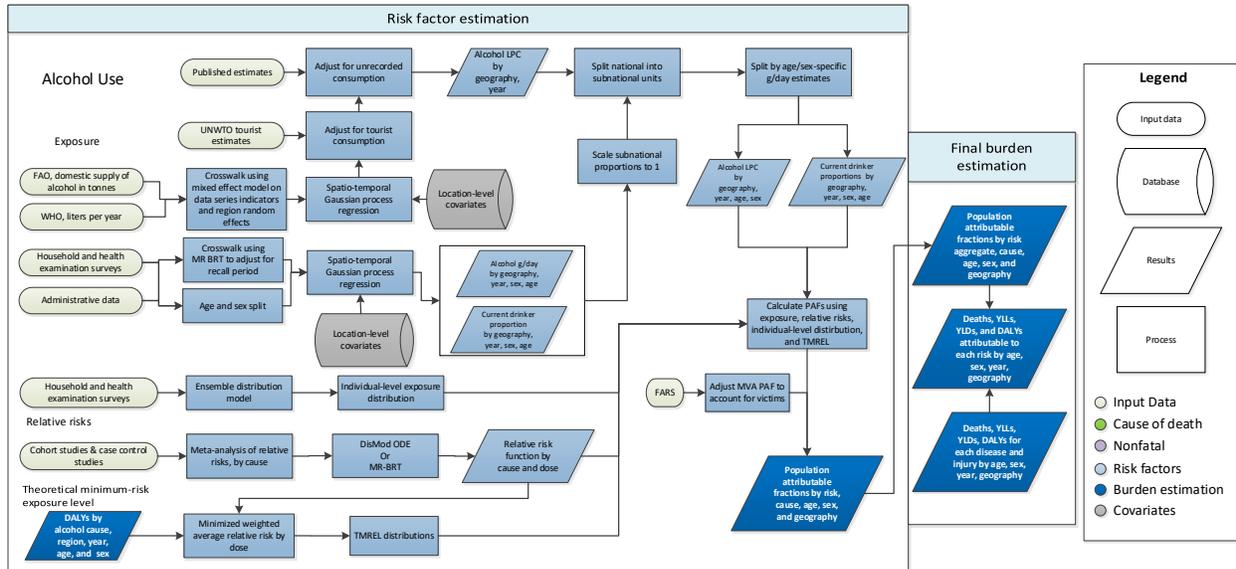


# High alcohol use

## Flowchart



## Input data and methodological summary

### Definition

### Exposure

High alcohol use is defined as alcohol consumption in excess of the theoretical minimum risk exposure level (TMREL), the level of alcohol consumption at which all-cause risk is minimised. Prior to GBD 2020, this risk factor was simply “Alcohol use” and quantified the burden of alcohol consumption over the entire exposure range. More details on the changes to the methodology can be found in the TMREL and “Population attributable fraction” sections of this appendix.

We defined exposure as the grams per day of pure alcohol consumed among current drinkers. We constructed this exposure using the indicators outlined below:

1. Current drinkers, defined as the proportion of individuals who have consumed at least one alcoholic beverage (or some approximation) in a 12-month period.
2. Alcohol consumption (in grams per day), defined as grams of alcohol consumed by current drinkers, per day, over a 12-month period.
3. Alcohol litres per capita (LPC) stock, defined in LBC of pure alcohol, over a 12-month period.

We also used three additional indicators to adjust alcohol exposure estimates to account for different types of bias:

1. Number of tourists within a location, defined as the total amount of visitors to a location within a 12-month period.
2. Tourists' duration of stay, defined as the number of days resided in a hosting country.
3. Unrecorded alcohol stock, defined as a percentage of the total alcohol stock produced outside established markets.

## Input data

### Exposure

A systematic review of the literature was performed to extract data on our primary indicators. The Global Health Exchange (GHDx), IHME's online database of health-related data, was searched for population survey data containing participant-level information from which we could formulate the required alcohol use indicators on current drinkers and alcohol consumption. Data sources were included if they captured a sample representative of the geographical location under study. We documented relevant survey variables from each data source in a spreadsheet and extracted using STATA 13.1 and R 3.3. A total of 6926 potential data sources were available in the GHDx, of which 5764 have been screened and 1206 accepted.

