,49]

### Systematic literature reviews for parameter elicitation

Two systematic literature reviews will be performed to elicit model parameters of the microsimulation. Preference for inclusion will be given to high-quality studies based on US data reporting on parameters that are specific to SES, race/ethnicity, and sex. The first review will focus on relative risks of alcohol use on cause-specific mortality (Objective I), the second review will be performed to elicit evidence-based alcohol control intervention effects on individual drinking behavior (Objective II).

### Statistical analysis

***Objective I.*** A dynamic microsimulation model will be used to model life course changes in SES and alcohol use and cause-specific mortality attributable to alcohol use by SES, race/ethnicity, age, and sex.[50,51] The baseline date of the microsimulation will be January 1, 2000.

*Synthetic population*. The synthetic baseline population will be representative of the US population on the national and the state level in 2000 regarding the joint frequency distribution of SES, race/ethnicity, age, sex, and drinking patterns. The synthetic population will be created using iterative proportional fitting,[52,53] combining multiple data sources (**Figure 1**).

*Microsimulation*. The microsimulation model will progress the synthetic population forward in time by adding new individuals to account for births and inward migration and removing individuals to account for deaths and outward migration.[6] It will use transition probabilities to inform individual trajectories through levels of SES and drinking patterns. Annual transition probabilities will be estimated using continuous-time, multistate Markov models with transition intensity covariates adjusted for SES, race/ethnicity, age, and sex.[54]

*Simulation runs and uncertainty estimation*. The microsimulation will be calibrated within a Bayesian probabilistic framework that accounts for uncertainty in the evidence base. In this approach, the estimated transition probabilities for demographic change, drinking pattern change, and mortality are interpreted as prior beliefs about the true value of these model parameters. An Approximate Bayesian Computation[55] approach will be used to estimate posterior beliefs about these model parameters based on observed population-level change from an independent source (i.e., target data). Specifically, we will condition beliefs about model parameters using an 'implausibility' goodness-of-fit metric that accounts for the observed differences between simulated and target data, as well as the sampling uncertainty relating to both the simulation and target.[56] This calibrated model will then be used to quantify YLL in each subgroup of the

population. The proportion of YLL attributable to alcohol by cause of death, SES, race/ethnicity, age, and sex over time will be calculated using lifetime abstention as the counterfactual scenario.[57,58] Similarly, the proportion of inequalities in YLL rates concerning SES and race/ethnicity that can be explained by alcohol exposure over time will be estimated using the same microsimulation.

***Objective II.*** The microsimulation model will be expanded to the state level to generate projections regarding the impact of alcohol control interventions on cause-specific, age-standardized YLL rates, by SES, race/ethnicity, age, and sex. The intervention impacts will be modeled for a 10-year intervention planning horizon (2019 to 2028) for 15 selected states. The status quo (i.e., alcohol consumption under current alcohol policies) will be used as the baseline scenario to contrast the impact of reduced alcohol use under different counterfactual intervention scenarios. The intervention model is based on an existing conceptual framework for examining the impact of alcohol use on population health and disparities.[29] As such, we will investigate three upstream interventions: alcohol taxation (with and without inflation correction),[49] minimum unit pricing, and regulation of the availability of alcohol concerning spirits sales after 10 pm, as well as screening and brief interventions as a downstream intervention.

*Forecasting component*. To model YLL for a 10-year intervention planning horizon, the synthetic population will be projected through 2028. Individual-level receipt of screening and brief interventions will be added to the synthetic population using BRFSS data.[41,53] Projections before intervention will be calibrated to existing population and mortality projections.[59-61]

*Intervention component*. The estimated impacts of all relevant alcohol control interventions will be combined with the posterior estimates for transition probabilities between drinking patterns. As the effects of alcohol policies will be beverage-specific, their impact on the transition probabilities is mediated by beverage preferences. We will follow the approach used in the Sheffield Alcohol Policy Model to account for beverage preferences.[28] This will enable a forward estimation of the impact of alcohol control interventions on alcohol use and alcohol-attributable mortality under different counterfactual intervention scenarios compared to the baseline scenario (status quo). Machine learning methods[62] will be used to identify the most parsimonious intervention scenarios leading to a) the highest decreases in alcohol-attributable mortality overall; b) the highest reduction of inequality in mortality; and c) a reversal of the declining trends in life expectancy.

