
This appendix provides further methodological detail, supplemental figures, and more detailed results for risk factors. The appendix is organized into broad sections following the structure of the main paper.
## Table of Contents

Preamble.................................................................................................................................................. 7

**Section 1. GBD overview**.................................................................................................................. 8

  - Geographic units of the analysis........................................................................................................ 8
  - GBD risk factor hierarchy.................................................................................................................... 8
  - Time periods of the analysis................................................................................................................ 8
  - List of abbreviations............................................................................................................................ 8

**Section 2. Risk factor estimation overview**....................................................................................... 11

  - Overview........................................................................................................................................ 11
  - Step 1. Effect size estimation............................................................................................................ 12
  - Step 2 Exposure estimation................................................................................................................ 15
    - DisMod-MR 2.1 estimation................................................................................................................ 18
    - Spatiotemporal Gaussian process regression................................................................................ 21
  - Step 3. Estimate summary exposure values..................................................................................... 25
  - Step 4. Theoretical minimum-risk exposure level......................................................................... 26
  - Step 5. Estimate population attributable fractions......................................................................... 26
  - Step 6. Mediation............................................................................................................................... 28
  - Step 7. Estimate attributable burden............................................................................................... 36
  - Other analysis: Decomposition of deaths and DALYs.................................................................... 36
  - Other analysis: Socio-demographic Index (SDI) analysis and epidemiological transition............ 38

References for Sections 1 and 2........................................................................................................... 40

**Section 3. Risk-specific estimation**.................................................................................................. 44

  - Unsafe water..................................................................................................................................... 45
  - Unsafe sanitation............................................................................................................................... 46
  - Unsafe hygiene................................................................................................................................. 49
  - Ambient particulate matter pollution............................................................................................. 52
  - Household air pollution.................................................................................................................... 62
  - Ambient ozone pollution.................................................................................................................. 64
Section 4. Supplemental Appendix Materials and Detailed Results for Risk Factors

Appendix figures

Appendix Figure 1. A more general causal web of the causes of health outcomes with the categories of causes included in this analysis shown in blue.

Appendix Figure 2. Analytical flowchart of the comparative risk assessment for the estimation of population attributable fractions by geography, age, sex, and year for GBD 2015. Ovals represent data inputs, rectangular boxes represent analytical steps, cylinders represent databases, and parallelograms represent intermediate and final results. GBD=Global Burden of Disease. SEVs=Summary exposure values. TMREL=Theoretical
minimum-risk exposure level. PAFs=Population attributable fractions. YLLs=years of life lost. YLDs=years lived with disability. DALYs=disability-adjusted life-years.

Appendix Figure 3. Types of Comparative Risk Assessments (CRA) based on the time perspective and the nature of the counterfactual level or distribution of exposure. The shaded box represents the type of CRA currently undertaken in GBD 2015. GBD=Global Burden of Disease.

Appendix Figure 4. Global decomposition of changes in level 3 cause-specific DALYs for all risk factors combined from 1990 to 2015 due to population growth, population ageing, risk exposure and the risk-deleted DALY rate. Causes are reported in order of percent change in the number of DALYs from 1990 to 2015. This figure excludes cervical cancer, HIV/AIDS, and sexually transmitted diseases DALYs because they are not estimated based on exposure and relative risk. DALYs=disability-adjusted life-years.

Appendix Figure 5. Global age-standardised percent change in SEVs for high and high-middle Socio-demographic Index (SDI) geographies versus middle, low-middle, and low SDI geographies, for males (A) and females (B), 1990 to 2015. Socio-demographic Index (SDI) is calculated for each geography as a function of lag dependent income per capita, average educational attainment in the population over age 15, and the total fertility rate (TFR). SDI units are interpretable; a zero represents the lowest level of income per capita and educational attainment and highest TFR observed from 1980 to 2015, and a one represents the highest income per capita and educational attainment and lowest TFR observed in the same period. Cut-offs on the SDI scale for the quintiles have been selected based on examining the entire distribution of geographies from 1980 to 2015. Annualised rate of change in the age-standardised SEV 1990-2015 in high SDI geographies compared to all other geographies. SEV=summary exposure value.

Appendix Figure 6. Diagram showing the proportion of all-cause DALYs to behavioural, environmental and occupational, and metabolic risk factors and their overlaps for all ages in 2015. DALYs=disability-adjusted life-years.

Appendix Figure 7. DALYs attributable to level 2 risk factors for the low Socio-demographic Index (SDI) quintile (A) and for the high SDI quintile (B), for both sexes combined, 2015. Socio-demographic Index (SDI) is calculated for each geography as a function of lag dependent income per capita, average educational attainment in the population over age 15, and the total fertility rate (TFR). SDI units are interpretable; a zero represents the lowest level of income per capita, educational attainment, and highest TFR observed 1980-2015 and a one represents the highest income per capita, educational attainment, and lowest TFR observed in the same period. Cut-offs on the SDI scale for the quintiles have been selected based on examining the entire distribution of geographies 1980-2015. Annualized rate of change in the age-standardized SEV 1990-2015 in high SDI geographies compared to all other geographies. DALYs=disability-adjusted life-years.

Appendix Figure 8. Leading 10 level 3 global risk factors for DALYs in 2015 by age group. Each cause is colored by the percent change in age specific DALYs from 2005 to 2015. DALYs=disability-adjusted life-years.
Appendix Figure 9. Observed SEVs compared to the value expected on the basis of SDI alone, across SDI quintiles for 61 risk factors included in the GBD 2015. Each SDI quintile is coloured-coded, and coloured lines represent expected levels, on the basis of SDI, for risk-specific SEVs. To enhance readability, SEVs in this figure have been scaled such that the lowest observed SEV for a given risk equals 0 and the highest observed SEV equals 1. Each circular symbol represents observed SEVs at the country level in 2015, with colours aligning with SDI quintile. Each risk factor corresponds with a vertical line. The ordering of risk factors was determined by the difference in expected SEVs for low SDI (the red line) and high SDI (the purple line). Risks proceed clockwise from those with the largest decline in SEV to those with the largest increase in SEV as SDI increases. SDI = Socio-demographic index. SEV = Summary exposure value.

Appendix Figure 10. Global map for Level 3 risk factors in 2013 of attributable DALYs for males (A) and females (B). DALYs = disability-adjusted life-years. ATG = Antigua and Barbuda. VCT = Saint Vincent and the Grenadines. BRB = Barbados. COM = Comoros. DMA = Dominica. GRD = Grenada. MDV = Maldives. MUS = Mauritius. LCA = Saint Lucia. TTO = Trinidad and Tobago. SYC = Seychelles. MLT = Malta. SGP = Singapore. MHL = Marshall Islands. KIR = Kiribati. SLB = Solomon Islands. FSM = Federated States of Micronesia. VUT = Vanuatu. WSM = Samoa. FJI = Fiji. TON = Tonga.

Appendix Figure 11. Map of socio-demographic index (SDI) classifications by location. ATG = Antigua and Barbuda. VCT = Saint Vincent and the Grenadines. BRB = Barbados. COM = Comoros. DMA = Dominica. GRD = Grenada. MDV = Maldives. MUS = Mauritius. LCA = Saint Lucia. TTO = Trinidad and Tobago. SYC = Seychelles. MLT = Malta. SGP = Singapore. MHL = Marshall Islands. KIR = Kiribati. SLB = Solomon Islands. FSM = Federated States of Micronesia. VUT = Vanuatu. WSM = Samoa. FJI = Fiji. TON = Tonga.

Appendix tables

Appendix Table 1. Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) 18-items checklist with description of compliance and location of information for GBD 2015 risk factors capstone.

Appendix Table 2. GBD 2015 geography hierarchy with levels.

Appendix Table 3. GBD 2015 risk factor hierarchy with levels, modeling strategies, and the main type of data sources used to estimate exposure levels.

Appendix Table 4. Socio-demographic Index (SDI) groupings by geography, based on 2015 values.

Appendix Table 5. Socio-demographic Index (SDI) values for all estimated GBD locations, 1980 to 2015.

Appendix Table 6a. Relative risks used by age and sex and for each outcome for all risk factors except for ambient air pollution and alcohol.

Appendix Table 6b. Relative risks used by age and sex and for each outcome for the particulate matter integrated exposure response curve.

Appendix Table 6c. Relative risks used by age and sex and for each outcome for alcohol use.
Appendix Table 7. All-age deaths, death PAFs, DALYs, DALY PAFs, and median percent change for each risk factor and outcome for both sexes combined in 2005 and 2015, by geography..... 69

Appendix Table 8. Age-standardised PAFs attributable to the joint distribution of all risk factors for all ages, both sexes combined for each GBD Level 1 and Level 2 cause and all causes for 2015, by geography. Results are shown for deaths, YLLs, YLDs and DALYs.............. 3233

Appendix Table 9. Global age-standardised Summary Exposure Values (SEVs) for all risk factors and median percentage change, by sex, 1990, 2005, and 2015, by geography......................... 4363

Appendix Table 10. Citations for Table 1. Summary of epidemiologic evidence supporting causality between risk-outcome pairs included in GBD 2015................................................................. 4815
Preamble

This appendix provides methodological detail, supplemental figures and more detailed results for risk factors. The appendix is organized into broad sections following the structure of the main paper. This study complies with the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) recommendations, and this appendix is more comprehensive and encyclopedic than previous Global Burden of Disease appendices. It includes detailed tables, figures, cause modeling write-ups and flowcharts, and information on data in an effort to maximize transparency in our estimation processes and provide a comprehensive description of analytical steps. Components of this document are the same as described in the appendix to our GBD 2013 risk factors paper; substantial components of this appendix are new text. We intend this to be a living document, to be updated with each annual iteration of the Global Burden of Disease.
Section 1. GBD overview

Geographic units of the analysis
In the GBD framework, geographies have been arranged as a set of hierarchical categories: seven super-regions; 21 regions nested within the seven super-regions; and 195 countries and territories nested in the 21 regions. High-quality vital registration data made it possible to expand the geographies considered in the comparative risk assessment of GBD 2015 to include American Samoa, Bermuda, Greenland, Guam, Northern Mariana Islands, Puerto Rico, and the US Virgin Islands. These territories were not previously included in the national totals of the United States (US), United Kingdom (UK), and Denmark, and instead were included only in regional totals in GBD 2013. Additionally, GBD collaborator interest and availability of data resulted in an expansion of countries for which we disaggregate our estimates at the subnational level, including 26 states and one district for Brazil; 34 provinces and municipalities for China; 31 states and union territory groupings for India that include 62 rural and urban units; 47 prefectures for Japan; 47 counties for Kenya; 32 federal entities for Mexico; 13 provinces for Saudi Arabia; nine provinces for South Africa; two regions for Sweden; 13 regions for the UK (Northern Ireland, Scotland, Wales, England, and nine subregions of England); and 51 states and districts for the US. At the first level of subnational division, 256 geographic units are now included in GBD 2015. For this paper we present results for the 195 national and territory-level geographies.

GBD risk factor hierarchy
In this analysis, we focus on three groups of risk factors: behavioural, environmental and occupational, and metabolic. The GBD 2015 risk factors hierarchy and levels are summarized in Appendix Table 3.

Time periods of the analysis
We produced a complete set of age-, sex-, cause-, and location-specific estimates of risk factor exposure and attributable burden for 1990, 1995, 2000, 2005, 2010, and 2015 for included risk factors. Online data visualizations at http://vizhub.healthdata.org/gbd-compare provide access to results for all GBD metrics, including risk factor results, for all years for which estimates were computed from 1990 through 2015.

List of abbreviations
BMI: body-mass index
BMD: bone mineral density
CKD: chronic kidney disease
COD: causes of death
CODEm: cause of death ensemble modeling
COPD: chronic obstructive pulmonary disease
CSA: childhood sexual abuse
CRA: comparative risk assessment
CVD: cardiovascular disease
DALY: disability-adjusted life-year
DRI: data representativeness index
FAO: Food and Agriculture Organization
GATHER: Guidelines for Accurate and Transparent Health Estimates Reporting
GBD: Global Burden of Disease
IER: integrated exposure response
IHD: ischemic heart disease
ILO: International Labour Organization
IPV: intimate partner violence
LDI: lag distributed income per capita
LRI: lower respiratory infection
MDG: Millennium Development Goal
NCD: non-communicable disease
PAF: population attributable fraction
PM2.5: particulate matter <2.5µm in diameter
RCT: randomised controlled trial
RMSE: root mean square error
SBP: systolic blood pressure
SD: standard deviation
SDG: sustainable development goal
SDI: Socio-demographic Index
SEER: Surveillance, Epidemiology, and End Results Program
SEV: summary exposure value
SIR: smoking impact ratio
SSB: sugar-sweetened beverages
ST-GPR: spatiotemporal Gaussian process regression
TB: tuberculosis
UI: uncertainty interval
WHO: World Health Organization
YLD: years lived with disability
YLL: years of life lost
Section 2. Risk factor estimation overview

Overview

The CRA conceptual framework was developed by Murray and Lopez, who established a causal web of hierarchically organised risks or causes that contribute to health outcomes (Figure 1), which allows for quantification of risks or causes at any level in the framework. In GBD 2015, as in previous iterations of the GBD study, we evaluated a set of behavioural, environmental and occupational, and metabolic risks, where risk-outcome pairs were included based on evidence rules (see appendix p 7). These risks were organised in four hierarchical levels, where level 1 represents the overarching categories (behavioural, environmental and occupational, and metabolic) nested within level 1 risks; level 2 contains both single risks and risk clusters (such as child and maternal malnutrition); level 3 contains the disaggregated single risks from within level 2 risk clusters; and level 4 details risks with the most granular disaggregation, such as for specific occupational carcinogens, the subcomponents of childhood undernutrition (stunting, wasting, underweight), and suboptimal breastfeeding (discontinued and non-exclusive breastfeeding).

At each level of risk, we evaluated whether risk combinations were additive, multiplicative, or shared common pathways for intervention. This approach allows the quantification of the proportion of risk-attributable burden shared with another risk or combination of risks and the measurement of potential overlaps between behavioural, environmental and occupational, and metabolic risks. To date, we have not quantified in the GBD the contribution of other classes of risk factors illustrated in Figure 1; we provide through an analysis of the relationship between risk exposures and development measured using the socio-demographic index (see below for details) some insights into the potential magnitude of distal social, cultural and economic factors.