**Table 1: Data inputs for exposure for alcohol use**

	<b>Countries with data</b>	<b>New sources</b>	<b>Total sources</b>
Exposure	202	323	10,724

### Relative risk

For relative risks, in GBD 2016 we performed a systematic literature review of all cohort and case-control studies reporting a relative risk, hazard ratio, or odds ratio for any risk-outcome pairs studied in GBD 2016. Studies were included if they reported a categorical or continuous dose for alcohol consumption, as well as uncertainty measures for their outcomes, and the population under study was representative.

In GBD 2020, we undertook an effort to update the relative risk curves, beginning with six risk-outcome pairs that were among those associated with the greatest burden: ischaemic heart disease, ischaemic stroke, intracerebral haemorrhage, diabetes mellitus type II, lower respiratory infection, and tuberculosis. We refined the search strings to capture a larger number of studies than was identified by previous search strings. Studies published between 01/01/1970 and 12/31/2019 were reviewed. Of those articles captured, cohort and case-control studies were included if they reported an association between alcohol use and a GBD outcome, a continuous dose for alcohol consumption, and effect size (relative risk, hazard ratio, or odds ratio) with uncertainty. Information on study type, confounders controlled for, sample representativeness, and measurement of exposure and outcomes was also extracted.

**Table 2: Data inputs for relative risks for alcohol use**

	<b>Countries with data</b>	<b>New sources</b>	<b>Total sources</b>
Relative risks	63	110	566

## Data processing

Estimates of current drinking prevalence were split by age and sex where necessary. First, studies that reported prevalence for both sexes were split using a region-specific sex ratio estimated using meta-regression—Bayesian, regularised, trimmed (MR-BRT). Second, where studies reported estimates across non-GBD age groups, these were split into standard five-year age groups using the global age pattern estimated by ST-GPR.

**Table 3: MR-BRT sex splitting adjustment factors for current drinking**

Data input	Gamma	Beta coefficient, log (95% CI)	Adjustment factor*
Female: Male	0	-0.16 (-0.17 to -0.14)	0.85
Age <50	0	0.06 (0.06–0.06)	1.07
East Asia	0.36	-1.02 (-1.74 to -0.29)	0.36
Southeast Asia	0.64	-1.06 (-2.34 to 0.22)	0.35
Central Asia	0.41	-0.35 (-1.16 to 0.46)	0.70
Central Europe	0.18	-0.21 (-0.58 to 0.14)	0.80
Eastern Europe	0.10	-0.07 (-0.28 to 0.14)	0.93
High-income Asia Pacific	1.27	-1.11 (-4.90 to 2.68)	0.33
Western Europe	0.08	0.03 (-0.14 to 0.20)	1.03
Southern Latin America	1.26	-0.67 (-4.18 to 2.84)	0.51
High-income North America	0.09	-0.07 (-0.26 to 0.11)	0.93
Caribbean	0.25	-0.52 (-1.02 to -0.03)	0.59
Andean Latin America	0.76	-0.16 (-1.66 to 1.34)	0.85
Central Latin America	0.30	-0.52 (-1.12 to 0.08)	0.59
Tropical Latin America	0.08	-0.61 (-0.79 to -0.44)	0.54
North Africa and Middle East	1.21	-1.44 (-3.91 to 1.03)	0.24
South Asia	0.71	-1.17 (-2.57 to 0.23)	0.31
Eastern sub-Saharan Africa	0.28	-0.53 (-1.10 to 0.03)	0.58
Southern sub-Saharan Africa	0.20	-0.16 (-0.56 to 0.23)	0.85
Western sub-Saharan Africa	0.32	-0.19 (-0.83 to 0.45)	0.83
Oceania	0.94	-0.54 (-2.42 to 1.34)	0.58

\*Adjustment factor is the transformed beta coefficient in normal space and can be interpreted as the factor by which the alternative case definition is adjusted to reflect the ratio by which both-sex datapoints were split.

To allow for the inclusion of data that did not meet our reference definition for current drinking, two crosswalks were performed using MR-BRT. The first crosswalk converted estimates of one-month drinking prevalence to what they would be if data represented estimates of 12-month drinking prevalence. This crosswalk incorporated two binary covariates: male and age  $\geq 50$ . The second crosswalk converted estimates of one-week drinking prevalence to 12-month drinking prevalence. This crosswalk incorporated age  $< 20$  and male as covariates. The covariates utilised in both crosswalks were included as both x and z covariates. A uniform prior of 0 was set as the upper bound for the beta coefficients to enforce the logical constraint that one-month and one-week prevalence could not be greater than 12-month prevalence.