## PRELIMINARY FINDINGS

Preliminary analyses on the demographic foundation of the microsimulation model, including aging, transitioning between levels of SES (educational attainment), migration, and cause-specific mortality, were performed on the national level. The individual-level characteristics implemented in the synthetic population include drinking patterns, educational attainment, race/ethnicity, age, and sex. The synthetic population of 1,000,000 individuals was generated using iterative proportional fitting to integrate data from the BRFSS, US Census, ACS, and PSID and represented approximately 0.6% of adults in the US in 2000.

The synthetic population was projected forwards in time by adding new synthetic individuals each year and removing individuals based on Monte Carlo sampling using estimated net outward migration and mortality rates. Transition probabilities were used to simulate educational attainment by sex and race/ethnicity. The latter were estimated using a continuous-time, multi-state Markov model based on data from PSID.[63] Overall, the microsimulation model accounted for population developments between 2000 and 2018, including births, migration, and deaths in each subgroup of the population.[64] Preliminary findings shown do not include changes in drinking patterns and use mortality rates rather than YLL; these will be added in later iterations of the model. For the first microsimulation model shown here, uncertainty in the transition probabilities between levels of educational attainment was considered to demonstrate uncertainty estimates in the resulting population trends over time (see Web Appendix 2 for details). Population estimates based on ACS, Census, and PSID data were used as target data for comparison. Observed mortality rates calculated from NVSS and March CPS data and used as target data. While we did not perform a full model calibration within a Bayesian probabilistic framework yet, all model runs shown correspond to prior beliefs about the model parameters, all of which were informed by empirical data. Root mean squared errors (RMSE) were calculated to summarize differences between modelled and observed target data. Given the space limitation, our preliminary findings are focused on SES, results for race/ethnicity are shown in Web Tables 2, 3, and 4 and Web Figures 1 and 2.

**Figure 2** shows the prevalence of the four drinking categories in the synthetic baseline population and based on BRFSS data by sex and educational attainment. Overall, the prevalence of any current alcohol use matched the BRFSS target data well, with a higher prevalence in men

and women with higher levels of educational attainment. Among men, the prevalence of category III or category IV drinking was lower among those with higher educational attainment; the highest prevalence of category III or category IV drinking (combined) in the synthetic population was about 6% among men with high school degree or less. Among women, the prevalence of category III or category IV drinking was at around 2% for women with high school degree or less as well as women with college degree or more and at about 3% for women with some college education.

**Figure 3** depicts the microsimulation of the synthetic adult US population from 2000 to 2018. Specifically, the proportion of the population in each subgroup defined by educational attainment and sex is shown over time in comparison to observed data (US Census, ACS, and PSID). Some survey observations lie outside our uncertainty intervals (UI) due to differences in category definitions (see Web Appendix 1 for details). Overall, the proportion of individuals in each SES category in the microsimulation showed a good fit to the proportion of individuals in each education category by sex in the different data sources, namely, US Census (RMSE=1.7%), ACS (RMSE=7.1%), and PSID (RMSE=5.6%).

According to the microsimulation model, the proportion of men with a high school degree or less decreased from 47.3% in 2000 (US Census data) to 37.7% in 2018 (modelled data; 95% UI 34.2%-40.0%). The decreases were even stronger among women, declining from 47.0% (US Census data) in 2000 to 32.8% in 2018 (modelled data; 95% UI 30.2%-35.2%). The proportion of men with a college degree or more increased from 25.1% in 2000 (US Census data) to 30.2% in 2018 (modelled data; 95% UI 28.0%-34.3%). Over the same period, the proportion of women with college degree or more caught up with the proportion among men, increasing from 21.9% in 2000 (US Census data) to 31.6% in 2018 (modelled data; 95% UI 27.9%-36.4%). The microsimulation also showed a good fit with educational attainment data split by race/ethnicity and sex (Web Table 2, Web Figure 1).

Results for cause-specific mortality rates by sex and educational attainment modelled by the microsimulation and compared to observed data are shown in **Figure 4**. Mortality rates for all nine cause of death categories in the microsimulation model were a good fit to the observed data when split by education categories (Web Table 5). For individuals with a high school degree or less, the RMSE between modelled and observed mortality rates varied by causes between 3.8 deaths per 100,000 (ischemic heart disease) and 0.6 deaths per 100,000 (alcohol use disorders;