Two types of risk assessments are possible within the CRA framework: attributable burden and avoidable burden. Attributable burden is the reduction in current disease burden that would have been possible if past population exposure had shifted to an alternative or counterfactual distribution of risk exposure. Avoidable burden is the potential reduction in future disease burden that could be achieved by changing the current distribution of exposure to a counterfactual distribution of exposure. Murray and Lopez identified four types of counterfactual exposure distributions: (1) theoretical minimum risk; (2) plausible minimum risk; (3) feasible minimum risk; and (4) cost-effective minimum risk. The theoretical minimum risk level (TMREL) is the level of risk exposure that minimises risk at the population level, or the level of risk that captures the maximum attributable burden. Other possible forms of risk quantification include plausible minimum risk – which reflects the distribution of risk that is conceivably possible and would minimise population-level risk if achieved – while feasible minimum risk describes the lowest risk distribution that has been attained within a population, and the cost-effective minimum risk is the lowest risk distribution for a population that can be attained in a cost-effective manner. Because no robust set of forecasts for all components of the GBD is available, in this study we focus on quantifying attributable burden using the theoretical minimum risk counterfactual distribution. Figure 2 shows the eight possible types of risk quantification within the CRA framework, with the hatched box representing the type of CRA currently undertaken by the GBD study. As per the definition of avoidable burden, risk reversibility would be incorporated into this type of assessment, as it would involve reducing risk to the counterfactual for the index year, given a history of past risk exposure. Given the focus in this study on attributable burden, risk reversibility is not a criteria used in estimation here.
In general, this analysis follows the CRA methods used in GBD 2013. The methods described here provide a high-level overview of the analytical logic with a focus on areas of notable change from the methods employed in GBD 2013. Key methodological refinements include improved spatial calibration of satellite measures of atmospheric particulate matter <2.5μm in diameter (PM2.5) to ground measurements; an updated integrated exposure response (IER) curve for all outcomes of PM2.5; the development of age-specific relative risks for diet risks based on high systolic blood pressure and cholesterol age curves; a lower TMREL for total cholesterol and for high body mass index (BMI); the incorporation of new data to improve estimation of tourism consumption for alcohol; improvements in exposure data standardization such as age-splitting and severity-splitting for several risks; and the selection of the maximum level of relative risk from dose-response studies for diet and metabolic risks. Here we aim to provide sufficient detail on these methodological improvements to understand the overall structure of the estimation process – greater detail of inputs, analytical processes and outputs, and methods specific to each risk-outcome pairing are now maintained as a single source available as an appendix. This study complies with the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) recommendations proposed by the World Health Organization (WHO) and others, which include recommendations on documentation of data sources, estimation methods, and statistical analysis (Table 1).

### Step 1. Effect size estimation

#### 1a. Collate relative risk data

Criteria for inclusion of risk-outcome pairs

In this study, as in GBD 2013, we have included risk-outcome pairs that we have assessed as meeting the World Cancer Research Fund (WCRF) grades of convincing or probable evidence. In this framework, convincing evidence consists of biologically plausible associations between exposure and disease established from multiple epidemiological studies in different populations. Evidentiary studies must be substantial, include prospective observational studies, and where relevant, randomised controlled trials (RCTs) of sufficient size, duration, and quality, and showing consistent effects. Probable evidence is similarly based on epidemiological studies with consistent associations between exposure and disease, but for which shortcomings in the evidence exist, such as insufficient trials (or prospective observational studies) available.

The World Cancer Research Fund grading system

**Convincing evidence**

Evidence based on epidemiological studies showing consistent associations between exposure and disease, with little or no evidence to the contrary. The available evidence is based on a substantial number of studies including prospective observational studies and where relevant, randomized controlled trials of sufficient size, duration, and quality showing consistent effects. The association should be biologically plausible.

**Probable evidence**

Evidence based on epidemiological studies showing fairly consistent associations between exposure and disease, but for which there are perceived shortcomings in the available evidence or some evidence to the contrary, which precludes a more definite judgment. Shortcomings in the evidence may be any of...
the following: insufficient duration of trials (or studies); insufficient trials (or studies) available; inadequate sample sizes; or incomplete follow-up. Laboratory evidence is usually supportive. The association should be biologically plausible.

Possible evidence
Evidence based mainly on findings from case-control and cross-sectional studies. Insufficient randomized controlled trials, observational studies, or non-randomized controlled trials are available. Evidence based on non-epidemiological studies, such as clinical and laboratory investigations, is supportive. More trials are needed to support the tentative associations, which should be biologically plausible.

Insufficient evidence
Evidence based on findings of a few studies which are suggestive, but insufficient to establish an association between exposure and disease. Little or no evidence is available from randomized controlled trials. More well-designed research is needed to support the tentative association.

2a. Determine relative risks
Effect size estimation
The relative risk by level of exposure, or by cause, for mortality or morbidity can be found in published and unpublished primary studies or in secondary studies that summarize relative risks. In Step 1a of the analytical process (Figure 3 in manuscript), we collated information from randomized controlled trials, cohort, pooled cohort, and case control studies, and in Step 1b, used these data to determine the relative risk for the risk-outcome pairs included in GBD 2015. For most risks, data from pooled cohorts, or meta-analyses of cohorts, were used; in the case of the risk of cataracts from household air pollution cohort data were not available, and instead we used case control data. We estimated relative risks of mortality and morbidity for 59 risk factors for which we determined attributable burden using relative risk and exposure. We incorporated relative risks from studies that controlled for confounding but not for factors along the causal pathway between exposure and outcome. For risk-outcome pairs with evidence available for only one of mortality or morbidity, we generally assumed that the estimated relative risks applied equally to both. Given evidence of statistically different relative risks for mortality and morbidity, we incorporated different relative risks for each. We did not find that relative risks were consistently higher or lower for mortality compared with morbidity. Details and citation information for the data sources used for relative risks are provided in searchable form through a new web-tool (http://ghdx.healthdata.org/).

Available data sources for determining relative risks varied across risks. Relative risks for metabolic risks were established from pooled cohorts that were similar across definitions, study design, and control of major confounding factors. Meta-analyses were used in determining the relative risk for zinc deficiency (zinc supplementation trials) and diarrhoea, and meta-analyses of cohort studies were used for the relative risk for cancers from BMI. For both asthmagens and child sexual abuse, one cohort study with strong controls for confounding or biases was available and was used (see data source tool for citation details, http://ghdx.healthdata.org/). We re-estimated the integrated exposure-response curves for six outcomes (lower respiratory infections, lung cancer, ischaemic heart disease, ischaemic stroke, hemorrhagic stroke, and chronic obstructive pulmonary disease [COPD]) from the risk of exposure to
ambient particulate matter pollution (PM2.5). For four outcomes (breast cancer, diabetes, ischaemic heart disease, ischaemic stroke) paired with the risk of low physical activity, we updated relative risk using recently published studies. Relative risks by age and sex for each risk factor and outcome pair are provided in Appendix Table 6.

For the following seven risk factors and a portion of their related outcomes, evidence of the direct relationship between a risk factor and a disease outcome was sparse or lacking and instead we interpolated relative risks from analogous or related outcomes. Suitable studies for the risk from unsafe water, unsafe sanitation, and unsafe handwashing for enteric diseases (eg, salmonella pathogens, typhoid, and paratyphoid fevers) were not located; because of the similar fecal-oral pathway of transmission, we substituted the effect size for diarrhoea. For certain other risks a single effect size was applied to groups of related outcomes where specific relative risks were unavailable. For the risk of atrial fibrillation and peripheral vascular disease associated with high systolic blood pressure we substituted the relative risk from “other cardiovascular diseases” (other CVD), and for the risk for endocarditis, cardiomyopathy, and myocarditis paired with high BMI, we used the relative risk of inflammatory heart disease. For BMI and colon and rectum cancers, we combined the relative risks proportional to the incidence of each cancer in the Surveillance, Epidemiology, and End Results Program (SEER) cancer registry data into a single relative risk for colorectal cancer.

For all outcomes related to unsafe sex, the relative risk and exposure framework was not used to estimate attributable burden. For unsafe sex and HIV, we used a direct attribution approach to address the lack of data on unsafe sexual practices in most populations. The proportion of HIV attributable to unsafe sex was modelled directly using DisMod-MR 2.1 from data on the fraction of cases identified as being through sexual transmission, intravenous drug use, or blood transfusion.

For risks estimated from a continuous exposure distribution where the effect size is reported by categories in pooled or meta-analysis studies, we converted those categories to relative risk per unit increase in exposure. This implies a linear increase in the log of the relative risk and exposure; various studies have suggested this is a reasonable approximation of the dose-response curve for many risks. An example of this is high systolic blood pressure, where data from the Prospective Cohort Study (PSC) and the Asia-Pacific Cohort Studies Collaboration (APCSC) were well described by a linear increase in the logarithm of the relative risk by a 10-unit increase in high systolic blood pressure. This approximately log-linear relationship suggests that the proportional difference in the age-specific risk of stroke death associated with a given absolute difference in exposure is about the same at all levels of risk. Many meta-analyses convert relative risks to per unit increase for convenience, particularly when studies choose different categories that could not otherwise be compared. The log-linear approximation appears plausible’ even where there is limited consensus on the appropriate TMREL. Where there were insufficient samples in the primary studies at high levels of exposure to inform the shape of the tail of the distribution, we applied a cap to the maximum relative risk using the midpoint of the last category for which a relative risk was reported.
Step 2. Exposure estimation
2a. Collate exposure data

**Systematic reviews**

GBD 2010 collaborators undertook initial systematic reviews for the majority of risk factors; for GBD2013, updates to these reviews were conducted for all risk factors at IHME using data available through January of 2014. For the GBD 2015 study no data or studies were extracted after January 2016. Household surveys including the Demographic and Health Surveys, Multiple Indicator Cluster Surveys, Living Standards Measurement Surveys, Reproductive Health Surveys, and various national health surveys included in the Global Health Data Exchange (ghdx.healthdata.org) were systematically screened for data relevant to sequelae. For some risk factors, only a small fraction of the existing data appear in the published literature and other sources predominate such as survey data and satellite data. The new source tool in GHDx offers a comprehensive view of data sources used in GBD 2015.

**Search terms**

Search terms for updates of systematic reviews for GBD 2015 are shown by risk in Section 3 of this appendix.

**Survey data preparation**

For GBD 2015, survey data constitutes a substantial part of the underlying data used in the estimation process. During extraction, we concentrate on demographic variables (such as location, gender, age), survey design variables (such as sampling strategy and sampling weights), and the variables used to define the population estimate (such as prevalence or a proportion) and a measure of uncertainty (standard error, confidence interval or sample size and number of cases).

2b. Adjust exposure data

A number of adjustments were applied to extracted exposure sources in order to make the data more consistent and suitable for modelling. Commonly applied adjustments included age-sex splitting, adding study-level covariates, and bias correction. Age-sex splitting was applied to literature data reported by age or sex but not by age and sex assuring that the total number of cases remained as reported. If a source did not report sample size by age or sex we applied the age-sex distribution of the population for the same location and year to the reported total sample size. We relied on the metaregression component of DisMod-MR 2.1 for most of the bias correction of data for variations in study attributes such as case definitions and measurement method. DisMod-MR 2.1 calculates a single adjustment that is applied regardless of age, sex, or location. If enough data were available to differentiate these adjustments by age, sex, or location, or if detailed survey data were available to make more precise adjustments between different thresholds on a biochemical measure, we applied bias corrections to the data before entry into DisMod-MR 2.1.

2c. Estimate exposure

**Mean exposure estimation**

In Step 2a of the estimation process, we used systematic literature reviews to identify risk factor exposure studies published or identified since GBD 2013 and combined these with existing data from household and health examination surveys, census, morbidity, or satellite imagery and ground sensor data (used for PM2.5 estimation). Certain risks, such as diet and alcohol consumption, also incorporated
administrative record systems. Sources of data used in estimating risk factor exposure can be accessed through the data source tool at http://ghdx.healthdata.org/.

A geographic and temporal data representativeness index (DRI) for risk factor exposure estimation was determined for each risk factor as the fraction of countries for which we have identified any data for the risk factor. The DRI is a minimalist measure which does not take into account the quality of the available data, only whether any data for an interval are available. For aggregate causes, the DRI score reflects the availability of any data from any component cause. Table 3 provides the DRI for each risk factor in the GBD hierarchy for three time periods: prior to 2005, 2005 to 2015, and the total for all years. Overall, DRI ranged from a low of 16.2% for diet low in whole grains – indicating the lack of available data for this risk factor from most geographies included in the GBD – to 100% for each of ambient ozone pollution and ambient particulate matter pollution. The DRI for PM2.5 is 100% because data are available for all countries and all years. Once data were collected and compiled, step 2b of the analytical flowchart describes the adjustments applied, where necessary, to correct for bias. Examples of these adjustments include: use of urban studies for lead; crosswalks between different measurements, methods, and definitions, such as for self-report of obesity and glycated hemoglobin (HbA1C) for diabetes; and age-sex splitting of data, such as for fasting plasma glucose, cholesterol, and systolic blood pressure that may be reported from broad age-groups.

For the GBD we developed two modelling approaches, a Bayesian meta-regression model (DisMod-MR 2.1) and a spatiotemporal Gaussian process regression model (ST-GPR), to pool data from different sources, control and adjust for bias in data, and incorporate other types of information such as country-level covariates. DisMod-MR 2.1 and ST-GPR are mixed effect models that borrow information across age, time, and geographies to synthesise multiple sources of data into unified estimates of levels and trends. DisMod-MR 2.1 is an improved version of the method used in GBD 2013 (DisMod-MR 2.0). Key updates from the previous version include improvements in how country covariates differentiate estimates with sparse data, and consolidation of the code into a single language, among other computational efficiencies. A detailed description of the likelihood used for estimation, and a full description of improvements made for DisMod-MR 2.1, are detailed by Vos and colleagues with additional detail in the appendix to that paper. The ST-GPR model has three main hyper-parameters that control for smoothing across time, age, and geography. Values for these hyper-parameters were selected based on cross-validation. Cross-validation tests were conducted for different combinations of the hyper-parameters. In each test, 30% of the data were held out and the performance of each combination of hyper-parameters evaluated on the held out data. For each hyper-parameter combination, 25 cross-validation tests were conducted. The performance of each model in predicting the withheld 30% of the data was evaluated using a combined measure based on root mean square error (RMSE) and uncertainty interval coverage.

The main difference between these methods is their power to include unstructured types of data by sex and age group and in their degree of flexibility. Step 2c in Figure 3 outlines the use of DisMod-MR 2.1 for 23 risk factors where data were available by different age intervals or mixed sex groups; DisMod-MR 2.1 is the preferred tool in these cases because of its ability to integrate over age and adjust for different exposure definitions in the data; however, the use of Bayesian Markov Chain Monte Carlo (MCMC) simulations with large volumes of data renders the analysis computationally intensive and reduces the number of iterations that are possible. If large volumes of standard age-group data are available – as is generally the case for metabolic risks – using ST-GPR becomes the preferred approach.
In some cases, we adapted our methods of modelling exposure to risks where necessary to account for complexities in the risk-outcome relationship or the need for particular handling of data. For dietary risks, we first used ST-GPR to model the national availability of foods and nutrients in all countries based on Food and Agriculture (FAO) data from the United Nations. Then, we used DisMod-MR 2.1 to model the intake of each food group and nutrient and to conduct crosswalks between different methods of dietary assessment (food frequency questionnaire, 24-hour diet recall, and diet record), various definitions of food groups or nutrients, and different levels of dietary assessment (national availability, household availability, individual level intake); additional details on crosswalk methodology is supplied for individual risks in the Appendix, Section 3.