**Table 4: MR-BRT crosswalk adjustment factors for alcohol use current drinking model**

Data input	Reference or alternative case definition	Gamma	Beta coefficient, logit (95% UI)*	Adjustment factor**
12-month prevalence	Ref	---	---	---
1-month prevalence	Alt	0.22	-0.60 (-1.05, -0.16)	0.55 (0.35, 0.85)
Age ≥50		0.13	0.16 (-0.10, 0.43)	1.17 (0.9, 1.54)
Male		0.29	0.01 (-0.57, 0.59)	1.01 (0.57, 1.8)
1-week prevalence	Alt	0.46	-1.51 (-2.42, -0.59)	0.22 (0.09, 0.55)
Age <20		0.47	-0.29 (-1.34, 0.76)	0.75 (0.26, 2.14)
Male		0.00	0.38 (0.15, 0.60)	1.46 (1.16, 1.82)

\*MR-BRT crosswalk adjustments can be interpreted as the factor the alternative case definition is adjusted by to reflect what it would have been had it been measured using the reference case definition. If the log/logit beta coefficient is negative, then the alternative is adjusted up to the reference. If the log/logit beta coefficient is positive, then the alternative is adjusted down to the reference.

\*\*The adjustment factor column is the exponentiated beta coefficient. For log beta coefficients, this is the relative rate between the two case definitions. For logit beta coefficients, this is the relative odds between the two case definitions.

The raw data used in the supply-side model are domestic supply (WHO GISAH; FAO) and retail supply (Euromonitor) of litres of pure ethanol consumed. Domestic supply is calculated as the sum of production and imports, subtracting exports. The WHO and FAO sources were combined, so that FAO data were only used if there were no data available for that location-year from WHO. This was done because the WHO source takes into consideration FAO values when available. Since the WHO data are given in more granular alcohol types, the following adjustments were made:

$$LPC \text{ Pure Ethanol} = 0.13 * \left( \frac{Wine}{0.973} \right)$$

$$LPC \text{ Pure Ethanol} = 0.05 * \left( \frac{Beer}{0.989} \right)$$

$$LPC \text{ Pure Ethanol} = 0.4 * \left( \frac{Spirits}{0.91} \right)$$

Three outliering strategies are used to omit implausible datapoints and data that created implausible model fluctuations. First, estimates from the current drinking model are used to calculate the grams of alcohol consumed per drinker per day. A point is outliered if the grams of pure ethanol per drinker per day for a given source-location-year is greater than 100 (approximately 10 drinks). These thresholds were chosen by using expert knowledge about reasonable consumption levels. In the second round of outliering, the mean LPC value over a 10-year window is calculated. If a point is over 70% of that mean value away from the mean value, it is outliered. The 70% limit was chosen using histograms of these distances. Additionally, some manual outliering is performed to account for edge cases. Finally, data smoothing is performed by taking a three-year rolling mean over each location-year.

Next, an imputation to fill in missing years is performed for all series to remove compositional bias from our final estimates. Since the data from our main sources cover different time periods, by imputing a complete time series for each data series, we reduce the probability that compositional bias of the

sources is leading to biased final estimates. To impute the missing years for each series, we model the log ratio of each pair of sources as a function of an intercept and nested random effects on super-region, region, and location. The appropriate predicted ratio is multiplied by the source that we do have, which generates an estimated value for the missing source. For some locations where there was limited overlap between series, the predicted ratio did not make sense, and a regional ratio was used.

Finally, variance was calculated both across series (within a location-year) as well as across years (within a location-source). Additionally, if a location-year had one imputed point, the variance was multiplied by 2. If a location-year had two imputed points, the variance was multiplied by 4. The average estimates in each location-year were the input to an ST-GPR model. This uses a mixed-effects model modelled in log space with nested location random effects.

We obtained data on the number of tourists and their duration of stay from the UN World Tourism Organization.<sup>3</sup> We applied a crosswalk across different tourist categories, similar to the one used for the LPC data, to arrive at a consistent definition (ie, visitors to a country).

We obtained estimates on unrecorded alcohol stock from data available in WHO GISAH database,<sup>2</sup> consisting of 189 locations. For locations with no data available, the national or regional average was used.