ischemic stroke). Model fit for individuals with some college was comparable for some categories with an RMSE of 0.7 deaths per 100,000 (alcohol use disorders) and worse for others with an RMSE of 15.5 deaths per 100,000 (ischemic heart disease). For individuals with college degree or more, all causes of death were well represented by the microsimulation and RMSE ranged from 0.4 (alcohol use disorders) to 2.9 deaths per 100,000 (ischemic heart disease).
The microsimulation showed (in accordance with observed data) that mortality rates for causes of death closely related to alcohol use increased between 2000 and 2018 for both sexes, with overall higher increases among individuals with only a high school degree or less (**Figure 4**). This included alcohol use disorders, liver disease and cirrhosis, suicide, and other unintentional injuries but not motor vehicle accidents. The most notable declines in mortality rates were observed for ischemic heart disease. The latter was also the cause of death with the largest absolute inequalities between education groups with a difference of 148 (men) and 65 (women) deaths per 100,000 for individuals with high school degree or less compared to those with college degree or more in 2000. This absolute rate difference declined to 116 (men) and 46 deaths per 100,000 (women) in 2018. In relative terms, the inequalities between individuals with a high school degree or less compared to individuals with a college degree or more increased universally for all nine causes of death, with some variation across the years. The largest relative inequalities among men were observed for motor vehicle accidents which increased from 2.9-fold higher rates among men with high school degree or less in 2000 to 5.2-fold higher rates in 2018. Among women, the relative inequalities were largest for diabetes mellitus starting with a 2.7-fold higher rate among women with high school degree or less compared to women with college degree or more in reaching 3.5-fold higher mortality rates in 2018.
Overall, the microsimulation was also a good fit to mortality rates split by race/ethnicity (Web Table 3, Web Figure 2) and to mortality rates split by race/ethnicity and education categories (Web Table 4). The model fit was best for non-Hispanic White individuals and worst for non-Hispanic Other (Web Table 3).

## DISCUSSION

Premature mortality in the US has recently been increasing among specific socio-demographic subgroups, and especially for causes of death closely related to alcohol use. The SIMAH project will use a rigorous approach applying innovative microsimulation methodology to investigate trends in alcohol-attributable mortality for major causes of death by SES, race/ethnicity, age, and

sex concurrently (Figure 5). This approach represents an advance over approaches commonly used such as traditional burden of disease analyses, based on comparative risk assessments.[65] The latter do not account for differences in exposure by factors such as SES, which can affect exposure, risk, or baseline mortality rates.[66] Furthermore, the SIMAH microsimulation approach will allow for dynamic modelling of diverse intervention scenarios to inform public health policy decisions.

The preliminary results indicate that the crucial microsimulation component provides a good fit to observed demographic changes in the population including changes in cause-specific mortality by sex, educational attainment, and race/ethnicity, providing a robust baseline model for further simulation work. The microsimulation provided a generally good fit to mortality rates for non-Hispanic White and Black individuals, but a comparatively poorer fit for the mixed race/ethnicity category of non-Hispanic Other. This is due to this group representing a smaller proportion of the population, and as the simulation relies on random number sampling this can lead to higher inaccuracy with smaller numbers. Additionally, the model showed a comparatively poorer fit for the "some college" educational category for some causes of death, i.e., ischemic heart disease. Such inaccuracies will be reduced in future modelling by running the model with a larger sample of individuals to improve model estimations.

As a model can only be as good as its input data, some limitations have to be acknowledged. First, representativeness and validity of survey data are limited by the sampling frame which may exclude parts of the population such as homeless individuals or individuals living in institutions, low and declining response rates,[67] and mis- or under-reporting, e.g., of alcohol use.[68,69] A key limitation of the mortality data is the assessment of education and race/ethnicity through a funeral director consulting next of kin, leading to inaccuracies in assessment and systematic differences compared to self-reported data.[46,47,70] This can lead to "dual data" bias due to a mismatch between population and mortality data, impacting, for example, the accuracy of the modelled inequalities.[71,72] Despite these limitations, the preliminary findings demonstrate the feasibility of this novel approach. With this, the current public health policy modeling paradigm of static, one-factor-at-a-time analyses can be supplemented by the new simulation approach proposed by SIMAH that offers superior integration of relevant empirical evidence.