For the GBD 2015 study, our estimates for exposure to ambient air pollution incorporated an updated database of 6,003 ground measurements based on the recently released WHO Air Pollution in Cities database\(^9\) and updated satellite-based estimates.\(^10\) These estimates combine aerosol optical depth retrievals from multiple satellites with the chemical transport model, GEOS Chem.\(^11\) For GBD 2010 and GBD 2013 a single function was used to calibrate available ground measurements to the mean of gridded satellite-based and chemical transport model values. This use of a single, global calibration led to underestimation of ground measurements in some locations\(^12\) and we therefore applied a Bayesian hierarchical modelling approach, which allowed the calibration to vary spatially and to enable the inclusion of information on land use and other factors related to air quality. The within-sample model fit and out-of-sample assessment of predictive ability were used to inform modelling. Predictive validity was assessed using 25 sets of training data, where holdouts were determined by randomly selecting 20% of sites based on sampling probabilities for super-regions and tabulated PM2.5 categories – returning a validation set with the same distribution of PM2.5 exposure and super-regions as the training dataset. Improvements were seen for countries and regions with limited ground monitoring. This process resulted in an improvement in both within-sample fit; with an increase in R\(^2\) from 0.64 (reported in GBD 2013) to 0.91, and out-of-sample predictive ability; with a population-weighted RMSE of 12.1 µg/m\(^3\) compared to 23.1 µg/m\(^3\) when using the GBD 2013 model.

To evaluate exposure to smoking for cancers and COPD, we calculated a smoking impact ratio (SIR)\(^13\) using the lung cancer mortality rate estimated for every population compared to the mortality rate in nonsmokers and never-smokers from the few cohorts available; for example, the Cancer Prevention Study II.\(^14\) Estimating exposure directly from smoking prevalence would be preferable but at present is hindered by the variation in tobacco content in cigarettes and other products, filtering, cigarettes per smoker, and other factors which contribute to the effect of tobacco on cancers. The SIR based on observed lung cancer death rates is meant to capture the lifetime cumulative effect of smoking.

In modelling exposure to alcohol consumption, we extracted exposure data from general population surveys reported in both unpublished and published literature; however, these surveys tend to underestimate the amount of alcohol consumption due to self-report.\(^15,16\) To correct for the underreporting of alcohol consumption in surveys, we estimated total alcohol consumed in every country using data of alcohol sales and FAO data of available alcohol for drinking and then adjusted for sales to tourists visiting each country and the estimated volume of illicit production from survey data.\(^17\) Survey-based estimates of consumption by age and sex have been scaled up to match our estimates of total consumption. A complete list of risks and the analytical method used is reported in Appendix Table 3. Additional details for adjustments or adaptations to particular risk models are located in Appendix section 3.
Exposure distributions

In order to select an appropriate distribution for risk factors measured on a continuous scale we used mean and standard deviation (SD) for our models, because these statistics are available in nearly all published studies. We found strong predictive validity (smaller RMSE) between the mean and SD using out-of-sample cross validation compared with the alternative of modelling the coefficient of variation. A correlation coefficient of at least 96% (R²) was found between the SD and mean of dietary and metabolic risks from survey populations.

In analyses conducted for GBD 2013, we tested normal, beta, lognormal, and gamma distributions for their fit to metabolic and diet risks using individual record datasets (such as the National Health and Nutrition Examination Study [NHANES], the Cebu Longitudinal Health and Nutrition Survey [CLHNS], or the National Income Dynamics Study [NIDS]; see data source tool for details), and found that the lognormal distribution fit the available data best for all but three risk factors: iron deficiency and low bone mineral density, high BMI, and high systolic blood pressure. For iron deficiency and low bone mineral density, the best fit was provided by the normal distribution. For high BMI, GBD authors Ng and colleagues demonstrated that the best fit was provided by a beta distribution fit to the mean prevalence of overweight and prevalence of obesity, constrained such that skewness could not be negative. For high systolic blood pressure, relative risks were corrected for regression dilution bias; exposure SD was corrected for a measure of inter-temporal variance in blood pressure observed in cohort studies to ensure the estimates reflected “usual” systolic blood pressure. We did not use a relative risk per unit increase in exposure where the relative risk substantially deviated from a log-linear approximation. For example, for the integrated exposure response curve (PM2.5 exposure and relative risk of outcomes) we fit a nonlinear curve and estimated the relative risk for every level of PM2.5.

DisMod-MR 2.1 Estimation

An estimation method used for modeling the exposure to many risk factors is the Bayesian meta-regression method DisMod-MR 2.1.

DisMod-MR 2.1 description

Until GBD 2010, non-fatal estimates were based on a single data source on prevalence, incidence, remission or a mortality risk selected by the researcher as most relevant to a particular geography and time. For GBD 2010, we set a more ambitious goal: to evaluate all available information on a disease that passes a minimum quality standard. That required a different analytical tool that would be able to pool disparate information presented in varying age groupings and from data sources using different methods. The DisMod-MR 1.0 tool used in GBD 2010 evaluated and pooled all available data, adjusted data for systematic bias associated with methods that varied from the reference and produced estimates by world regions with uncertainty intervals using Bayesian statistical methods. For GBD 2013, the improved DisMod-MR 2.0 had increased computational speed allowing computations that were consistent between all disease parameters at the country rather than region level. The hundred-fold increase in speed of DisMod-MR 2.0 was partly due to a more efficient re-write of the code in C++ but also by changing to a model specification using log rates rather than a negative binomial model used in DisMod-MR 1.0. In cross-validation tests, the log rates specification worked as well or better than the negative binomial specification. For GBD2015, the computational engine (DisMod-MR 2.1) remained substantively unchanged but we re-wrote the ‘wrapper’ code that organizes the flow of data and settings at each level.
of the analytical cascade. The sequence of estimation occurs at 5 levels: global, super-region, region, country and where applicable subnational geographical unit. The super-region priors are generated at the global level with mixed-effects, non-linear regression using all available data; the super-region fit, in turn, informs the region fit, and so on down the cascade. The wrapper gives analysts the choice to branch the cascade in terms of time and sex at different levels depending on data density. The default used in most models is to branch by sex after the global fit but to retain all years of data until the lowest level in the cascade. For GBD 2015, we generated fits for the years 1990, 1995, 2000, 2005, 2010, and 2015.

In updating the ‘wrapper,’ we consolidated the code base into a single language, Python, to make the code more transparent and efficient and to better deal with subnational estimation. The computational engine is limited to three levels of random effects; we differentiate estimates at the super-region, region and country level. In GBD 2013, the subnational units of China, the UK and Mexico were treated as ‘countries’ such that a random effect was estimated for every geography with contributing data. However, the lack of a hierarchy between country and subnational units meant that the fit to country data contributed as much to the estimation of a subnational unit as the fits for all other countries in the region. We found inconsistency between the country fit and the aggregation of subnational estimates when the country’s epidemiology varied from the average of the region. Adding an additional level of random effects required a prohibitively comprehensive rewrite of the underlying DisMod-MR engine. Instead, we added a fifth layer to the cascade, with subnational estimation informed by the country fit and country covariates, plus an adjustment based on the average of the residuals between the subnational unit’s available data and its prior. This mimicked the impact of a random effect on estimates between sub nationals.

For GBD 2015 we improved how country covariates differentiate non-fatal estimates for diseases with sparse data. The coefficients for country covariates are re-estimated at each level of the cascade. For a given geography, country coefficients are calculated using both data and prior information available for that geography. In the absence of data, the coefficient of its parent geography is used, in order to utilize the predictive power of our covariates in data sparse situations.

**DisMod-MR 2.1 likelihood estimation**

Analysts have the choice of using a Gaussian, log-Gaussian, Laplace or Log-Laplace likelihood function in DisMod-MR 2.0. The default log-Gaussian equation for the data likelihood is:

\[
-\log \left( p(y_j|\Phi) \right) = \log(\sqrt{2\pi}) + \log(\delta_j + s_j) + \frac{1}{2} \left( \frac{\log(a_j + \eta_j) - \log(m_j + \eta_j)}{\delta_j + s_j} \right)^2
\]

where, \( y_j \) is a ‘measurement value’ (i.e. data point); \( \Phi \) denotes all model random variables; \( \eta_j \) is the offset value, eta, for a particular ‘integrand’ (prevalence, incidence, remission, excess mortality rate, with-condition mortality rate, cause-specific mortality rate, relative risk or standardized mortality ratio) and \( a_j \) is the adjusted measurement for data point \( j \), defined by:

\[
a_j = e^{(-u_j - c_j)}y_j
\]
where \( u_j \) is the total ‘area effect’ (i.e. the sum of the random effects at three levels of the cascade: super-region, region and country) and \( c_j \) is the total covariate effect (i.e. the mean combined fixed effects for sex, study level and country level covariates), defined by:

\[
c_j = \sum_{k=0}^{K[I(j)]-1} \beta_{I(j),k} \hat{x}_{k,j}
\]

with standard deviation

\[
s_j = \sum_{l=0}^{L[I(j)]-1} \zeta_{I(j),l} \hat{z}_{k,j}
\]

where \( k \) denotes the mean value of each data point in relation to a covariate (also called x-covariate); \( I(j) \) denotes a data point for a particular integrand, \( j \); \( \beta_{I(j),k} \) is the multiplier of the \( k^{\text{th}} \) x-covariate for the \( i^{\text{th}} \) integrand; \( \hat{x}_{k,j} \) is the covariate value corresponding to the data point \( j \) for covariate \( k \); \( l \) denotes the standard deviation of each data point in relation to a covariate (also called z-covariate); \( \zeta_{I(j),k} \) is the multiplier of the \( l^{\text{th}} \) z-covariate for the \( i^{\text{th}} \) integrand; and \( \delta_j \) is the standard deviation for adjusted measurement \( j \), defined by:

\[
\delta_j = \log[y_j + e^{(-u_j-c_j)\eta_j + c_j}] - \log[y_j + e^{(-u_j-c_j)\eta_j}]
\]

Where \( m_j \) denotes the model for the \( j^{\text{th}} \) measurement, not counting effects or measurement noise and defined by:

\[
m_j = \frac{1}{B(j)-A(j)} \int_{A(j)}^{B(j)} I_j(a) \, da
\]

where \( A(j) \) is the lower bound of the age range for a data point; \( B(j) \) is the upper bound of the age range for a data point; and \( I_j \) denotes the function of age corresponding to the integrand for data point \( j \).

**Modelling Dietary Risks in DisMod-MR 2.1**

We used DisMod-MR 2.1 to estimate the intake of each dietary component by age, sex, and year in each country and subnational unit. For each dietary factor, we included in our models study level covariates that provided information about the method of dietary assessment (i.e., 24-hour diet recall, food frequency questionnaire, household budget surveys, FAO Food Balance Sheets), definition of the dietary factor (whether it is consistent with the definition of GBD or not), and representativeness of survey (whether it is representative of the geographical unit or not). We considered data from representative 24-hour diet recall as optimal and adjusted all other data sources accordingly. For some dietary risks, we used relevant country level covariates to help improve our estimates where we had missing data. For example, we used national availability of red meat and pig meat as covariates for processed meat; national availability of hydrogenated oil as a covariate for trans fatty acids; and national availability of sugar for sugar-sweetened beverages.
Spaciotemporal Gaussian process regression

Spaciotemporal Gaussian process regression (ST-GPR) has been used for risk factors where the data density is sufficient to estimate a very flexible time trend. The approach is a stochastic modeling technique that is designed to detect signals amidst noisy data. It also serves as a powerful tool for interpolating non-linear trends.\(^{22,23}\) Unlike classical linear models that assume that the trend underlying data follows a definitive functional form, GPR assumes that the specific trend of interest follows a Gaussian Process, which is defined by a mean function \(m(\cdot)\) and a covariance function \(Cov(\cdot)\). For example, let \(p_{c,a,s,t}\) be the exposure, in normal, log, or logit space, observed in country \(c\), for age group \(a\), and sex \(s\) at time \(t\):

\[
(p_{c,a,s,t}) = g_{c,a,s}(t) + \epsilon_{c,a,s,t}
\]

where

\[
\epsilon_{c,a,s,t} \sim \text{Normal}(0, \sigma_{p}^2),
\]

\[
g_{c,a,s}(t) \sim GP\left(m_{c,a,s}(t), Cov\left(g_{c,a,s}(t)\right)\right).
\]

The derivation of the mean and covariance functions, \(m_{c,a,s}(t)\) and \(Cov\left(g_{c,a,s}(t)\right)\), along with a more detailed description of the error variance \((\sigma_{p}^2)\), is described below.

**Estimating mean functions**

We estimated mean functions using a two-step approach. To be more specific, \(m_{c,a,s}(t)\) can be expressed, depending on the exposure transformation, as:

\[
\log(p_{c,a,s}(t)) = X_{c,a,s}\beta + h(r_{c,a,s,t})
\]

\[
\logit(p_{c,a,s}(t)) = X_{c,a,s}\beta + h(r_{c,a,s,t})
\]
\[ p_{c,a,s}(t) = X_{c,a,s}\beta + h(r_{c,a,s,t}) \]

where \( X\beta \) is the summation of the components of a hierarchical mixed-effects linear regression, including the intercept and the product of covariates with their corresponding fixed effect coefficients. For a majority of models, predictions were not made using the random effects component of the linear model. The second part of the equation, \( h(r_{c,a,s,t}) \), is a smoothing function for the residuals, \( r_{c,a,s,t} \), derived from the linear model. Descriptions of exposure transformations and which covariates were used in linear models can be found in Section 3. Risk-specific estimation.

While the linear component captures the general trend in exposures over time, much of the data variability may still not be adequately accounted for. To address this, we fit a locally weighted polynomial regression (LOESS) function \( h(r_{c,a,s,t}) \) to systematically estimate this residual variability by borrowing strength across time, age, and space patterns (the spatio-temporal component of ST-GPR). The time adjustment parameter, defined by \( \lambda \), aims to borrow strength from neighboring time points (i.e. the exposure in this year is highly correlated with exposure in the previous year but less so further back in time). The age adjustment parameter, defined by \( \omega \), borrows strength from data in neighboring age groups. The space adjustment parameter, defined by \( \xi \), aims to borrow strength across the hierarchy of geographical locations.

Let \( w_{c,a,s,t} \) be the final weight assigned to observation \( r_{c,a,s,t} \) with reference to a focal observation \( r_{c_0,a_0,s_0,t_0} \). We first generated a preliminary weight \( w'_{c,a,s,t} \) for smoothing over time, which was based on the scaled distance along the time dimension of the two observations:

\[ w'_{c,a,s,t} = \left(1 - \left(\frac{|t - t_0|}{1 + \max|t - t_0|}\right)^\lambda \right)^3 \]

Next, we calculated the weight \( w''_{c,a,s,t} \) to smooth over age, which is based on a distance along the age dimension of two observations. For a point between the age \( a \) of the observation \( r_{c,a,s,t} \) and a focal observation \( r_{c_0,a_0,s_0,t_0} \), the weight is defined as follows:

\[ w''_{c,a,s,t} = \frac{1}{e^{\omega|a-a_0|}} \]

Finally, these combined weights were multiplied and further adjusted to account for geographic patterns.