## Modelling strategy

### Exposure

While population-based surveys provide accurate estimates of the prevalence of current drinkers, they typically underestimate real alcohol consumption levels.<sup>10-12</sup> As a result, we considered the LPC input to be a better estimate of overall volume of consumption. Per capita consumption, however, does not provide age- and sex-specific consumption estimates needed to compute alcohol-attributable burden of disease. Therefore, we use the age-sex pattern of consumption among drinkers modelled from the population survey data and the overall volume of consumption from FAO, GISAH, and Euromonitor to determine the total amount of alcohol consumed within a location. In the paragraphs that follow, we outline how we estimated each primary input in the alcohol exposure model, as well as how we combined these inputs to arrive at our final estimate of grams per day of pure alcohol. We estimated all models below using 1000 draws.

For data obtained through surveys, we used spatiotemporal Gaussian process regression (ST-GPR) to construct estimates for each location/year/age/sex. We chose to use ST-GPR due to its ability to leverage information across the nearby locations or time periods. We also modelled the alcohol LPC data, as well as the total number of tourists, using ST-GPR. To improve the LPC model fit in years beyond those in which data was available, we forecasted ST-GPR estimates using a damped holt function.

Given the heterogeneous nature of the estimates on unrecorded consumption, as well as the wide variation across countries and time periods, we took 1000 draws from the uniform distribution of the lowest and highest estimates available for a given country. We did this to incorporate the diffuse uncertainty within the unrecorded estimates reported. We used these 1000 draws in the equation below.

We adjusted the alcohol LPC for unrecorded consumption using the following equation:

$$\text{Alcohol LPC} = \frac{\text{Alcohol LPC}}{(1 - \% \text{ Unrecorded})}$$

We then adjusted the estimates for alcohol LPC for tourist consumption by adding in the per capita rate of consumption abroad and subtracting the per capita rate of tourist consumption domestically.

$$\text{Alcohol LPC}_d = \text{Unadjusted Alcohol LPC}_d + \text{Alcohol LPC}_{\text{Domestic consumption abroad}} - \text{Alcohol LPC}_{\text{Tourist consumption domestically}}$$

$$\text{Alcohol LPC}_i = \frac{\sum_l \text{Tourist Population}_l * \text{Proportion of tourists}_{i,l} * \text{Unadjusted Alcohol LPC}_l * \frac{\text{Average length of stay}_{i,l}}{365}}{\text{Population}_d}$$

where:

$l$  is the set of all locations,  $i$  is either Domestic consumption abroad or Tourist consumption domestically, and  $d$  is a domestic location.

After adjusting alcohol LPC by tourist consumption and unrecorded consumption for all location/years reported, sex-specific and age-specific estimates were generated by incorporating estimates modelled in ST-GPR for percentage of current drinkers within a location/year/sex/age, as well as consumption trends modelled in the ST-GPR grams per day model. We do this by first calculating the proportion of total consumption for a given location/year by age and sex, using the estimates of alcohol consumed per day, the population size, and the percentage of current drinkers. We then multiply this proportion of total stock for a given location/year/sex/age by the total stock for a given location/year to calculate the consumption in terms of LPC for a given location/year/sex/age. We then convert these estimates to be in terms of grams/per day. The following equations describe these calculations:

$$= \frac{\text{Proportion of total consumption}_{l,y,s,a} * \text{Alcohol g/day}_{l,y,s,a} * \text{Population}_{l,y,s,a} * \% \text{ Current drinkers}_{l,y,s,a}}{\sum_{s,a} \text{Alcohol g/day}_{l,y,s,a} * \text{Population}_{l,y,s,a} * \% \text{ Current drinkers}_{l,y,s,a}}$$

$$\text{Alcohol LPC}_{l,y,s,a} = \frac{\text{Alcohol LPC}_{l,y} * \text{Population}_{l,y} * \text{Proportion of total consumption}_{l,y,s,a}}{\% \text{ Current drinkers}_{l,y,s,a} * \text{Population}_{l,y,s,a}}$$

$$\text{Alcohol g/day}_{l,y,s,a} = \text{Alcohol LPC}_{l,y,s,a} * \frac{789 \text{ g/L}}{365}$$

where:

$l$  is a location,  $y$  is a year,  $s$  is a sex, and  $a$  is an age group.

We then used the gamma distribution to estimate individual-level variation within location, year, sex, age drinking populations, following the recommendations of other published alcohol studies.<sup>7,8</sup> We chose parameters of the gamma distribution based on the mean and standard deviation of the 1000 draws of alcohol g/day exposure for a given population. Standard deviation was calculated using the following formula.<sup>15</sup> We tested several alternative models using our data and found this model performed best.

$$\text{standard deviation} = \text{mean} * (0.087 * \text{female} + 1.171)$$

### Theoretical minimum risk exposure level

The methods for calculating the TMREL were updated for GBD 2020. Previously, one global estimate of the TMREL was calculated. However, the contributions of each cause to overall health loss vary over geography, age, time, and sex, suggesting that the amount of alcohol that minimises health loss similarly varies over these domains. For this reason, in GBD 2020 we estimated an individual TMREL for each region, age, sex, and year.