In the next step, we will model transition probabilities between drinking patterns and the relationship between alcohol use and the specific causes of death (Objective I). In the later

phases of the project, scenarios of several alcohol-control interventions will be investigated to evaluate their ability to reverse current decreases in life expectancy with a 10-year forward modelling time-horizon (Objective II). Thus, the microsimulation can offer new perspectives on much-debated US public health policies (such as alcohol taxation) in addition to appraising two novel interventions for reducing inequalities in alcohol-related mortality: minimum unit pricing and a primary care program of screening and brief interventions for harmful alcohol use.[22] However, one key challenge will be to model the impacts of the COVID-19 pandemic that has affected nearly all aspects relevant to the model including alcohol consumption and socioeconomic health inequalities.[73,74]

The final microsimulation model will cast light on those subgroups of the population that experienced the highest increases in (alcohol-attributable) mortality, but that have been neglected by the broad-brushed modeling approaches currently available.[25] The use of Bayesian methods to propagate uncertainty in the evidence base through to simulation outputs will provide knowledge users with a robust perspective on the likely direction and magnitude of intervention impacts over time. Instead of providing results in the form of a single point estimate, the simulation model has the potential to flexibly analyze intervention scenarios upon request of stakeholders. To that end, the microsimulation is an open-source platform that can be expanded on and used by other researchers to explore the impacts of other exposure variables on alcohol-related mortality and major causes of death. As a result, a fine-tuned knowledge translation will be facilitated by the project that can be tailored to the needs of public health authorities on the state level.

## REFERENCES


1. Wilde P, Huang Y, Sy S, Abrahams-Gessel S, Jardim TV, Paarlberg R, et al. Cost-effectiveness of a US national sugar-sweetened beverage tax with a multistakeholder approach: Who pays and who benefits. Am J Public Health. 2019;109(2):276-84. PMCID: PMC6336039.
2. Basu S, Wagner RG, Sewpaul R, Reddy P, Davies J. Implications of scaling up cardiovascular disease treatment in South Africa: a microsimulation and cost-effectiveness analysis. Lancet Glob Health. 2019;7(2):e270-e80.
3. Basu S, Yudkin JS, Kehlenbrink S, Davies JI, Wild SH, Lipska KJ, et al. Estimation of global insulin use for type 2 diabetes, 2018-30: a microsimulation analysis. Lancet Diabetes Endocrinol. 2019;7(1):25-33.
4. Julien J, Ayer T, Bethea ED, Tapper EB, Chhatwal J. Projected prevalence and mortality associated with alcohol-related liver disease in the USA, 2019-40: a modelling study. The Lancet Public Health. 2020;5(6):e316-e23.
5. Zhang X. Application of discrete event simulation in health care: A systematic review. BMC Health Serv Res. 2018;18(1):687. PMCID: PMC6123911.
6. Brennan A, Buckley C, Vu TM, Probst C, Nielsen A, Bai H, et al. Introducing CASCADEPOP: an open-source sociodemographic simulation platform for US health policy appraisal. IJM. 2020;13(2):21-60.
7. Deaton A. *The great escape: Health, wealth, and the origins of inequality*. Princeton, USA: Princeton University Press; 2013.
8. Case A, Deaton A. Mortality and morbidity in the 21(st) century. Brookings Pap Econ Act. 2017;2017:397-476.
9. Woolf SH, Chapman DA, Sabo RT, Zimmerman EB. Excess Deaths From COVID-19 and Other Causes in the US, March 1, 2020, to January 2, 2021. JAMA. 2021.
10. Shiels MS, Chernyavskiy P, Anderson WF, Best AF, Haozous EA, Hartge P, et al. Trends in premature mortality in the USA by sex, race, and ethnicity from 1999 to 2014: An analysis of death certificate data. Lancet. 2017;389(10073):1043-54. PMCID: PMC5388357.
11. Chetty R, Stepner M, Abraham S, Lin S, Scuderi B, Turner N, et al. The association between income and life expectancy in the United States, 2001-2014. JAMA. 2016;315(16):1750-66. PMCID: PMC4866586.
12. Probst C, Könen M, Rehm J, Sudharsanan N. Alcohol-Attributable Deaths Help Drive Growing Socioeconomic Inequalities In US Life Expectancy, 2000-18. Health Affairs. 2022;41(8):1160-8.
13. Rehm J, Gmel GE, Gmel G, Hasan OSM, Imtiaz S, Popova S, et al. The relationship between different dimensions of alcohol use and the burden of disease-an update. Addiction. 2017;112(6):968-1001. PMCID: PMC5434904.
14. Shield K, Manthey J, Rylett M, Probst C, Wettlaufer A, Parry CDH, et al. National, regional, and global burdens of disease from 2000 to 2016 attributable to alcohol use: a comparative risk assessment study. Lancet Public Health. 2020;5(1):e51-e61.
15. Marmot M. *The Health Gap: The Challenge of an Unequal World*. London, UK: Bloomsbury Publishing; 2015.
16. Mackenbach JP, Kulhánová I, Bopp M, Borrell C, Deboosere P, Kovács K, et al. Inequalities in alcohol-related mortality in 17 European countries: A retrospective analysis of mortality registers. PLoS Med. 2015;12(12):e1001909. PMCID: PMC4666661.