Specifically, we defined a geospatial relationship by categorizing data based on the GBD location hierarchy. We adapted the weighting strategy used in previous studies estimating time series of global indicators to be more flexible with respect to estimating subnational locations and to borrow strength from all levels. A vector of spatial weights corresponding to each level of the location hierarchy was derived as \([\xi, \xi \times (1 - \xi)^{n_1-1}, \ldots, \xi \times (1 - \xi)^{n_i-1}, (1 - \xi)^{n_i}]\), where the vector is expanded to include the number, \( n_i \), levels in the location hierarchy between the location being estimated and global, which receives a pre-rescaling weight of \((1 - \xi)^{n_i}\). For example, estimating a country would use the following weighting scheme:

- Country data: \( \xi \)
- Regional data not from the country being estimated: \( \xi \times (1 - \xi) \)
Data from other regions in the same super region: $\xi \cdot (1 - \xi)^2$

Global data from other super regions: $(1 - \xi)^3$

A full derivation of weights for each category follow, assuming the location being estimated was a country, follows:

1) If the observation $r_{c,t}$ belongs to the same country $c_0$ of the focal observation $r_{c_0,t_0}$:

$$w_{c,a,s,t} = \frac{\xi (w'_{c,a,s,t} w''_{c,a,s,t})}{\sum_{c=c_0} (w'_{c,a,s,t} w''_{c,a,s,t})} \quad \forall c = c_0$$

2) If the observation $r_{c,t}$ belongs to a different country than the focal observation $r_{c_0,t_0}$, but both belong to the same region $R$:

$$w_{c,a,s,t} = \frac{\xi \cdot (1 - \xi) \cdot (w'_{c,a,s,t} w''_{c,a,s,t})}{\sum_{c \neq c_0} (w'_{c,a,s,t} w''_{c,a,s,t})} \quad \forall c \neq c_0 \cap R[c] = R[c_0]$$

3) If the observation $r_{c,t}$ belongs to the same super region $SR$ but to a both different country $c_0$ and region $R[c_0]$ than the focal observation $r_{c_0,t_0}$:

$$w_{c,a,s,t} = \frac{\xi \cdot (1 - \xi)^2 \cdot (w'_{c,a,s,t} w''_{c,a,s,t})}{\sum_{c \neq c_0} (w'_{c,a,s,t} w''_{c,a,s,t})} \quad \forall c \neq c_0 \cap R[c] \neq R[c_0] \cap SR[c] = SR[c_0]$$

4) If the observation $r_{c,t}$ is from a different super region than the focal observation $r_{c_0,t_0}$ (ie. all other data currently not receiving a weight):

$$w_{c,a,s,t} = \frac{(1 - \xi)^3 \cdot (w'_{c,a,s,t} w''_{c,a,s,t})}{\sum_{c \neq c_0} (w'_{c,a,s,t} w''_{c,a,s,t})} \quad \forall c \neq c_0 \cap R[c] \neq R[c_0] \cap SR[c] \neq SR[c_0]$$

To allow additional flexibility and specificity in weighting schemes, we allowed for two different $\xi$ to be defined. The higher $\xi$ was applied when at least one age-sex group in the country of estimation had at least five unique data points. The lower $\xi$ was applied when estimating data-scarce countries.

Observations could be downweighted by a factor of 0.1, usually because they were not geographically representative at the unit of estimation. Details of reasons for downweighting can be found in risk-specific modeling summaries. The final weights were then normalized such that the sum of weights across age, time, and geographic hierarchy for a reference group was 1.

**Estimating error variance**

$\sigma_p^2$ represents the error variance in normal or transformed space including sampling variance of the estimates and predication error from any crosswalks performed. First, variance was systematically imputed if the data extraction did not include any measure of uncertainty. When some sample sizes for data were available, missing sample sizes were imputed as the 5th percentile of available sample sizes. Missing variances were then calculated as $\sigma_p^2 = \frac{p(1-p)}{n}$ for proportions and using the global coefficient of variation for continuous exposures. When sample sizes were entirely missing and could not be imputed, the 95th percentile of available variances at the most granular geographic level (ie. first
country, then region, etc.) were used to impute missing variances. For proportions where \( p \times n \) or \((1-p)\times n\) is < 20, variance was replaced using the Wilson Interval Score method.

Next, if the exposure was modeled as a log transformation, the error variance was transformed into log-space using the delta method approximation as follows,

\[
\sigma_p^2 \approx \frac{\sigma_{p, t}^2}{p_{c,a,s,t}^2}
\]

where \( \sigma_p^2 \) represents the error variance in normal space. If the exposure was modeled as a logit transformation, the error variance was transformed into logit-space using the delta method approximation as follows,

\[
\sigma_p^2 \approx \frac{\sigma_{p, t}^2}{(p_{c,a,s,t} \times (1 - p_{c,a,s,t}))^2}
\]

Finally, prior to GPR, an approximation of non-sampling variance was added to the error variance. Calculations of non-sampling variance were performed on normal-space variances, and before GPR variances were again transformed using the delta method approximation, if necessary. Non-sampling variance was calculated as the variance of inverse-variance weighted residuals from ST at a given location level hierarchy. If there were fewer than 5 data points at a given level of the location hierarchy the non-sampling variance was replaced with that of the next highest geography level with more than 5 data points.

**Estimating the covariance function**

The final input into GPR is the covariance function, which defines the shape and distribution of the trends. Here, we have chosen the Matern-Euclidian covariance function, which offers the flexibility to model a wide spectrum of trends with varying degrees of smoothness. The function is defined as follows:

\[
M(t, t') = \sigma^2 2^{1-v} \left( \frac{d(t, t') \sqrt{2v}}{l} \right)^v K_v \left( \frac{d(t, t') \sqrt{2v}}{l} \right)
\]

where \( d(\cdot) \) is a distance function; \( \sigma^2, v, l, \) and \( K_v \) are hyperparameters of the covariance function—specifically \( \sigma^2 \) is the marginal variance, \( v \) is the smoothness parameter that defines the differentiability of the function, \( l \) is the length scale, which roughly defines the distance between which two points become uncorrelated, and \( K_v \) is the Bessel function. Based on previous applications of ST-GPR, we approximated \( \sigma^2 \) by \( MADN(r_{c,t}') \), which is the normalized absolute deviation of the residuals from the smoothing step for each country, region, or super-region depending on the data coverage at a given location hierarchy level. Here, we have used the parameter specifications \( v = 2 \) and \( l = 20 \).

**Prediction using GPR**

Based on the specifications stated above, we integrated over \( g_{c,t}(t_*) \) to predict the full time series of mean SBP for country \( c \), age \( a \), sex \( s \), and the prediction time \( t_* \):

\[
\log(p_{c,a,s}(t_*)) \sim N\left(m_{c,a,s,t}(t_*), \sigma_p^2 l + Cov\left(g_{c,a,s,t}(t_*)\right)\right)
\]
Random draws of 1000 samples were obtained from the distributions above for every country for a given indicator. The final estimated mean for each country was the mean of the draws. In addition, 95% uncertainty intervals were calculated by taking the 2.5 and 97.5 percentile of the sample distribution. The entire modeling process was performed in log space and back-transformed to obtain final estimates in the original scale. The linear modeling process was implemented using the lmer4 package in R, and the ST-GPR analysis was implemented through the PyMC2 package in Python.

Subnational Scaling and Aggregation
To ensure consistency of the estimates between countries and their respective subnational locations, national estimates were either created by population-weighted aggregation or subnational estimates were adjusted by population-weighted scaling to the national estimates, depending on the data coverage of a given country compared to that of its subnational locations. For example, if there was better data coverage at the national level, relative to its corresponding subnational locations, for a given country and risk across age, sex, and time, estimates were raked to the national level. Conversely, if there was better data coverage at the subnational level, estimates for its parent country were created through population-weighted aggregation.

3. Estimate summary exposure values
In prior GBD studies, we did not report comparable exposure metrics for the risk factors included because of the complexity of quantifying polytomous and continuous risks. Because of this challenge, prior GBD studies have largely reported attributable deaths or DALYs in rates or numbers. For dichotomous exposures (tobacco or obesity), we previously published separate analyses of the trends in exposure to these risks. Because of substantial interest in this type of exposure trend analysis, for the present analysis we generated a summary measure of exposure for each risk at step 3 of our analytical process (Figure 3). This summary measure, called the summary exposure value (SEV), is the risk-weighted prevalence of exposure. More formally, it is defined:

\[
SEV = \sum_{i=1}^{n} \frac{Pr_i \cdot RR_i - 1}{RR_{max} - 1}
\]  

(3)

where \(Pr_i\) is prevalence of category \(i\) exposure; \(RR_i\) is relative risk of the category \(i\); and \(RR_{max}\) is the maximum relative risk observed (between categories). This quantity is estimated for each age, sex, location, and year. In the case of dichotomous exposure, SEV is equal to prevalence. For continuous risks:

\[
SEV = \frac{\int_{x=1}^{u} RR(x) \cdot P(x) \, dx - 1}{RR_{max} - 1}
\]  

(4)

where \(P(x)\) is the density of exposure at level \(x\) of exposure; \(RR(x)\) is relative risk of the level \(x\); and \(RR_{max}\) is the highest relative risk that is supported by data and reflects a level where more than 1% of the population are exposed globally.

SEV takes the value zero when there is no excess risk for a population and takes the value 1 when the population is at the highest level of risk; we report SEV on a scale from 0% to 100% to emphasise that it is risk-weighted prevalence. Because risk exposure distributions can include individuals with extremely high levels of exposure that are often inflated by measurement error, and because few cohort studies provide valid relative risks at the highest level of exposure, we computed \(RR_{max}\) as the level for
exposure with the highest relative risk supported by cohort or trial data and for which at least 1% or more of the global population is exposed. For comparison purposes, we have also computed age-standardised SEVs for every risk factor from the most detailed level using the GBD population standard.

4. Theoretical minimum-risk exposure level
In this and all previous GBD studies, the counterfactual level of risk exposure used is the risk exposure that is both theoretically possible and minimizes risk in the exposed population that consequently captures the maximum population attributable burden. For each risk evaluated in GBD 2015, Step 4 of the analytical flowchart describes the use of the best available epidemiological evidence from published and unpublished relative risks by level of exposure and the lowest observed level of exposure from cohorts, used to select a single level of risk exposure that minimises risk from all causes of DALYs combined to establish the TMREL. In principle, the TMREL for a given risk may vary by age, sex, and location if supported by clear evidence. Based on the available evidence, the TMREL itself can be uncertain, which is reflected in the 95% uncertainty intervals (UIs) in Table 3. An estimation of uncertainty was derived by resampling from a uniform distribution of TMRELs where evidence supporting the selection of the TMREL was uncertain (for example, elevated systolic blood pressure or cholesterol).

Following substantive debate over the appropriate selection of TMREL for sodium, the UIs for the TMREL for sodium intake were widened for GBD 2013 but were not adjusted further for the present study. For other dietary risks, we used a two-step approach to determine the TMREL. First, we estimated the level of intake associated with the lowest risk for each outcome based on the published reports from cohorts and RCTs evaluating that risk-outcome pair. Then, we calculated the TMREL as the weighted average of these estimates. The weight was estimated by dividing the number of deaths due to each outcome by the total number of deaths from all the outcomes related to the exposure at the global level. Sufficient evidence has accumulated to justify adjusting the TMREL for bone mineral density (BMD) with age; we used the 99th percentile of age-sex subgroups of the NHANES data to capture the decrease in bone density with age while also including the excess risk of fracture resulting from lower BMD in older age groups. For GBD 2015, we altered the TMREL for total cholesterol in light of new evidence from statin trials at low levels of cholesterol; a recent meta-analysis found that cardiovascular outcomes could be improved even at low levels of LDL-cholesterol, below 1.3 mmol/litre. We used the strong correlation between LDL-cholesterol and total cholesterol to map the proposed LDL-cholesterol TMREL of 0.7-1.3 mmol/litre to a TMREL for total cholesterol of 2.8-3.4 mmol/litre. We also revised the TMREL for PM2.5, previously set as the lowest 5th percentile of observed values in cohorts evaluated for GBD 2010 and GBD 2013; the publication of new cohorts led us to decrease the TMREL for particulate matter air pollution to 2.4-5.9 µg/m³ from the value previously used (5.9-8.7 µg/m³). There is insufficient evidence that risk exists below this new TMREL or that a lower level is achievable or even theoretically possible.

5. Estimate population attributable fractions
Risks are categorised on the basis of how exposure was measured: dichotomous, polytomous, and continuous. High total cholesterol is an example of a risk measured on a continuous scale. The population attributable fraction (PAF), which represents the proportion of risk that would be reduced in
a given year if the exposure to a risk factor in the past were reduced to an ideal exposure scenario, is
defined for a continuous risk factor as: \( G_{PAF} \)

\[
P_{PAF} = \frac{\int_{x=1}^{u} RR_{joasg}(x)P_{jasg}(x)dx - RR_{joasg}(TMREL_{jast})}{\int_{x=1}^{u} RR_{joasg}(x)P_{jasg}(x)dx}
\]

Where \( P_{PAF} \) is the population attributable fraction for cause \( o \) due to risk factor \( j \) for age group \( a \), sex \( s \), geography \( l \), and year \( t \). \( RR_{joasg}(x) \) is the relative risk as a function of exposure level \( x \) for risk factor \( j \) for cause \( o \), age group \( a \), sex \( s \), and geography \( g \) with the lowest level of observed exposure as \( l \) and the highest as \( u \); \( P_{jasg}(x) \) is the distribution of exposure at \( x \) for age group \( a \), sex \( s \), geography \( g \), and year \( t \); \( TMREL_{jast} \) is the TMREL for risk factor \( j \), age group \( a \), and sex \( s \).

The \( P_{PAF} \) for dichotomous and polytomous risk factors for every country is defined as:

\[
P_{PAF} = \frac{\sum_{x=1}^{u} RR_{joast}(x)P_{jasg}(x) - RR_{joasg}(TMREL_{jast})}{\sum_{x=1}^{u} RR_{joasg}(x)P_{jasg}(x)}
\]

Where \( P_{PAF} \) is the population attributable fraction for cause \( o \) due to risk factor \( j \) for age group \( a \), sex \( s \), geography \( g \), and year \( t \). \( RR_{joasg}(x) \) is the relative risk as a function of exposure level \( x \) for risk factor \( j \) for cause \( o \), age group \( a \), sex \( s \), and geography \( g \) on a plausible range of exposure levels from \( l \) to \( u \). \( P_{jasg}(x) \) is the proportion of population in risk group (prevalence), for age group \( a \), sex \( s \), geography \( g \), and year \( t \); \( TMREL_{jast} \) is the TMREL for risk factor \( j \), age group \( a \), and sex \( s \).
6. Mediation

Summary
The portion of the burden of disease that is attributable to various combinations of risk factors or to all risk factors combined has been a topic of broad interest. Assumptions about how one risk factor is mediated through other risk factors are needed in order to estimate the joint risk factor burden for combinations of metabolic risks and behavioural or environmental risks. To accomplish this, in Step 6 of the estimation process, for every two risk factors for an outcome, we used published studies to estimate the fraction of risk that was mediated through the other risk. This resulted in a matrix of parameters containing each possible pairing of risk factors included in the GBD 2015. Using this matrix, we computed the aggregated burden of disease at each level of the GBD 2015 hierarchy and for all risk factors using the following formula:

\[
P_{\text{PAF}_{j}} = 1 - \prod_{j=1}^{J} \left( 1 - P_{\text{PAF}_{j}} \prod_{i=1}^{I} (1 - M_{j}) \right)
\]

where \( J \) is a set of risk factors for the aggregation; \( P_{\text{PAF}_{j}} \) is the PAF for risk \( j \) for age group \( a \), sex \( s \), geography \( g \), and year \( t \); and \( M_{j} \) is the mediation factor for risk \( j \) mediated through \( i \) for cause \( o \).