For each region, age, sex, and year, we calculated TMREL by first calculating the overall risk attributable to alcohol. We did this by weighting each relative risk curve by the share of overall DALYs for a given cause. We then took the minimum of this overall-risk curve as the TMREL of alcohol use. More formally,

$$TMREL = \text{argmin average overall risk}_{\omega}(g/day)$$

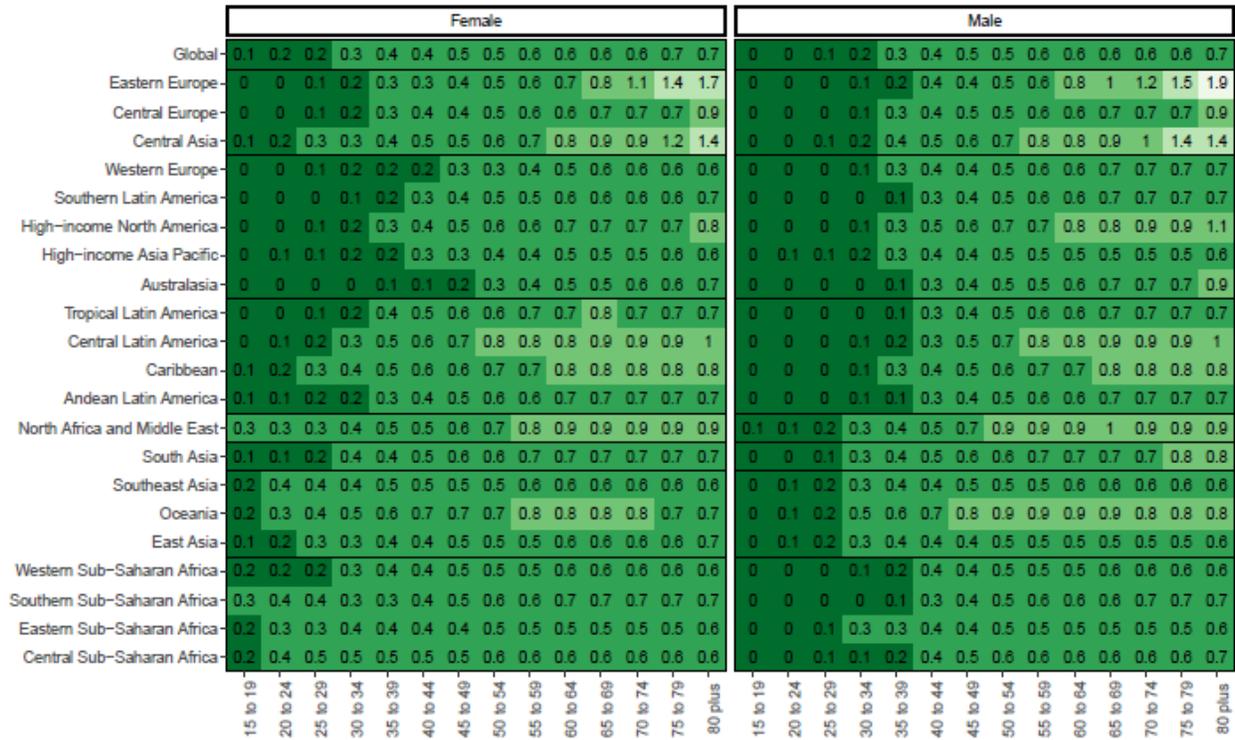
$$\text{Average overall risk}_{\omega,l,y,a,s}(g/day) = \sum_i^{\omega} \log(RR_i(g/day)) * \frac{DALY_{i,l,y,a,s}}{\sum_i^{\omega} DALY_{i,l,y,a,s}}$$

Where:

$\omega$  is the set of causes associated with alcohol,  $i$  is a given cause from that set,  $l$  is a location,  $y$  is a year,  $s$  is a sex,  $a$  is an age group, DALY is the DALY rate, and RR is the dose response curve for a given cause and exposure level in grams per day.