17. Probst C, Lange S, Kilian C, Saul C, Rehm J. The dose-response relationship between socioeconomic status and alcohol-attributable mortality risk - a systematic review and meta-analysis. BMC Medicine. 2021;19:268.
18. Probst C, Roerecke M, Behrendt S, Rehm J. Socioeconomic differences in alcohol-attributable mortality compared with all-cause mortality: A systematic review and meta-analysis. Int J Epidemiol. 2014;43(4):1314-27. PMCID: PMC4258771.
19. Babor T, Caetano R, Casswell S, Edwards G, Giesbrecht N, Graham K. *Alcohol: No Ordinary Commodity: Research and Public Policy*. Oxford, UK: Oxford University Press; 2010.
20. Anderson P, Chisholm D, Fuhr DC. Effectiveness and cost-effectiveness of policies and programmes to reduce the harm caused by alcohol. Lancet. 2009;373(9682):2234-46.
21. Patel V, Chisholm D, Dua T, Laxminarayan R, Medina-Mora ML, Vos T, eds. *Disease Control Priorities, Third Edition (Volume 4): Mental, Neurological, and Substance Use Disorders*. Washington, DC: World Bank Publications; 2016.
22. O'Donnell A, Anderson P, Jané-Llopis E, Manthey J, Kaner E, Rehm J. Immediate impact of minimum unit pricing on alcohol purchases in Scotland: controlled interrupted time series analysis for 2015-18. BMJ. 2019;366:l5274.
23. World Health Organization. *Global action plan for the prevention and control of noncommunicable diseases 2013–2020*. Geneva, Switzerland: World Health Organization; 2013.
24. Wagenaar AC, Salois MJ, Komro KA. Effects of beverage alcohol price and tax levels on drinking: A meta-analysis of 1003 estimates from 112 studies. Addiction. 2009;104(2):179-90.
25. Chisholm D, Moro D, Bertram M, Pretorius C, Gmel G, Shield K, et al. Are the "Best Buys" for alcohol control still valid? An update on the comparative cost-effectiveness of alcohol control strategies at the global level. J Stud Alcohol Drugs. 2018;79(4):514-22.
26. Tansil KA, Esser MB, Sandhu P, Reynolds JA, Elder RW, Williamson RS, et al. Alcohol Electronic Screening and Brief Intervention: A Community Guide Systematic Review. Am J Prev Med. 2016;51(5):801-11. PMCID: PMC5082433.
27. Purshouse RC, Brennan A, Rafia R, Latimer NR, Archer RJ, Angus CR, et al. Modelling the Cost-Effectiveness of Alcohol Screening and Brief Interventions in Primary Care in England. Alcohol and Alcoholism. 2012;48(2):180-8.
28. Holmes J, Meng Y, Meier PS, Brennan A, Angus C, Campbell-Burton A, et al. Effects of minimum unit pricing for alcohol on different income and socioeconomic groups: A modelling study. Lancet. 2014;383(9929):1655-64. PMCID: PMC4018486.
29. Mulia N, Jones-Webb R. Alcohol policy: A tool for addressing health disparities? In: Giesbrecht N, Bosma L, editors. Preventing alcohol-related Problems: evidence and community-based initiatives. Washington, DC: APHA Press; 2017. p. 377-95.
30. Roche A, Kostadinov V, Fischer J, Nicholas R, O'Rourke K, Pidd K, et al. Addressing inequities in alcohol consumption and related harms. Health Promot Int. 2015;30(supplement 2):ii20-ii35.
31. Purshouse RC, Ally AK, Brennan A, Moyo D, Norman P. Evolutionary parameter estimation for a theory of planned behaviour microsimulation of alcohol consumption dynamics in an English birth cohort 2003 to 2010. Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation; July 12-16, 2014; Vancouver, BC, Canada: ACM; 2014. p. 1159-66.