Additional detail
In GBD 2010 we only aggregated the burden of risk factors for some clusters of risks including access to improved water and sanitation, child and maternal malnutrition, tobacco smoking, alcohol use, dietary risk factors, occupational risk factors, and sexual abuse and violence. We did not aggregate air pollution and metabolic risk factors. In GBD 2013 and GBD 2015, we aggregated all risk factors into three large categories: behavioral, environmental and occupational, and metabolic risks -- as well as aggregating all GBD risk factors into a single attributable fraction for each diseases and eventually for all-causes of burden.

Aggregating risk factors at different levels share three essential challenges:

1. Risk factor coexistence or aggregation: for example, metabolic risk factors often occur together or high-risk behaviors are related such as drug abuse and unsafe sex.
2. Mediation: a risk factor may effect another risk factor that lies in the physiological pathway to a disease outcome. It can be inside a cluster of risk factors such as the effect of obesity through an increase in fasting plasma glucose (FPG) and later cardiovascular disease outcomes, or between clusters of risk factors such as the effect of fiber on cholesterol.
3. The formula to calculate the aggregated PAF.

The aggregation method is conceptually applicable to other aggregations such as socioeconomic factors, education, homelessness and refugee status that are being considered for inclusion in future GBD iterations. In the next section, we explain our approach to deal with these challenges.

There are three patterns of associations between risk factors to take into consideration. The first concerns confounding; risk B affects risk A and outcome C (Pattern 1 in Figure. Patterns of associations between risk factors). In these cases the relative risk (RR) for A should be adjusted for B, for example the fruit RR is adjusted for smoking. If part of the effect of A is through B, a mediator, we do not adjust the
effect of A for B. For example, we do not adjust the RR of body mass index (BMI) for cholesterol as cholesterol lies in the biological pathway between BMI and cardiovascular outcomes (Pattern 2 in Figure. Patterns of associations between risk factors). The third pattern occurs when risks A and B are proxies of a third variable Z and aggregation aims to estimate the total effect of a latent variable Z, on C. An example is childhood undernutrition, which is measured by stunting, wasting, and underweight as proxies.

Figure. Patterns of associations between risk factors

Calculating burden of multiple risk factors
Validation studies have reported congruency between the true risk associated with multiple risk factors affecting the same outcome and a multiplicative aggregation of the population attributable fractions of the individual risk factors (formula below).\(^{32}\)

\[
P_{AF_{1..l}} = 1 - \prod_{i=1}^{n} (1 - P_{AF_{i}})
\]

Where \(P_{AF}\) is the population attributable fraction and \(i\) is each individual risk factor. The same validation studies also found that the overestimation from ignoring the covariance between risk factors is small. This was important to note as there are few data sources from which we can draw information on covariance.
We endeavored to evaluate RRs that were controlled for confounders. However, as we had to rely on the literature for many RRs we did not always have full control over the choice of confounders controlled for in each study.

**Adjusting for mediation**

When aggregating the effects of multiple risk factors, we included a mediation factor if a part of the effect of one risk factor was included in the effect estimated for in the mediator. First we prepared a list of possible mediations especially between metabolic risk factors and other risk factors. We found limited data primarily for these categories. We did not assume any mediation effect between risk factors for cancers except for sugar sweetened beverages and BMI.

Danaei and colleagues assumed that part of the effect of BMI on ischemic heart disease (IHD) is through high systolic blood pressure (SBP), cholesterol and FPG. The proportion of the BMI effect that can be explained by other metabolic risk factors is the amount of mediation. The difference between the crude RR of BMI on IHD with the RR adjusted for SBP, FPG, and cholesterol reflects the amount of BMI effect on IHD that is mediated and already included in SBP, FPG, and cholesterol:

\[
MF = \frac{RR_{crude} - RR_{adjusted}}{RR_{crude} - 1}
\]

We used this approach for estimating mediation factors to adjust PAFs before aggregation.

\[
MF = \frac{R^+_c - R^+_a}{R^+_c - R^-_c}
\]

So:

\[
R^+_a = R^+_c - MF \times (R^+_c - R^-_c)
\]

\[
PAF_c = \frac{p \times (R^+_c - R^-_c)}{p \times R^+_c + (1 - p) \times R^-_c} = \frac{p \times (R^+_c - R^-_c)}{R_T}
\]

If \( R^+_c \): crude risk of outcome in exposed population

\( R^-_c \): crude risk of outcome in non-exposed population

\( R^+_a \): adjusted risk of outcome in exposed population

\( R^-_a \): adjusted risk of outcome in non-exposed population

\( R_T \) is the overall rate of the outcome in the population. Since we are interested in the part which is from BMI but through cholesterol, the total risk in the population will be the same for the adjusted RR, so the unmediated part of the risk factor would be:

\[
PAF_a = \frac{p \times (R^+_a - R^-_a)}{R_T} = \frac{p \times (R^+_c - MF \times (R^+_c - R^-_c) - R^-_c)}{R_T} = \frac{p \times (R^+_c - R^-_c) \times (1 - MF)}{R_T} = PAF_c \times (1 - MF)
\]

So for aggregating the PAF of multiple risk factors, we first calculated the part of the effect of every risk factor that is not mediated and then aggregated these assuming they are independent.

Therefore the aggregated PAF would be:

If MF is mediation factor of R2 through R1:
\[
P_{AF_{1,2}} = 1 - (1 - P_{AF_1}) \cdot \left(1 - P_{AF_2} \cdot (1 - MF_{2/1}) \right)
\]

and a generalization for multiple pathways of R1 through other RFs:

\[
P_{AF_{1,i}} = 1 - \prod_{i=1}^{n} \left(1 - P_{AF_i} \cdot \left(1 - \prod_{j=1}^{n} \left(1 - MF_{i/j} \right) \right) \right)
\]

For every risk factor outcome pair, the matrix of possible mediations was calculated and used. In the example the matrix of mediation when we aggregate BMI, cholesterol, FPG, and SBP would be:

<table>
<thead>
<tr>
<th></th>
<th>BMI</th>
<th>Cholesterol</th>
<th>FPG</th>
<th>SBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>0</td>
<td>0.111</td>
<td>0.148</td>
<td>0.296</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FPG</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SBP</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Calculating mediation factor

1 – **Comparing crude RR versus mediator-adjusted RR**

The best example is the mediation of BMI through SBP, FPG, and cholesterol reported by Danaei et al.33 In their meta-analysis, they report the adjusted and unadjusted RR of BMI on IHD and stroke based on combined data from individual cohorts. They calculated the mediation factor using equation 4 and we used it directly as mediation factor in risk factor aggregation.

For some risk factor aggregations we just simply added PAFs. For example, the total burden of smoking including smoking and secondhand smoke is the sum of the estimates of the individual risks because we estimate the burden of secondhand smoke in non-smokers only.

2 – **Estimating the mediation factor by pathway of the effect**

For many other risk factors there are no data available to use the first method. Instead, we searched studies to estimate the effect of the risk factor (for example fruit) on the mediator (SBP) and finally the expected increase in IHD risk. We pooled available studies to calculate the unit increase in the mediator per unit increase in the risk factor to calculate the size of the IHD RR.

**Figure. Example of pathway between fruit, high systolic blood pressure, and cardiovascular diseases**
We have RRs for the effect of A on C and B on C in GBD from a meta-analysis of studies in the literature. The effect of A on B was estimated by analysis of diet trials.

\[ RR_{ABC} = RR_{BC}^\Delta 
\]

\( RR_{ABC} \) is expected effect of A through B on C

\( RR_{BC} \) is relative risk of each unit increase in mediator on outcome C

\( \Delta \) is change in mediator level B per each unit change in A

If \( RR_{AB} \) is the overall effect of A on B then:

The mediation factor would be

\[ MF = \frac{RR_{ABC} - 1}{RR_{AB} - 1} \]

We kept uncertainty of each parameter by generating and following 1,000 draws of the estimates to calculate 1,000 draws of the posterior distribution of the mediation factor. We did not include risk-mediator pairs if the mediation factor was not significant at 5% level (more than 50 out of 1,000 draws were negative). We truncated the mediation factor distribution at 1 where the whole effect of the risk factor on the outcome would be assumed to be through the mediator pathway.

Some mediation factors equal 1 where the whole effect was calculated through other risk factor e.g. the effect of sugar-sweetened beverages through BMI or salt through SBP or when we assumed other risk factors are sources of the exposure, for example fiber is provided by consuming fruit, vegetable, and whole grains.

**Dietary risk factors**

We searched for diet trials that reported change in SBP, cholesterol, and FPG by change in dietary risk factors, for example fruit and vegetables. We did a systematic search to find clinical trials that reported the baseline values or change in diet levels. We also searched for a list of important hypothetical mediations primarily through metabolic risk factors because of the great burden of metabolic risk factors and a need to aggregate and control for double-counting of the burden, especially for cardiovascular diseases (CVD).
Considering that outcome of metabolic changes such as SBP and cholesterol are measured objectively (compared with subjective measurements that might be affected by patient or physician knowledge about the intervention group), and there are no issues with blinding and analytical concerns like type of analysis (intention-to-treat or per-protocol), we think they provide a sufficient data on the short-term effect of diet on metabolic risk factors.

Long-term effects are more difficult to capture, given that there is little data available on long-term effects in the literature. Future analysis of cohort studies will be necessary to understand the long-term effects of diet on metabolic risk factors.

We modeled change in a given mediator (e.g. cholesterol) per unit change in diet components. The best possible approach would be controlling for other dietary changes, but it is not possible because of few data points and uncertainty levels for both diet and metabolic risk change. With a limited number of studies providing data points for the analysis and no access to micro-data from diet trials, it is not possible to control for other diet components.

In cases in which there were very few data points, such as for unsaturated fatty acids and trans fats, or if we could not find trials, mediations were excluded. Also, BMI was excluded because our diet analyses are adjusted for a 2,000 calorie diet, thereby addressing mediation through BMI and obesity.

We did not include possible mediation/interaction of diet with many other risk factors and outcomes besides metabolic risks. Fruit and vegetables could have interaction with smoking and possibly air pollution on cancers, but we did not identify sufficient evidence for such an analysis. We assumed all effects of fiber are captured in fruit, vegetables, whole grain and nuts and seeds, so we assumed complete mediation.

In the case of fibre, the mediation is counted as one mechanism of producing covariance between risk factors, and the calculation depends on the concept and direction of mediation. To be consistent with the methodology employed in the GBD, we must aggregate and avoid double-counting of burden and we should control covariance. Covariance might be with or without interaction and mediation is one way of subsequent double-counting of the burden. We think that through mediation analysis we are able to quantify non-random and biologically plausible covariance and improve risk factor aggregation.

Physical activity
We found cohort studies on the effect of physical activity on FPG. The data was more on the effect of physical activity on diabetes incidence, so we calculated the shift in FPG using the provided RR value. We used this to calculate the mediated part of effect of physical activity on CVD.34–40

Air pollution
We looked for cohort and time series studies but the data were limited. We found only one study with the effect of last year average of particle pollution (PM) 2.5 on SBP, FPG and cholesterol.41 However, the effects through FPG and cholesterol were bigger than the effect expected for that level of PM2.5, indicating significant overestimation of the mediation. We found time series studies with different PM2.5 lag (by day) that show very short-term and confounded effects. So we decided to add this when stronger evidence is available.
Assumed mediations
For the risk factors with PAFs of 100% such as FPG and diabetes, low estimated glomerular filtration rate and chronic kidney disease, hypertension and hypertensive heart disease, alcohol and alcohol disorders, childhood underweight and protein-energy malnutrition, and childhood wasting and protein-energy malnutrition, and drug use and drug use disorders, no mediation is needed.

3 – Piecewise aggregation (Pattern 3)
There are three anthropometric indicators that are highly correlated: childhood underweight, stunting, and wasting, as demonstrated in Figure. Venn diagram demonstrating the correlation between childhood underweight, stunting, and wasting. Available RRs for each indicator are not adjusted for the other two because there is a high correlation between these indicators and also interaction where the majority of the burden occurs. Estimating the total burden due to undernutrition, a latent variable, is difficult. The three anthropometric indicators are not independent, so the covariance between them should be considered. This was the main reason that GBD 2010 only included childhood underweight. If covariance between these indicators is significant (as is shown in the Figure below), aggregating these indicators assuming independence would overestimate the total burden significantly.

To use the best available data, we adjusted observed RRs reported by Olofin et al for underweight, stunting and wasting by simulating the joint distribution of the three indicators using the distribution of each indicator and covariance between indicators in the countries included in the meta-analysis (extracted from Demographic and Health Survey (DHS) micro-data). Based on the analysis done by McDonald et al, we assumed there is an interaction between the three indicators, and extracted the interaction terms from the corresponding analysis. We calculated the adjusted RRs by minimizing the error between observed crude RRs (from meta-analysis) and expected crude RRs derived from adjusted RRs. To use the best available data, we adjusted observed RRs reported by Olofin et al for underweight, stunting and wasting by simulating the joint distribution of the three indicators using the distribution of each indicator and covariance between indicators in the countries included in the meta-analysis (extracted from Demographic and Health Survey (DHS) micro-data). Based on the analysis done by McDonald et al, we assumed there is an interaction between the three indicators, and extracted the interaction terms from the corresponding analysis. We calculated the adjusted RRs by minimizing the error between observed crude RRs (from meta-analysis) and expected crude RRs from adjusted RRs, interaction terms, and joint distribution of the risk factors.
After adjusting for the three risk factors, we calculated the PAFs and aggregated underweight, stunting and wasting burden.

**Uncertainty of aggregated and mediated PAFs**

We generated 1000 draws of posterior distribution of mediation factor calculated by different methods to use beside draws of other inputs to the PAF aggregation.

**Important assumptions in aggregating risk factors and including mediation**

1 – The mediation factors or PAF adjustments are similar across countries, age, sex, and years. While it is quite likely that the size of mediation is different in different populations, there is little data to inform the covariance between different risk factors or the mediation factor amount by age and countries. For example in some countries, the size of the mediated BMI-IHD PAF through cholesterol, calculated by the mediation factor, was even bigger than the total burden of cholesterol, indicating that less effect of BMI is mediated through cholesterol and mediation factors are not similar across countries.

2 – For many risk-mediator-outcome pairs, there are no data available, so we assumed the mediation is zero.

3 – Since the covariance between undernutrition indicators is different by countries (and across time, results were not reported), and there is an interaction between these indicators, the total burden might be underestimated.