In other words, we chose TMREL as being the exposure that minimises the risk of suffering burden from any given cause related to alcohol. We weight the risk for a particular cause in our aggregation by the proportion of DALYs due to that cause (eg, since more observed people die from ischaemic heart disease, we weight the risk for ischaemic heart disease more in the above calculation of average risk compared to, say, diabetes, even if both have the same relative risk for a given level of consumption).

Figure 1: TMREL by region, age, and sex, 2020



### Relative risk

For GBD 2016 through 2019, we used the studies identified through a systematic review to calculate a dose–response, modelled using DisMod ODE. We chose DisMod ODE rather than a conventional mixed-effects meta-regression because of its ability to estimate non-parametric splines over doses (ie, for most alcohol causes, there is a non-linear relationship with different doses) and incorporate heterogeneous doses through dose-integration (ie, most studies report doses categorically in wide ranges. DisMod ODE estimates specific doses when categories overlap across studies, through an integration step.). We used the results of the meta-regression to estimate a non-parametric curve for all doses between zero and 100 g/day and their corresponding relative risks. For all causes, we assumed the relative risk was the same for all ages and sexes.

For GBD 2020, we used the studies identified through the updated systematic review to estimate new dose–response curves using MR-BRT for six outcomes among those associated with the greatest burden: ischaemic heart disease, ischaemic stroke, intracerebral haemorrhage, diabetes mellitus type II, lower respiratory infection, and tuberculosis. The relative risk curves for the remaining outcomes will be modelled using MR-BRT instead of DisMod ODE in the coming GBD rounds. Importantly, this new method takes into account the risk of biases in the relative risk estimation and incorporates unexplained between-study heterogeneity into the uncertainty of the relative risk estimates. The results of the meta-regression were used to estimate a non-parametric curve for all doses between zero and 100 g/day and their corresponding relative risks.

We implemented the Fisher Scoring correction to the heterogeneity parameter, which corrects for data-sparse situations. In such cases, the between-study heterogeneity parameter estimate may be 0, simply from lack of data. The Fisher Scoring correction uses a quantile of gamma, which is sensitive to the number of studies, study design, and reported uncertainty.

We have also added methodology that can detect and flag publication bias. The approach is based on the classic Egger's Regression strategy, which is applied to the residuals in our model. In the current implementation, we do not correct for publication bias, but flag the risk-outcome pairs where the risk for publication bias is significant.

In the table below, we list each risk-outcome pair that is updated in GBD 2020 along with several of the key modelling parameters and results. The formulation for MR-BRT is described in detail in the MR-BRT section of the appendix.

**Table 5: MR-BRT splines and priors by type of risk**

Risk-outcome	Type of risk	Spline degree, # interior knots	Priors and constraints
Ischaemic heart disease	J-shaped	Quadratic, 2   knots	No monotonicity constraint
Ischaemic stroke	J-shaped	Quadratic, 3   knots	No monotonicity constraint, right linear tail
Intracerebral haemorrhage	J-shaped	Cubic, 3   knots	No monotonicity constraint, right linear tail
Type II diabetes mellitus	J-shaped	Cubic, 3   knots	No monotonicity constraint, right linear tail
Tuberculosis	Harmful	Quadratic, 3   knots	Monotonic increasing, right linear tail, Gaussian max derivative prior on the right tail (0, 0.001)
Lower respiratory infection	Harmful	Quadratic, 3   knots	Monotonic increasing, right linear tail, Gaussian max derivative prior on the right tail (0, 0.001)

**Table 6: MR-BRT parameters by risk-outcome pair**

Risk-outcome	Type of risk	Selected covariates	Mean gamma solution	Publication bias result
Ischaemic heart disease	J-shaped	cv_incidence	0.158	No publication bias
Ischaemic stroke	J-shaped	cv_incidence	0.234	No publication bias
Intracerebral haemorrhage	J-shaped	cv_adjusted_2, cv_adjusted_1	0.09	No publication bias
Type II diabetes mellitus	J-shaped	None	0.117	No publication bias
Tuberculosis	Harmful	cv_sick_quitters, cv_incidence	19.488	No publication bias
Lower respiratory infection	Harmful	None	0	No publication bias

After evaluating all available evidence, we found insufficient evidence for a relationship between alcohol use and lower respiratory infection. Specifically, a simplified log-linear model was run, including only exposed and reference group dose data and study id as covariates, and a one-sided z-test was performed for the fixed-effects only model at alpha value set to 0.1. Based on this test, we removed

alcohol use vs. lower respiratory infection as a risk-outcome pair for GBD 2020. Regarding injuries outcomes, we constructed relative risks based on chronic exposure to alcohol rather than acute exposure immediately preceding injury, which has a weaker relationship to the outcome, though still significant.<sup>15,16,18-21</sup> We decided to use chronic exposure given the lack of available data on acute exposure, as well as the lack of cohort studies using acute exposure as a metric. Further, using chronic exposure allowed us to construct relative risks curves for unintentional injuries, interpersonal violence, motor vehicle accidents, and self-harm using the same method as reported above.