32. Fone D, Hollinghurst S, Temple M, Round A, Lester N, Weightman A, et al. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. J Public Health Med. 2003;25(4):325-35.
33. Mulia N, Karriker-Jaffe KJ, Witbrodt J, Bond J, Williams E, Zemore SE. Racial/ethnic differences in 30-year trajectories of heavy drinking in a nationally representative U.S. sample. Drug Alcohol Depend. 2017;170:133-41. PMCID: PMC5270645.
34. World Health Organization. *International Guide for Monitoring Alcohol Consumption and Related Harm*. Geneva, Switzerland: World Health Organization; 2001.
35. Center for Economic and Policy Research. March CPS Uniform Extracts, Version 1.0. http://ceprdata.org/cps-uniform-data-extracts/march-cps-supplement/march-cps-data/. [Published 2018]. [Accessed 01/03/2019].
36. U.S. Census Bureau. *Current Population Survey Design and Methodology Technical Paper 77*. Washington, DC: U.S. Bureau of Labor Statistics and U.S. Census Bureau; 2019 https://www2.census.gov/programs-surveys/cps/methodology/CPS-Tech-Paper-77.pdf [Accessed 08/18/2022].
37. Ruggles S, Flood S, Goeken R, Grover J, Meyer E, Pacas J, et al. IPUMS USA: Version 10.0 [dataset]. In: IPUMS, editor. Minneapolis, MN2020.
38. Johnson DS, Freedman VA, Sastry N, McGonagle KA, Brown C, Fomby P, et al. Panel Study of Income Dynamics (PSID): Main Interview, 1968-2015. https://www.icpsr.umich.edu/web/DSDR/studies/37142/summary. [Published 2018]. [Accessed 08/18/2022].
39. McGonagle KA, Schoeni RF, Sastry N, Freedman VA. The Panel Study of Income Dynamics: Overview, Recent Innovations, and Potential for Life Course Research. Longit Life Course Stud. 2012;3(2). PMCID: PMC3591471.
40. Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying Cause of Death 1999-2016 on CDC WONDER Online Database, released December, 2017. http://wonder.cdc.gov/ucd-icd10.html. [Published 2017]. [Accessed 05/27/2018].
41. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System. http://www.cdc.gov/brfss. [Published 2018]. [Accessed 01/07/19].
42. Hasin DS, Grant BF. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: Review and summary of findings. Soc Psychiatry Psychiatr Epidemiol. 2015;50(11):1609-40. PMCID: PMC4618096.
43. Chen CM, Slater ME, Castle I-JP, Grant BF. *Alcohol use and alcohol use disorders in the United States: Main findings from the 2012–2013 National Epidemiologic Survey on Alcohol and Related Conditions-III (NESARC-III)*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2016 (NIH Publication No. 16-AA-8020) https://pubs.niaaa.nih.gov/publications/NESARC_DRM/NESARCDRM.pdf [Accessed 08/18/2-22].
44. Martinez P, Kerr WC, Subbaraman MS, Roberts SCM. New estimates of the mean ethanol content of beer, wine, and spirits sold in the U.S. show a greater increase in per capita alcohol consumption than previous estimates. Alcohol Clin Exp Res. 2019;43(3):509-21.
45. Kehoe T, Gmel G, Shield KD, Gmel G, Rehm J. Determining the best population-level alcohol consumption model and its impact on estimates of alcohol-attributable harms. Popul Health Metr. 2012;10:6. PMCID: PMC3352241.