4 – It is assumed that there is no significant covariance between PAFs, which might not be true between some risk factors such as between metabolic risk factors. While this overestimation is controlled by using adjusted RRs, using crude RRs for BMI and other metabolic risk factors may cause significant overestimation of aggregated metabolic risks burden.
7. Estimate attributable burden

Four key components are included in estimation of the burden attributable to a given risk factor: the metric of burden being assessed (the number of deaths, years of life lost [YLLs], years lived with disability [YLDs], or DALYs [the sum of YLLs and YLDs]); the exposure levels for a risk factor; the relative risk of a given outcome due to exposure; and the counterfactual level of risk factor exposure. Estimates of attributable burden as DALYs for risk-outcome pairs were generated using the following model:

\[ AB_{jasgt} = \sum_{o=1}^{W} DALY_{joasgt} PAF_{joasgt} \]

where \( AB_{jasgt} \) is the attributable burden for risk factor \( j \) for age group \( a \), sex \( s \), geography \( g \), and year \( t \); \( DALY_{joasgt} \) is total DALYs for cause \( o \) (of \( W \) relevant outcomes for risk factor \( j \)) for age group \( a \), sex \( s \), geography \( g \), and year \( t \); \( PAF_{joasgt} \) is the population attributable fraction (PAF) for cause \( o \) due to risk factor \( j \) for age group \( a \), sex \( s \), geography \( g \), and year \( t \). The proportion of deaths, YLLs, or YLDs attributable to a given risk factor or risk factor cluster were analogously computed by sequentially substituting each metric in place of DALYs in the equation above.

Other analysis: Decomposition of deaths and DALYs

We conducted two related decomposition analyses of changes in DALYs from 1990 to 2015: (1) decomposing changes in cause-specific DALYs due to changes in population growth, population age structure, exposure to all risks for a disease, and risk-deleted death and DALY rates; and (2) decomposing changes in risk-attributable all-cause DALYs due to changes in population growth, population age structure, risk exposure to the single risk factor, and risk-deleted DALY rates. In this case, risk-deleted rates are the rates after removing the effect of a risk factor or combination of risk factors; in other words, observed DALY rates multiplied by one minus the PAF for the risk or set of risks. Our decomposition analyses draw from methods developed by Das Gupta\(^44\) to provide a computationally tractable solution to isolating drivers of burden changes whereby all combinations of possible pathways are averaged across factors. Both the total burden and the attributable burden are determined, following the methods of Das Gupta, as a product of four factors such that:

\[ T_{asgt} = (A_{asgt} B_{asgt} C_{asgt} D_{asgt}) \]

where \( T_{asgt} \) represents either the total burden or the attributable burden at year \( t \); \( A_{asgt} \) is the all-age population size for a given sex \( s \) and geography \( g \) at year \( t \); \( B_{asgt} \) is the proportion of the population in the age group for a given age group \( a \), sex \( s \) and geography \( g \) at year \( t \); \( C_{asgt} \) is the underlying rate of the outcome unrelated to the risk factor or observed rate, multiplied by \( 1 - PAF \) for a given age group \( a \), sex \( s \) and geography \( g \) at year \( t \); and where \( D_{asgt} \) is the ratio of total burden (or attributable burden) to the underlying rate, which reflects the risk effect for a given age group \( a \), sex \( s \), and geography \( g \) at year \( t \) defined as \( 1/(1 - PAF) \) in the case of total burden or as \( PAF/(1 - PAF) \) in the case of decomposing attributable burden to a risk. The contribution of each factor to total change in either total burden or attributable burden was determined by changing the level of one factor from time \( t_0 \) to \( t_1 \) – here 1990 to 2015 – with all other factors held constant. Thus, the effect of any of the four factors, for example \( A_{asgt} \) on the change of total burden between 1990 (\( A_{90} \)) and 2015 (\( A_{15} \)) is calculated as:

\[ E_A = (A_{15} - A_{90}) \left( \frac{B_{90}C_{90}D_{90}+B_{15}C_{15}D_{15}}{4} + \frac{B_{90}C_{90}D_{15}+B_{15}C_{90}D_{90}+B_{15}C_{90}D_{90}+B_{15}C_{90}D_{15}}{12} \right) \]
Where $E_A$ is the proportion of change due to factor $A$, and the subscripts for each factor in the equation denote the year for each estimate. Since the effect depends on the order of entry of the factor, we calculated the average of all combinations of the four factors.$^{54}$

This four factor decomposition method does not work for risks where the PAF, by definition, is 100% (such as high fasting plasma glucose and diabetes) or where the PAF is directly estimated (such as for unsafe sex and HIV). In the cases of childhood underweight and protein-energy malnutrition, childhood wasting and protein-energy malnutrition, vitamin A deficiency and vitamin A deficiency, alcohol use and cirrhosis and other chronic liver diseases due to alcohol use, alcohol use and alcohol use disorders, alcohol use and liver cancer due to alcohol use, drug use and drug use disorders, iron deficiency and iron-deficiency anemia, and low glomerular filtration rate and chronic kidney disease, we used a three factor decomposition method, which examines the contribution of population, ageing, and risk exposure. Effectively, we assume trends in these cases are driven by exposure, not change in the risk deleted rates. For FPG and diabetes, we used GBD estimates of the prevalence of diabetes and the excess DALY rate for each prevalent case of diabetes to decompose trends in diabetes into the contribution of the four factors. We were not able to include three outcomes in this analysis: cervical cancer, sexually transmitted diseases excluding HIV, and HIV/AIDS.
Other analysis: Socio-demographic Index (SDI) analysis & Epidemiological Transition

a. Development of revised SDI indicator

We began exploring the relationship between a composite indicator of socio-demographic development in GBD 2013 DALYs. We used lag distributed income per capita (LDI), average educational attainment over the age 15 years, total fertility rate (TFR), and mean population age and called it SDS, socio-demographic status. In response to feedback, we excluded mean population age due its strong relationship to mortality rates. We renamed the indicator Socio-demographic Index (SDI). SDI has an interpretable scale: zero represents the lowest income per capita, lowest educational attainment, and highest TFR observed across all GBD geographies from 1980 to 2015 and one represents the highest income per capita, highest educational attainment, and lowest TFR.

SDI was calculated using the Human Development Index (HDI) methodology, wherein an index value was determined for each covariate input (log LDI, average educational attainment in the population over age 15, and TFR):

\[ I_{c,ly} = \frac{C_{ly} - \min(C)}{\max(C) - \min(C)} (C_{ly} - \min(C)) / (\max(C) - \min(C)) \]

Where \( I_{c,ly} \) – the index for covariate \( C \), location \( l \), and year \( y \) – is equal to the difference between the value of that covariate in that location-year and the minimum observed value of the covariate (\( \min(C) \)) in any location over the 1980-2015 time interval divided by the observed range (\( \max(C) - \min(C) \)). An additional innovation for GBD 2015 was to incorporate subnational locations where estimated (resulting in 519 unique administrative units) for the entire estimation period of 1980-2015. The Socio-Demographic Index is then the geometric mean of these three indices:

\[ SDI = \sqrt[3]{I_{lnLDI} I_{educ} I_{TFR}} \]

In our mortality analyses, for LDI and TFR, we noted depreciating gains in life expectancy at birth and 5q0 at the higher and lower terminals, respectively. Due to the significance of these values in indexing, we aimed to identify the point at which increasing income or reducing fertility no longer resulted in improved child mortality or life expectancy. We tested various restrictions, and found that capping LDI at $60,000 and setting a TFR floor at 1 resulted in improved correlations with the resultant health indicators.

We further aimed to validate the use of SDI by regressing it in a variety of forms against life expectancy at birth, 5q0, 35q15, and 20q50. We found that SDI generally is as capable of predicting these demographic indicators as the previous SDS, and also as the inputs. We also found that in incorporating year, we did not substantially reduce the coefficients for SDI. Additionally, in testing lags of 2-10 years, we found the version with no lag to be the most predictive. Appendix Table 5 has SDI values by GBD geography over time and illustrates these results more in-depth.
b. Age-sex-specific relationships between SDI and SEVs

In order to evaluate the relationship between SDI and SEVs, we fit a simple least-squares regression using a smoothing spline on SDI for every cause in levels 1, 2, and 3 of the GBD cause hierarchy:

$$\ln(y_{l,y,a,s,c}) = \sum_{i=0}^{d+k} \beta_i B_{i,d}(SDI) + \gamma_U + \gamma_E + \gamma_C + \gamma_O + \epsilon_{l,y,a,s,c}$$

where:

- $\ln(y_{l,y,a,s,c})$ is the logit SEV in location $l$ and year $y$, and for age $a$, sex $s$, and cause $c$
- $\sum_{i=0}^{d+k} \beta_i B_{i,d}(SDI)$ is resultant parametric curve, of degree $d$ and interior knots $k$, of a linear combination of basis splines $B_{i,d}(SDI)$
- $\gamma_U$ is a dummy variable for the United States
- $\gamma_E$ is a dummy variable for the GBD region Eastern Europe
- $\gamma_C$ is a dummy variable for the GBD region Central Asia
- $\gamma_O$ is a dummy variable for the GBD region Oceania
- $\epsilon_{l,y,a,s,c}$ is the error term for location $l$, year $y$, age $a$, sex $s$, and cause $c$

Regressions were run separately by age, sex, and cause, using all location-years. Dummy variables were included for locations that were identified in modeling to skew fit due to significant deviation from levels of morbidity observed elsewhere at similar levels of SDI. In the case of the United States because of the inclusion of 50 states, the US collectively had an undue influence on the shape of the relationship which is why a separate dummy variable was included for the US. Because of the mortality crisis in Eastern Europe and Central Asia after the collapse of the Soviet Union, we included a dummy variable to adjust for the mean difference in these regions.

Having a complete set of age specific SEVs, we were then able to produce a full set of age-standardized rates for every SDI level. We evaluated this relationship at each centile value of SDI (i.e., by increments of 0.01). The SDI ranged from 0.060 in Mozambique in 1987 to 0.978 in the District of Columbia, United States in 2015.

We used the same modeling set up but used the logit of the share of population in each age-group as the dependent variable to estimate a smoothed relationship between population age-structure and SDI. Predictions for each age-group at each level of SDI were rescaled to sum to 100%.
References


Section 3. Risk-specific estimation

The risk-specific modeling write-ups follow the order of the risk factor hierarchy for GBD 2015. In some cases, multiple risk factors are addressed in a single write-up, for example childhood underweight, wasting, and stunting are all included in a single detailed write-up.
Unsafe Water Capstone Appendix

Flowchart

Unsafe Drinking Water

Input Data & Methodological Summary

Exposure

Case Definition
For GBD 2015, exposure to unsafe water is defined based on reported primary water source used by the household and use of household water treatment (HWT) to improve the quality of drinking water before consumption. Water sources were defined as improved based on the JMP designation (The WHO), which includes piped water as improved water, and households with access to piped water connection to the house, yard, or plot were defined as having access to piped water supply. Solar treatment, chlorine treatment, boiling, or the use of filters were all assumed to be effective point-of-use household water treatments, and based on effect sizes published by Wolf et al. (2014) boiling or filtering was the most effective form of water treatment.

Input Data
The search for usable household surveys and censuses was conducted using the Global Health Data Exchange (GHDx) database. All surveys through December 2015 that provide household level micro-data on water source were added. Tabulated and report data was lower priority and was only updated when time permitted. HWT input data was limited to two large survey series (DHS and MICS) due to time constraints. An update to HWT input data is a top priority for estimating exposure to unsafe water in future iterations.

Modeling
Water source data is modeled in two distinct categories: household prevalence of improved water and household proportion of piped water within improved population in order to prevent the population with...
Unsafe Sanitation Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case Definition
Exposure to unsafe sanitation were defined based on the primary toilet type used by households. Improved facilities are defined as such based on JMP designation (The WHO). Sewer connection toilets included flush toilets or any toilet with connection to the sewer or septic tank.

Input Data
The search for usable household surveys and censuses was conducted using the Global Health Data Exchange (GHDx) database. Searches were conducted from October 2015 to December 2015, with the final search household level micro-data on toilet type conducted on December 15, 2015. Due to the organized nature of the GHDx, the only search term used was “unsafe sanitation”, which yielded just under 1400 results, of which 795 were extracted and used as inputs for modeling. Tabulated and report data was lower priority and was only updated when time permitted.

Modeling
There were no substantive changes in the modeling process from GBD 2015. Two distinct models are produced from sanitation data: prevalence of households with improved sanitation and the proportion of households with a sewer connection over the total improved sanitation population. Prevalence of households with a sewer connection is modeling with improved sanitation prevalence as the denominator in order to prevent the population with access to sewer connection from exceeding the population with access to improved sanitation. By each geography-year, both models are generated using a 3-step modeling scheme of mixed effect linear regression followed by spatio-temporal Gaussian process
regression (ST-GPR), which outputs full time series estimates for each GBD 2015 location. Socio-demographic status (SDS), an index metric that includes a measure of education and income level, was used as a fixed effect in the linear regression since it proved to have significant coefficients. Random effects were placed at GBD 2015 region and super-region levels.

The process of vetting and validating models was accomplished primarily through an examination of ST-GPR scatter plots by GBD 2015 location from 1990-2015. Any unfitting data points were re-inspected for error at the level of extraction and survey implementation, and subsequently excluded from analysis if deemed appropriate. In addition to SDS, a number of different potential fixed effects were considered, including lag-distributed income and urbanicity, but SDS proved to be the strongest predictor of unsafe sanitation. Uncertainty in the estimates was initially formed based on standard deviation by survey, then propagated through ST-GPR modeling by means of confidence intervals around each data point that reflect the point-estimate specific variance.

Once models are fully vetted, full time series outputs from ST-GPR modeling are then converted from proportion to prevalence by year and geography and then rescaled to form 3 mutually exclusive categories that sum up to 1. The table below provides the final result of this rescaling.

<table>
<thead>
<tr>
<th>Category</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unimproved sanitation</td>
<td>Proportion of households that use unimproved sanitation facilities.</td>
</tr>
<tr>
<td>Improved sanitation, excluding sewer</td>
<td>Proportion of households that use improved sanitation facilities except those with sewer connection.</td>
</tr>
<tr>
<td>Sanitation facilities with sewer connection</td>
<td>Proportion of households that use toilet facilities with sewer connection.</td>
</tr>
</tbody>
</table>

Due to the nature of modeling sanitation with sewer connection as a proportion of total improved sanitation access, we are limited in only using sources for sewer connection that also include total improved sanitation values. It should be noted that high-income countries are assumed to have risk of unsafe sanitation which could lead to an underestimate of unsafe sanitation health burden in these countries. Another limitation that extends to the other two risk factors that comprise WaSH (unsafe water and unsafe hygiene) and can be improved upon in future iterations is taking into account covariance of access to water, sanitation and handwashing facilities. Currently, all three components of WaSH are modeled independently, which may lead to an overestimation of the burden of WaSH factors. High-income countries were assumed to have 0 risk of unsafe sanitation and the TMREL was applied to these countries.

**Theoretical minimum-risk exposure level**

The theoretical minimum-risk exposure level for unsafe sanitation was defined as all households have access to a sanitation facility with sewer connection. Since it is assumed that all households in high-income countries have access to sewer-connected sanitation, this counterfactual exposure level is applied to all households in high-income countries.
Relative risks
GBD 2015 employ the same relative risks for unsafe water as was done for GBD 2013. Three adverse health outcomes are paired with unsafe sanitation, which comprise of diarrheal diseases, typhoid fever, and paratyphoid fever. A meta-analysis by Wolf et al. 2014 provides relative risk evidence for the relationship between unsafe sanitation and diarrheal diseases. In the absence of better data, the relative risk for typhoid and paratyphoid fevers were assumed to be the same as the relative risk for diarrheal disease. Please refer to appendix tables for more information on relative risk values and citations.