## Population attributable fraction

We calculated population attributable fractions (PAFs) by setting the relative risk of alcohol consumption among abstainers and drinkers consuming alcohol below the TMREL to be 1. We then calculated PAFs for drinkers consuming alcohol in excess of the TMREL as we have previously. For each location, age, sex, year, and cause, we defined PAF as:

$$PAF(x) = \frac{P_A + \int_0^{TMREL} P(x) dx + \int_{TMREL}^{100} P(x) * RR_C(x) dx - RR_C(TMREL)}{P_A + \int_0^{TMREL} P(x) dx + \int_{TMREL}^{100} P(x) * RR_C(x) dx} \quad P(x) = P_C * \Gamma(\mathbf{p})$$

where:

$P_C$  is the prevalence of current drinkers,  $P_A$  is the prevalence of abstainers, and  $\mathbf{p}$  are parameters determined by the mean and sd of exposure for that location, age, sex, and year;  $RR_C(x)$  is the global relative risk function for current drinkers for a given cause, and  $TMREL$  is the theoretical minimum risk exposure level for that location's region, age, sex, and year

We performed the above equation for 1000 draws of the exposure and relative risk models. We then used the estimated PAF draws to calculate YLL, YLDs, and DALYs, as per the other risk factors.

For outcomes that are by definition caused by alcohol, such as liver cancer or cirrhosis due to alcohol use, PAFs are set to 1. PAFs for cirrhosis due to all causes that are in excess of the proportion of all cirrhosis burden due to alcohol are proportionally redistributed over cirrhosis due to hepatitis B, hepatitis C, and other causes. Similarly, PAFs for liver cancer due to all causes that are in excess of the proportion of all liver cancer burden due to alcohol are proportionally redistributed over liver cancer due to hepatitis B, hepatitis C, and other causes.

In the case of motor vehicle accidents, we adjusted the PAF to account for victims of drunk drivers who are involved in accidents. Using data from the Fatality Analysis Reporting System (FARS) in the US,<sup>17</sup> we calculated the average number of fatalities in a car crash involving alcohol, as well as the percentage of those fatalities distributed by age and sex (Figures 2 and 3). We aggregated FARS data across the years

1985–2015, given there was little variation in the data temporally and the number of cases in old age groups had too much variance when constructing estimates by year. To adjust PAFs, we multiplied attributable deaths by the average number of fatalities from FARS and redistributed the PAF among each population, based on the probability of being a victim to a certain drunk driver by age and sex, based on the FARS data. The following equation describes this process:

$$Adjusted\ PAF_i = \frac{\sum_d PAF_d * DALY_d * Avg\ Fatalities_d * P(i\ is\ a\ victim)_d}{DALY_i}$$

where:

*i* is a population by location, year, age, sex and  
*d* is the set of all age and sex exposed groups within that location and year.