46. Arias E, Heron M, Hakes J. The Validity of Race and Hispanic-origin Reporting on Death Certificates in the United States: An Update. Vital Health Stat 2. 2016(172):1-21.
47. Arias E, Schauman WS, Eschbach K, Sorlie PD, Backlund E. The validity of race and Hispanic origin reporting on death certificates in the United States. Vital Health Stat 2. 2008(148):1-23.
48. Bloss G. The Alcohol Policy Information System (APIS) and policy research At NIAAA. Alcohol Res Health. 2011;34(2):246-7. PMCID: PMC3860562.
49. Kerr W, Patterson D, Greenfield T. *State Alcohol Tax Rates. Spirits and Wine Tax Rates for the Control States: 2014 estimates based on retail price impact relative to license state pricing*. Alexandria, VA: National Alcohol Beverage Control Association (NABCA); 2016.
50. Li J. *Dynamic Microsimulation for Public Policy Analysis*. Maastricht, Netherlands: Boekenplan; 2011 https://books.google.ca/books?id=sHM8jN8hqWAC [Accessed 08/18/2022].
51. O'Donoghue C. *Handbook of Microsimulation Modelling*. Bingley, United Kingdom: Emerald Group Publishing Limited; 2014.
52. Lovelace R, Ballas D. 'Truncate, replicate, sample': A method for creating integer weights for spatial microsimulation. Comput Environ Urban Syst. 2013;41:1-11.
53. Lovelace R, Birkin M, Ballas D, van Leeuwen E. Evaluating the Performance of Iterative Proportional Fitting for Spatial Microsimulation: New Tests for an Established Technique. J Artif Soc Soc Simul. 2015;18(2):21.
54. Puka K, Buckley C, Mulia N, Purshouse RC, Lasserre AM, Kerr W, et al. Behavioral stability of alcohol consumption and sociodemographic correlates of change among a nationally representative cohort of US adults. Addiction. In press.
55. Jabot F, Faure T, Dumoulin N. EasyABC: performing efficient approximate Bayesian computation sampling schemes using R. Methods Ecol Evol. 2013;4(7):684-7.
56. Vernon I, Goldstein M, Bower R. Galaxy Formation: a Bayesian Uncertainty Analysis. Bayesian Analysis. 2010;5.
57. Rehm J, Room R, Monteiro M, Gmel G, Graham K, Rehn N, et al. Alcohol use. In: Ezzati M, Lopez AD, Rodgers A, Murray CJL, editors. Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. Geneva, Switzerland: World Health Organisation; 2004. p. 959-1108.
58. Ezzati M, Lopez AD, Rodgers A, Murray CJL, eds. *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Geneva, Switzerland: World Health Organization; 2004.
59. Hussar WJ, Bailey TM. *Projections of Education Statistics to 2026*. Washington, DC: National Center for Education Statistics; 2018 (NCES 2018-019).
60. U.S. Census Bureau. State 2017 National Population Projections Datasets. Projections for the United States: 2017 to 2060. https://www.census.gov/data/tables.html. [Published 2018]. [Accessed 01/24/2019].
61. Best AF, Haozous EA, de Gonzalez AB, Chernyavskiy P, Freedman ND, Hartge P, et al. Premature mortality projections in the USA through 2030: A modelling study. Lancet Public Health. 2018;3(8):e374-e84.
62. Deb K. *Multi-Objective Optimization Using Evolutionary Algorithms*. New York, New York: John Wiley & Sons; 2001.

63. Buckley C, Purshouse C, Lasserre AM, Puka K, Kerr WC, Mulia N, et al. The probability of transitioning to higher levels of education conditional on sex and race/ethnicity in the United States: a prospective cohort study from 1999 to 2019in review.
64. Manson S, Schroeder J, Riper DV, Ruggles S. IPUMS National Historical Geographic Information System: Version 14.0 [Database]. http://doi.org/10.18128/D050.V14.0. [Published 2019]].
65. Murray CJL, López AD. *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge, Massachusetts: Harvard University Press; 1996.
66. U.S. Burden of Disease Collaborators, Mokdad AH, Ballestros K, Echko M, Glenn S, Olsen HE, et al. The state of US health, 1990-2016: Burden of diseases, injuries, and risk factors among US states. JAMA. 2018;319(14):1444-72. PMCID: PMC5933332.
67. Schoeni RF, Stafford F, McGonagle KA, Andreski P. Response Rates in National Panel Surveys. The ANNALS of the American Academy of Political and Social Science. 2012;645(1):60-87.
68. Kilian C, Manthey J, Probst C, Brunborg GS, Bye EK, Ekholm O, et al. Why is per capita consumption underestimated in alcohol surveys? Results from 39 surveys in 23 European countries. Alc Alc. 2020;55(5):554–63.
69. Shield KD, Rehm J. Difficulties with telephone-based surveys on alcohol consumption in high-income countries: the Canadian example. Int J Methods Psychiatr Res. 2012;21(1):17-28. PMCID: PMC3561771.
70. Rostron BL, Boies JL, Arias E. Education reporting and classification on death certificates in the United States. Vital Health Stat 2. 2010(151):1-21.
71. Kunst AE, Groenhof F, Borgan JK, Costa G, Desplanques G, Faggiano F, et al. Socio-economic inequalities in mortality. Methodological problems illustrated with three examples from Europe. Rev épidémiol santé publique. 1998;46(6):467-79.
72. Hendi AS. Trends in Education-Specific Life Expectancy, Data Quality, and Shifting Education Distributions: A Note on Recent Research. Demography. 2017;54(3):1203-13.
73. Pollard MS, Tucker JS, Green HD, Jr. Changes in Adult Alcohol Use and Consequences During the COVID-19 Pandemic in the US. JAMA Network Open. 2020;3(9):e2022942-e.
74. Wyper GMA, Fletcher E, Grant I, Harding O, de Haro Moro MT, Stockton DL, et al. COVID-19 and prepandemic all-cause inequalities in disability-adjusted life-years due to multiple deprivation: a Scottish Burden of Disease study. The Lancet. 2021;398:S94.
75. U.S. Census Bureau. Educational Attainment in the United States: 2017. https://www.census.gov/data/tables/2017/demo/education-attainment/cps-detailed-tables.html. [Published 2018]. [Accessed 05/11/2018].