References
Unsafe Hygiene Capstone Appendix

Flowchart

Unsafe Handwashing

Input Data & Methodological Summary

Exposure

Case Definition
Unsafe hygiene is composed of global handwashing practices. Handwashing is defined as the observed prevalence of handwashing with soap and water after using a toilet or after contact with excreta, including children’s excreta. We estimate the burden of unsafe handwashing in both developed and developing settings.

Input Data
There were two main sources that were used in our estimation of handwashing practices, estimates from scientific literature and estimates from household survey series. Relevant literature on handwashing prevalence was gathered from a meta-analysis published recently by Freeman et al. (2014). Since water and soap availability data is very limited, only country-specific Demographic Health Surveys (DHS) and Malaria Indicator Survey Series (MICS) conducted after 2006 were able to be used as input data.

Modeling Strategy
Input data from scientific literature and input data from household survey series were modeled independently. Data from literature primarily measured a population’s handwashing practices under ideal conditions, such as when water and soap was readily available. Additionally, these estimates from literature would likely be susceptible to acquiescence bias. Alternatively, data from DHS and MICS only provide insight into the availability of water, soap, and washing stations, which, alone, does not indicate how often a person may wash their hands after contact with excreta. Thus, after modeling data from
literature and data from surveys independently, these values were multiplied together by location-year in order to gain a more accurate representation of true handwashing prevalence.

For GBD 2015, there was a shift away from the ST-GPR modeling used in 2013 toward a more basic one-step modeling approach. This change came in light of the data scarcity and concern that spatial and temporal smoothing within ST-GPR may capture spurious trends in hygiene prevalence. For modeling the act of handwashing under ideal conditions, a variance-weighted linear regression with a fixed effect on average years of education per capita was employed. For modeling the availability of water, soap, and wash station, a multi-level logistic regression with a fixed effect on lag-distributed income per capita and random effect at the GBD 2015 region level was chosen to be the most appropriate method. The fixed effects used in both models, education and LDI per capita, proved to be significant and in the expected direction. Uncertainty intervals were produced by generating a 1000 draws from a normal distribution of the beta, intercept, and random effect, if appropriate, from the variance-covariance matrix of the regression.

The process of vetting and validating models was accomplished primarily through an examination of ST-GPR scatter plots by GBD 2015 location from 1990-2015. Any unfitting data points were re-inspected for error at the level of extraction and survey implementation, and subsequently excluded from analysis if deemed appropriate. In addition to LDI per capita and education per capita, a number of different potential fixed effects were considered, including socio-demographic status and urbanicity, however LDI and education proved to be the strongest predictors of handwashing practices for their respective models. Once models were sufficiently vetted, full time series outputs from each of the models were multiplied together at each location-year.

A considerable limitation for when estimating handwashing practices for over 190 independent locations around the world is data sparseness. Even when data is published on handwashing prevalence, the definition is often altered from the GBD 2015 standard definition or it may only pertain to certain populations (such as hospital patients) and lacks representativeness at the geographic scale we require. The incorporation of questions about soap and water availability in DHS and MICS has added much-needed information but there remains a large data gap that must be filled if we are to become more certain in handwashing estimates.

**Theoretical minimum-risk exposure level**
The theoretical minimum-risk exposure level for unsafe hygiene is defined as all households engaging in handwashing with soap practices after any contact with excreta, including children’s excreta.

**Relative risks**
GBD 2015 use the same relative risks for unsafe hygiene as was done for GBD 2013. There are 3 adverse health outcomes paired with unsafe hygiene that include diarrheal diseases, typhoid fever, and paratyphoid fever. A meta-analysis by Freeman et al. 2014 provides relative risk evidence for the relationship between unsafe hygiene and diarrheal diseases. In the absence of adequate data, the relative risk for typhoid and paratyphoid fevers were assumed to be the same as the relative risk for diarrheal disease based on analogous transmission pathways (feco-oral pathway). Please refer to appendix tables for more information on relative risk values and citations.
References

Ambient Particulate Matter Pollution Capstone Appendix

Flowchart

Input data & Modelling strategy

Exposure

Definition

Exposure to ambient air pollution is defined as the population-weighted annual average mass concentration of particles with an aerodynamic diameter less than 2.5 micrometers (PM$_{2.5}$) in a cubic meter of air. This measurement is reported in µg/m$^3$.

Input data

The data to estimate exposure to ambient air pollution is drawn from estimates of annual concentration of PM$_{2.5}$ – generated using satellite observations of aerosols in the atmosphere. To correct for bias in the satellite modeling approach, a spatially-varying flexible framework is used to combine modeled concentrations with observations from ground-level monitoring of particles in more than 75 countries. All input data for GBD2015 was updated as follows:

Updated PM$_{2.5}$ ground measurement database

For the GBD2015 update we updated the database of annual average PM measurements to include more recent data and to incorporate additional locations where measurement data have become available. To facilitate this we collaborated with WHO and contributed to their recently released *WHO Air Pollution in Cities database*. We then used disaggregated (monitor-specific values and not the city averages that are reported by WHO) measurements from this database with additional site-specific information (e.g. all monitors in a city, monitor geo coordinates, monitor site type) such as that included in the GBD2013 database. In total measurements of concentrations of PM$_{10}$ and PM$_{2.5}$ were retrieved from 6,003 ground
monitors with the majority contributing measurements from 2014 (as there is a lag in reporting measurements, little data from 2015 were available). Where data were not available for 2014 (2760 monitors), data was used from 2015 (18 monitors), 2013 (2155), 2012 (564), 2011 (60), 2010 (375), 2009 (49), 2008 (21) and 2006 (1). For locations with only PM$_{10}$ measurements, PM$_{2.5}$ measurements were estimated from PM$_{10}$. This was done by a locally derived conversion factor (PM$_{2.5}$/PM$_{10}$ ratio) estimated as population-weighted averages of location-specific conversion factors for the country. Location-specific conversion factors were estimated as the mean ratio of PM$_{2.5}$ to PM$_{10}$ of stations for the same year. If national conversion factors were not available, regional ones were used, which were obtained by averaging country-specific conversion factors.

**Updated satellite-based estimates**

The updated satellite-based estimates are described in detail in van Donkelaar et al. 2016$^1$. These estimates (~11 x 11 km resolution at the equator) combine aerosol optical depth retrievals from multiple satellites with the GEOS Chem chemical transport model and land use information.

**Updated population data**

A comprehensive set of population data on a high-resolution grid was obtained from the Gridded Population of the World (GPW v4) database. These data are provided on a 0.0417°×0.0417° resolution. To aggregate these estimates of population to each 0.1°×0.1° grid cell, the central 3 × 3 population cells were summed. As this accounted for a resolution higher than necessary, the same was done four other times, offset by one cell in a North, South, East and West direction. The average of five quantities was used as the aggregated population estimate for each cell. Estimates of population for 2000, 2005, 2010, 2015 and 2020 were extracted from GPW version 4 and estimates for 1990 and 1995 were extracted from GPW version 3 as described previously for GBD2013$^3$.

**Modelling strategy**

The methodology used to estimate the burden of ambient particulate matter pollution has seen significant changes since GBD2013.

The GBD2010 and GBD2013 estimates both used a single global function to calibrate the mean of the chemical transport model and satellite-based estimates to available ground measurements. In both instances the approach taken was recognized at the time to be a compromise between what could be easily implemented under tight timeframes and one that most efficiently utilized all of the data sources. In particular, the GBD2013 exposure estimates were known to underestimate ground measurements in specific locations (see discussion in Brauer et al., 2015$^5$) such that it would be desirable to allow measurements to make a larger contribution to the final estimates where they were available. Therefore, for GBD2015 we implemented a Bayesian Hierarchical modelling approach using Integrated Nested Laplace Approximations (INLA) in which the satellite-based estimates, ground measurements and land use information are combined in a spatially varying flexible framework. Formal external evaluation using ground measurements was conducted and shown to lead to improved predictions of ground measurements in all super regions compared to GBD2013 estimates and an alternative geographically-weighted regression approach. Further, based on the external evaluation analyses, addition of the TM5
chemical transport model estimates of PM2.5 annual average did not improve the estimates and these were therefore not included.

Bayesian hierarchical models (BHM) provide an extremely useful and flexible framework in which to model complex relationships and dependencies in data. Uncertainty can also be propagated through the model allowing uncertainty arising from different components, both data sources and models, to be propagated through the models into estimates of uncertainty associated with the final estimates. In the hierarchical modeling approach coefficients associated with satellite-based estimates were estimated for each country. Where data were insufficient within a country, information can be ‘borrowed’ from a higher aggregation (region) and if enough information is still not available from an even higher level (super-region). Individual country level estimates were therefore based on a combination of information from the country, its region and super-region.

All modelling was performed on the log-scale with the unit of measurement being a grid cell. The model was constructed with the inclusion of all variables assessed statistically, based on model fit (DIC, a measure of the relative quality of a model and closely related to that of AIC but for Bayesian models) and predictive ability. The hierarchical structure was applied to the intercept and slope terms with all modelling on the log scale. The model was of the form

$$\log(\text{PM2.5}_i) = \beta_0 + \beta_1 \log \text{SAT}_i + \text{other variables} + \epsilon_i$$

where $i$ denotes the grid cell. The following sets of variables were considering in developing the models:

Continuous explanatory variables:

- (SAT) Estimate of PM$_{2.5}$ (in $\mu$g m$^{-3}$) for 2014 from satellite remote sensing on the log-scale.
- (CTM) Estimate of PM$_{2.5}$ (in $\mu$g m$^{-3}$) for 2014 from chemical transport models on the log-scale.
- Estimate of population for 2014 on the log-scale.
- (SNAOC) Estimate of the sum of sulfate, nitrate, ammonium and organic carbon as estimated from GEOS Chem
- (DST) Estimate of compositional concentrations for mineral dust from GEOS Chem
- (EDxDU) The log of the elevation difference between the elevation at the ground measurement location and the mean elevation within the GEOS Chem simulation grid cell multiplied by the inverse distance to the nearest urban land surface

Discrete explanatory variables:

- Binary variable indicting whether exact location of ground measurement is known
- Binary variable indicting whether exact type of ground monitor is known
- Binary variable indicting whether ground measurement is PM$_{2.5}$ or converted from PM$_{10}$

Random Effects:
- Grid cell random effects on the intercept to allow for multiple ground monitors in a grid cell.
- Country-region-super-region hierarchical random effects for the intercept
- Country-region-super-region hierarchical random effects for the satellite remote sensing term.
- Country-region-super-region hierarchical random effects for the coefficient associated with the difference between estimates from CTM and SAT.
- Country-region-super-region hierarchical random effects for the coefficient log(Pop)
- Country level random effects for intercept, satellite and difference between CTM and SAT are independent and identically distributed.
- Country level random effects for population uses a neighbourhood structure allowing specific borrowing of information from neighbouring countries.
- All region random effects are assumed to be independent and identically distributed.
- All super-region random effects are assumed to be independent and identically distributed.

Interactions:

- Interactions between the binary variables and the effects of log(SAT) and log(CTM/SAT)

Due to both the complexity of the models and the size of the data, notably the number of spatial predictions that are required in this setting, recently developed techniques that perform ‘approximate’ Bayesian inference based on integrated nested Laplace approximations (INLA) have been developed as a computationally attractive alternative to Markov Chain Monte Carlo methods. Computation was performed using the R interface to the INLA computational engine (R-INLA) with the size of the task of fitting the models and performing predictions for each of the ca. 1.4 million grid cells requiring the use of a high performance computing cluster (HPC) with high memory nodes. As in GBD2010 and GBD2013 the spatial model was built combining the different data sources for a single year (2014, corresponds to the most recent measurement data). The spatially-varying functions from this model were then applied to the satellite-based estimates from all other years - in other words assuming that the spatial relationship between the different data sources does not change over time. This is undoubtedly a simplification but to do otherwise would require assembling multi-year measurement databases which is not feasible given current data availability and computational constraints. As the spatial model was built using the most recently available (2014) measurement and satellite-based estimates, 2015 estimates were based on extrapolation. Instead of extrapolating using an exponential model based on a 1-year trend as in GBD2013, splines based on a 5 year trend (2010-2014) were fit and applied to the 2014 grid-cell values to estimate levels for 2015. This reduced the likelihood of 2015 estimates being overly influenced by meteorological events in a specific year and to better represent the duration of exposure relevant to the epidemiologic studies included in the integrated exposure-response functions.

Model Evaluation

Model evaluation and comparison was performed by fitting models on a training set and predicting exposures at locations for which measurements were known (the validation set). The selection of the training (20%) and validation (80%) set consisted of taking a random sample of the total number of sites measuring PM2.5 (or having a value converted from PM10 measurements). Sampling was performed
using sampling probabilities based on the cross-tabulation of PM2.5 categories (0-24.9, 25-49.9, 50-74.9, 75-99.9, 100+ µg/m³) and super-regions. The resulting hold-out evaluation data set was a sample of 20% of the sites that have the same distribution over PM2.5 categories and super-regions as the entire set of sites.

This process was used to generate multiple training and validation set combinations, allowing for example cross-validation to be performed. In the evaluation, 25 sets of training/validation data were used. The following models were considered in the evaluation phase:

(A) The GBD2013 model, using a simple linear regression with a fused estimate of SAT and CTM together with interactions with three binary variables representing whether the measurement was converted from PM10 and whether the exact site type and location is known.

(B) A hierarchical model with SAT, the TM5 CTM estimates, population and the three binary variables described above

(C) A hierarchical model with SAT, population, SNAOC, DST, EDxDU, population and the three binary variables
   - Estimate of population for 2014 on the log-scale.
   - Estimate of the sum of sulfate, nitrate, ammonium and organic carbon as estimated from GEOS Chem
   - Estimate of compositional concentrations for mineral dust from GEOS Chem
   - The log of the elevation difference between the elevation at the ground measurement location and the mean elevation within the GEOS Chem simulation grid cell multiplied by the inverse distance to the nearest urban land surface

For each training/evaluation set combination, model fit and prediction accuracy were evaluated for each of the 25 training/evaluation set combinations with the following metrics:

**Model fit**
- $R^2$
- DIC

**Predictive accuracy**
- $R^2$ arising from a linear regression of predicted vs actual measurements at each location
- RMSE – root mean squared error
- WRMSE – weighted (by population) root mean squared error
- MSE – mean square error
- MAE – mean absolute error

This evaluation indicated the final model improved predictions of ground measurements in all super regions compared to GBD2013 estimates (median global $R^2$ [population-weighted RMSE] 0.82 [12.1 µg/m³], 0.60 [13.5 µg.m³] for GBD2015 and GBD2013, respectively).

Error! Reference source not found. shows the RMSE (median from the 25 runs) for each of the three models, by super-region. The accuracy of the prediction varies between super-regions, with lower errors being observed in areas where there are more monitoring sites. In each of the super-regions, the largest errors are seen for model A which are considerably higher than those for models B and C, with model C showing a small improvement over B (except in super-region 5, North Afirca/Middle East).
Figure 2 shows scatter plots of the observed and predicted measurements using the three models for each super-region. The predicted measurements are the median values over those obtained from the 25 training sets. Predictions from the two Bayesian hierarchical models (B&C) match the observed values more closely than the linear model (A) with much less spread around a straight line (with slope one and zero intercept, shown in red). In Central Europe and Sub-Saharan Africa it is noticeable that, in addition to reduced spread, models B&C are much better at predicting higher values. The same patterns of results in predictive ability were seen when looking at regions and individual countries. In all cases, model C performed better than model B with both being considerable better than model A.