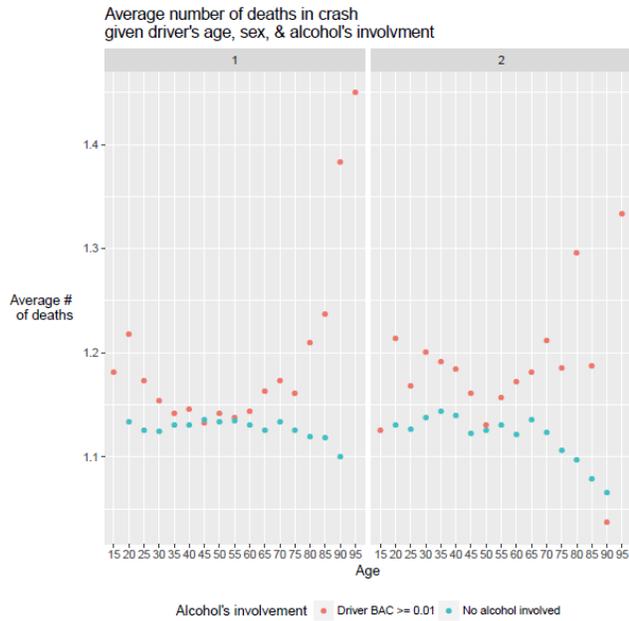


Figure 2

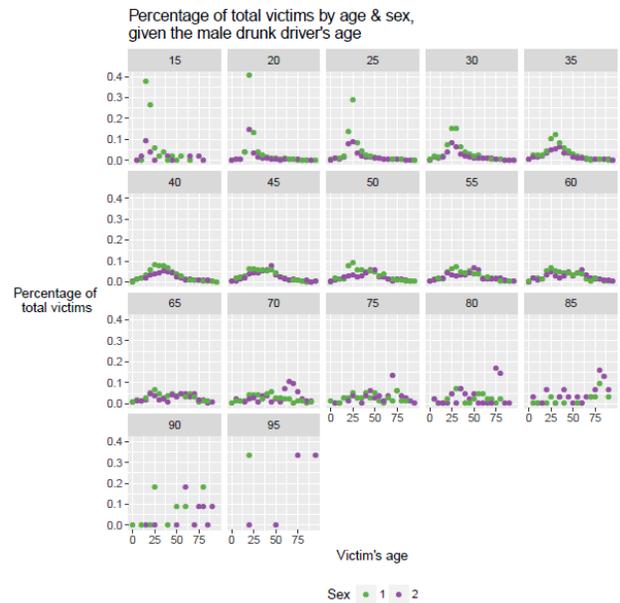


Figure 3

## Citations

1. Food and Agriculture Organization of the United Nations (FAO). FAOSTAT Food Balance Sheets, October 2014. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO).
2. World Health Organization (WHO). WHO Global Health Observatory - Recorded adult per capita alcohol consumption, Total per country. Geneva, Switzerland: World Health Organization (WHO).
3. UN World Tourism Organization (UNWTO). UN World Tourism Organization Compendium of Tourism Statistics 2015 [Electronic]. Madrid, Spain: UN World Tourism Organization (UNWTO), 2016.
4. Ramstedt, Mats. "How much alcohol do you buy? A comparison of self-reported alcohol purchases with actual sales." *Addiction* 105.4 (2010): 649-654.
5. Stockwell, Tim, et al. "Under-reporting of alcohol consumption in household surveys: a comparison of quantity–frequency, graduated–frequency and recent recall." *Addiction* 99.8 (2004): 1024-1033.
6. Kerr, William C., and Thomas K. Greenfield. "Distribution of alcohol consumption and expenditures and the impact of improved measurement on coverage of alcohol sales in the 2000 National Alcohol Survey." *Alcoholism: Clinical and Experimental Research* 31.10 (2007): 1714-1722.
7. Taylor, Bruce, et al. "The more you drink, the harder you fall: a systematic review and meta-analysis of how acute alcohol consumption and injury or collision risk increase together." *Drug and alcohol dependence* 110.1 (2010): 108-116.
8. Vinson, Daniel C., Guilherme Borges, and Cheryl J. Cherpitel. "The risk of intentional injury with acute and chronic alcohol exposures: a case-control and case-crossover study." *Journal of studies on alcohol* 64.3 (2003): 350-357.
9. Vinson, Daniel C., et al. "A population-based case-crossover and case-control study of alcohol and the risk of injury." *Journal of studies on alcohol* 64.3 (2003): 358-366.
10. Fatal Accident Reporting System (FARS). National Highway Traffic Safety Administration, National Center for Statistics and Analysis Data Reporting and Information Division (NVS-424); 1985, 1990, 1995, 2000, 2005, 2010, 2015
11. Chen, Li-Hui, Susan P. Baker, and Guohua Li. "Drinking history and risk of fatal injury: comparison among specific injury causes." *Accident Analysis & Prevention* 37.2 (2005): 245-251.
12. Bell, Nicole S., et al. "Self-reported risk-taking behaviors and hospitalization for motor vehicle injury among active duty army personnel." *American journal of preventive medicine* 18.3 (2000): 85-95.
13. Margolis, Karen L., et al. "Risk factors for motor vehicle crashes in older women." *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 57.3 (2002): M186-M191.
14. Sorock, Gary S., et al. "Alcohol-drinking history and fatal injury in older adults." *Alcohol* 40.3 (2006): 193-199.
15. Kehoe, Tara et al. "Determining the best population-level alcohol consumption model and its impact on estimates of alcohol-attributable harms." *Population health metrics* 10 6. (2012)