***Table 1.*** *Overview of data sources used in the SIMAH project*

| Data source;[a] host | Domain | Years | Design | Sampling |
|---|---|---|---|---|
| US Census; US Census Bureau[75] | Population | 2000, 2010 | Cross-sectional | Full assessment of the US population. |
| American Community Survey; US Census Bureau[37] | Population, migration | Annual, 2000-2018 | Cross-sectional | US civilian population, representative on the national and the state level. Institutionalized populations and people living in grouped quarters were included since 2006 |
| Current Population Surveys; US Census Bureau & US Bureau of Labor Statistics[35,36] | Population | Annual, 2000-2018 | Cross-sectional | US civilian, non-institutionalized population, representative on the national and the state level. |
| Panel Study of Income Dynamics; University of Michigan[38,39] | Education transitions | Biennial, 1999-2019 | Cohort | US civilian, non-institutionalized population, nationally representative |
| National Vital Statistics System; CDC[40] | Mortality | 2000-2018 | Registry | Full assessment of individual death records. |
| Behavioral Risk Factor Surveillance System; CDC[41] | Alcohol exposure, alcohol control interventions | Annual, 2000-2018 | Cross-sectional | US civilian, non-institutionalized population, representative on the national and the state level. |
| NESARC; National Institute on Alcohol Abuse and Alcoholism[42,43] | Alcohol exposure | 2001-2002, 2004-2005, 2012-2013 (NESARC I, II, and III) | Longitudinal (NESARC I and II), cross-sectional (NESARC III) | US civilian, non-institutionalized population. Nationally representative. |
| Alcohol per capita consumption; n/a[44] | Alcohol exposure | Annual, 2003-2016 | Modelled estimates | National and the state-level estimates. |
| Alcohol Policy Information System; Alcohol Policy Information System[48] | Alcohol control interventions | Exact dates, 2000-2018 | Collated information | State-level information. |

*a All data sources will include adults aged 18 years or older.*

*CDC Centers for Disease Control and Prevention; NESARC National Epidemiologic Survey on Alcohol and Related Conditions; SIMAH Simulation of Alcohol Control Policies for Health Equity; US United States*

## FIGURE CAPTIONS

***Figure 1.*** *Schematic overview of the study design, analysis steps, and data sources.*

ACS American Community Survey; APIS Alcohol Policy Information System; BRFSS Behavioral Risk Factor Surveillance System; NESARC National Epidemiologic Survey on Alcohol and Related Conditions; PSID Panel Study of Income Dynamics; SES socioeconomic status; YLL years of potential life lost

***Figure 2****. Prevalence of four drinking categories in the baseline synthetic population and Behavioral Risk Factor Surveillance System (BRFSS) data by sex and educational attainment in 2000. Category I: up to 20/40 (women/men) grams per day; category II: 21-40/41-60 (women/men) grams per day; category III: 41-60/61-100 (women/men) grams per day; category IV: >61/>100 (women/men) grams per day.*

***Figure 3.*** *The annual proportion of the subgroups of the population defined by sex and educational attainment as indicator of socioeconomic status over time as modelled via the microsimulation (2000 to 2018) compared to the Census (2000, 2010), the American Community Survey (ACS; annual data from 2000 to 2018) and the Panel Study of Income Dynamics (PSID; biannual data from 1999 to 2017). Grey shaded areas with dotted lines indicate the 95% uncertainty interval.*

***Figure 4.*** *Age-standardized mortality rates per 100,000 for nine cause of death categories between 2000 and 2018 as observed and (target data, dotted line) and as modelled by the microsimulation (solid line), by sex (men in black, women in gray) and socioeconomic status. Age-standardized to the population in 2018.*

AUD alcohol use disorders. IHD ischemic heart disease. HHD hypertensive heart disease. Liver C liver disease and cirrhosis. MVA motor vehicle accidents. Stroke ischemic stroke. UI unintentional injury.

***Figure 5.*** *Design figure.*