Figure 1: Comparison of RMSE from three models by super-region. Dots denote the median of the distribution from 25 training/evaluation sets and the vertical lines the range of values. Super-regions are

57
1: high income, 2: Central Europe, Eastern Europe, Central Asia, 3: Latin America and Caribbean, 4: Southeast Asia, East Asia and Oceania, 5: North Africa / Middle East, 6: Sub-Saharan Africa, 7: South Asia.

Figure 2: Comparison of observed and predicted measurements using three different models, by super-region. The red line has slope one and intercept zero.

Overall, the best model in terms of model fit and predictive ability was one with the following components:

- Estimates of PM$_{2.5}$ (in $\mu$gm$^{-3}$) from satellite remote sensing (SAT), population, and information on the GEOS Chem simulated chemical composition, elevation and distance to urban land use (SNAOC, DST and EDxDU).
- Binary variables indicating whether exact location and type of ground measurement is known, and whether the measurement was PM$_{2.5}$ or converted from PM$_{10}$.
- Interactions between the binary variables and the effects of estimates from satellite remote sensing.
- Grid cell random effects on the intercept to allow for multiple ground monitors in a grid cell.
Country-region-super-region hierarchical random effects for intercepts, satellite remote sensing and population terms.
- Country level random effects for population using a neighbourhood structure allowing specific borrowing of information from neighbouring countries.

**Theoretical minimum-risk exposure level**

The TMREL for ambient PM is estimated using a uniform distribution between the minimum and 5th percentile of exposure observed in the studies used to generate the GBD estimates. This estimate was updated for GBD2015 as new studies were added to the analysis and studies used previously were updated through continued follow-up. The newer estimates included several large studies that included exposure at lower levels of PM2.5. As a result, the TMREL for GBD2015 was \(~U(2.4, 5.9)\), lower than GBD2013’s distribution \(~U(5.9, 8.7)\), which had the effect, all things being equal, of increasing the estimated attributable burden relative to the GBD 2013 estimates.

**Relative Risk**

Relative risks are generated using integrated exposure-response functions (IER) that are fit to available epidemiologic data using a Bayesian MCMC approach and a modified power function. The IER are estimated based on published relative risks for long-term exposure to ambient PM2.5, household air pollution, second-hand smoking, and active (cigarette) smoking. The concentration of particulate matter for each type of exposure is estimated based on literature values and used to map the relative risks to a curve generated for the entire range of exposure from these sources. The input data for this curve fitting process has been updated since GBD2013, adding new studies that estimate exposure at finer spatial scales, including studies of within-city exposure that focus on traffic-related air pollution. In addition, changes were made to the curve-fitting process. In order to account for differences in study design, temporal patterns of exposure and other differences among the studies of the different sources of PM2.5, a source-specific heterogeneity parameter was added to the IER. This resulted in much wider, and, in our view, more realistic, uncertainty intervals for the burden estimates, by propagating through the entire process the current uncertainty regarding the mechanisms and magnitude of health impacts of exposure to PM2.5 from diverse sources.
IER Functional Form

Data Likelihood

\[ \log(RR_i) \sim \mathcal{N}(\mu_i, \sigma^2_i + \delta_{\text{source}_i}) \]

Model

\[ \mu_i = \log \left( \frac{1 + \alpha \times \left( 1 - e^{-\beta \times (\text{exposure}_i, \text{TMREEL})} \right)}{1 + \alpha \times \left( 1 - e^{-\beta \times (\text{counterfactual}_i, \text{TMREEL})} \right)} \right) \]

Data

- \( RR_i \): measured relative risk for data point \( i \)
- \( \sigma^2_i \): variance of data point \( i \) based on study information
- \( \text{source}_i \): exposure source type (outdoor/household air pollution, secondhand/active smoking)
- \( \text{TMREEL} \): theoretical minimum risk exposure level
- \( \text{exposure}_i \): measured exposure for data point \( i \)
- \( \text{counterfactual}_i \): counterfactual exposure for data point \( i \)

Priors

\[ \begin{align*}
\alpha & \sim \Gamma(1.0, 0.01) \\
\beta & \sim \Gamma(1.0, 0.01) \\
\gamma & \sim \Gamma(1.0, 0.01) \\
\delta & \sim \Gamma(1.0, 0.01)
\end{align*} \]

We also modified the way in which age-specific IER for IHD and stroke were estimated. In accordance with previously published work on other cardiovascular risk factors, the impact of air pollution on cardiovascular health is known to vary with age. To account for this phenomenon, age-specific RRs were based on a log-linear model of RR as a function of age, where the intercept (RR=1) is forced to age 110. In GBD2010 and GBD2013 the age for a relative risk estimate from a given study was estimated as the median age at death or disease incidence in that study. However, this may not accurately represent the age distribution of the entire study population so we re-estimated this variable as the mean age at enrollment + half of the average follow-up time to better represent the average age of the study population during the period of follow-up. When compared to GBD2013, this change produced RRs that were generally lower for younger age groups, given that median age at event tends to produce a higher anchor age than average age during follow-up.

The relative risks are generated on the grid-level based on estimated exposure, and then applied to generate PAFs. These PAFs are aggregated using the grid-level population to create population-weighted national estimates of attributable burden, using the following formula:

PM2.5 Aggregation Formula

\[ PAF_{A, C, L} = \frac{\sum ((RR_{A, C} - 1) \times Pop_i)}{\sum (RR_{A, C} \times Pop_i)} \]

\( A = \text{age group}, C = \text{cause}, L = \text{location}, i = \text{grid}, \ RR_{A, C} = \text{grid-level RR based on PM}_{2.5} \) and given age/cause IER curve.
References


Household Air Pollution Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case Definition
Exposure to household air pollution from solid fuels (HAP) is defined as the proportion of households using solid cooking fuels. The definition of solid fuel in our analysis includes coal, wood, charcoal, dung, and agricultural residues.

Input data
Data were extracted from the standard multi-country survey series such as Demographic and Health Surveys (DHS), Living Standards Measurement Surveys (LSMS), Multiple Indicator Cluster Surveys (MICS), and World Health Surveys (WHS), as well as country-specific survey series such as Kenya Welfare Monitoring Survey and South Africa General Household Survey. To fill the gaps of data in surveys and censuses, we also downloaded and updated HAP estimates from WHO Energy Database and extracted from literature through systematic review done in IHME. Each nationally or sub-nationally representative data point provided an estimate for the percentage of households using solid cooking fuels. Estimates for the usage of solid fuels for non-cooking purpose were excluded, i.e. primary fuels for lighting. The database, with estimates from 1980 to 2015, contained 685 studies from 150 countries. Updates to systematic reviews are performed on an ongoing schedule across all GBD causes and risk factors, an update for household air pollution will be performed in the next 1-2 iterations.
Modeling strategy
Household air pollution was modeled at household level using a three-step modeling strategy ST-GPR that uses linear regression, spatiotemporal regression and Gaussian Process Regression (GPR). The first step is a mixed-effect linear regression of logit-transformed proportion of households using solid cooking fuels. The linear model contains maternal education and proportion of population living in urban areas as covariates and has nested random effect by country, GBD region, and GBD super region respectively. The full ST-GPR process is specified elsewhere in this appendix.

Compared with GBD 2013, we have made changes in terms of the covariates utilized in the linear model. A variety of combinations of socioeconomic and environmental covariates in different transformation format were tested by running mixed-effect models with exposure data. The final list of covariates included in the exposure model are maternal education and the proportion of population living in urban area.

Theoretical minimum-risk exposure level
For outcomes where we extracted RR based on direct epidemiological evidence i.e. COPD, lung cancer, and cataract, TMREL was defined such that no households would report using solid fuel as their primary cooking fuel. For outcomes that utilize evidence based on the Integrated Exposure Response (IER), the TMREL is defined as uniform distribution between 33.3 and 41.9 ug/m^3. TMREL for household air pollution did not change from GBD 2013.

Relative risks
The disease-outcomes paired with household air pollution has not changed since GBD 2013. The list of outcomes paired with household air pollution has not changed since GBD 2013, which included lower respiratory infections (LRI), stroke, Ischemic Heart Disease (IHD), chronic obstructive pulmonary disease (COPD), lung cancer and cataract. The relative risks of all outcomes but not cataract were generated by using the integrated exposure-response functions (IER). The relative risks for cataract were extracted from a meta-analysis paper (1). The IER curves are updated to reflect the newly updated data and utilization of a new method that specified elsewhere.

PM2.5 mapping value
The relative risk estimates describing the association of HAP with outcomes including Ischemic Heart Disease (IHD), cardiovascular disease (CVD), and lower respiratory infections (LRI) were derived from the IER curves. This is done by first estimating the crosswalk values that map household use of solid fuel to PM2.5 exposure because the IER curve measures exposure using PM2.5. This step of the analysis relied on 67 studies conducted in 16 countries to generate the PM2.5 mapping values, which remain the same sources as GBD 2013. The PM2.5 exposure was then cross-walked to men, women and children by generating the ratio of personal exposure to average 24-hour kitchen PM2.5 concentration based on a study after the literature review in GBD 2013.

References
Ambient Ozone Pollution

Flowchart

Input data and Methodological Summary

Exposure

Case Definition

For GBD 2015, exposure to ozone pollution is defined as the number of parts-per-billion (ppb) of ozone (O\textsubscript{3}).

Input data

Data for estimating ozone exposure is derived from the TM5-FASST chemical transport model, which generates a 3-month running average of daily 1 hour maximum ozone values at the 0.1° x 0.1° for the years 1990, 2000, and 2010.\textsuperscript{1}

Modeling Strategy

The process for modeling ozone exposure has remained stable since GBD2010 and GBD2013. Natural cubic splines were used to interpolate for the years 1995, 2005, and 2011. Annualized rate of change was used to predict for the years 2013 and 2015. The uncertainty for exposure at the grid-level was assumed to be ±6% of the estimated concentration, in accordance with previous work. Uncertainty for ozone was calculated by assuming a +/- 6% uncertainty interval around the estimation concentration.
Theoretical minimum-risk exposure level
The TMREL of ozone was defined based on the exposure distribution from American Cancer Society CPS-II study, which was the source of the GBD 2015 ozone mortality RR estimate. As with PM2.5, a uniform distribution was drawn around the minimum and 5th percentile values experienced by the cohort. This value was not updated for GBD 2015, and continues to be defined as $\sim U(33.3, 41.9)$, in ppb.

No other significant changes were made from GBD 2013 to GBD 2015.

Relative Risks
The relative risk of ozone exposure for respiratory COPD was extracted from literature and was not updated for GBD 2015. The relative risk is applied linearly per 10 ppb of ozone exposure and is defined as 1.029 (1.010-1048).²

References
Radon Exposure Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case definition

Radon is a radioactive gas that is produced as a byproduct of the decay chain of uranium, occurring naturally within the Earth’s crust. Some fraction of this natural radon production escapes into the atmosphere, where it forms a low concentration unless build-up is caused by enclosed spaces like homes, mines, or caves. Radon exposure is expressed as average daily exposure to indoor air radon gas levels measured in Becquerels (disintegrations per second) per cubic meter (Bq/m³).

Input Data

Exposure to radon is determined using values curated by an expert group. These values are taken from a variety of sources including literature, government agencies, and monitoring stations. Their methodology is then inspected to determine if they are robust enough to be considered as country-level averages. This dataset was last updated for GBD2013 by adding new datapoints across time and space. No new datapoints were added for GBD2015.

Modelling Strategy

There has been minor change to the methodology to estimate radon exposure. The modelling process was previously updated by shifting it from a nested random effects model to spatial-temporal GPR. For GBD2015, the spatial-temporal GPR modelling methodology was updated as detailed in the appendix specific to this analytical technique, which is common to a variety of risk factors. Radon is naturally occurring, and is not considered to have much temporal fluctuation\(^1\). As such, we did not model radon over time, opting instead to use all datapoints for a single year, predict across space using our radon
database, and use the results for that year for the entire GBD time series. This eliminated any spurious time trends that might arise using the traditional ST-GPR approach. The only study level covariate was whether a datapoint was reported as geometric or arithmetic mean. Given the distribution of environmental measurements like radon tends to be skewed, the geometric mean is the preferred measurement. As such, measurements of the arithmetic mean were crosswalked during the linear regression. Uncertainty was extracted as measurement error from the data inputs and propagated through the modelling during the GPR stage. The final estimates of burden uncertainty also incorporate the reported uncertainty of the relative risk.

Theoretical minimum-risk exposure level

The TMREL was also taken directly from literature values that were not updated for GBD2015. Given that radon is naturally occurring, zero exposure would be impossible. As such, we continue to use a TMREL of 10 Bq/m³, which is equivalent to the outdoor concentration of radon³.

Relative Risks

The relative risk for radon exposure was extracted from literature values, a 2005 meta-analysis of case-control studies showing the association of radon with lung cancer². This value was used in GBD2010 and was not updated for GBD2013 or GBD2015.

References

Lead Exposure Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case definition
Exposure to lead is defined in two different ways according to the currently known pathways of health loss. Acute lead exposure, relevant to disease burden in children, is measured as the micrograms of lead per deciliter of blood (µg/dL). Long-term lead exposure, relevant to disease burden in adults given the manifestation of health impact through increased systolic blood pressure and hence a decline of cardiovascular health, is measured as the accumulation of lead in the bone as micrograms of lead per gram of bone (µg/g).

Input data
The input data for lead exposure is derived from values extracted from literature regarding blood lead. Typically, these values are produced by studies that take blood samples and analyze them using various techniques to determine the level of lead present. The blood lead database for GBD2010 was augmented with an updated literature review for the years 2008-2013. This combined approach yielded 1,573 usable
data points from 332 different studies, which spanned the years 1964 to 2013. More than 400 new data points were added, including 337 for children and 102 country-years. The update for children is particularly relevant since blood lead impacts child IQ. The database of literature values was modelled for data-sparse countries using spatio-temporal GPR (ST-GPR). These values were used as blood lead exposure. The second pathway of burden is related to bone lead exposure, which was estimated by calculating a cumulative blood lead index for cohorts using estimated blood lead over their lifetime. The cumulative blood lead index is then used to estimate bone lead using a scalar defined by the literature.

Modelling Strategy
There methodology to estimate lead exposure last underwent significant change in GBD2013. A literature review was conducted to update the exposure dataset, to include new studies and those missed by previous reviews. Global exposure was previously modelled using age-integrating Bayesian hierarchal modelling (DisMod-MR). The modelling process was previously updated for GBD 2013 by shifting to spatial-temporal GPR methodology. This allowed for estimates of all country-age-sex-year groups for single years instead of five year periods. This approach improved the granularity of estimates for bone lead, which requires back-estimation of previous blood lead to calculate a cumulative blood lead index.

For GBD2015, the spatial-temporal GPR modelling methodology was updated as detailed in the appendix specific to this analytical technique, which is common to a variety of risk factors. In order to predict blood lead in country-years with insufficient data, covariates that have been produced across the time and space relevant to this analysis were used. For blood lead exposure, the covariates determined to have predictive ability were lag distributed income per capita (in log) and a binary covariate indicating whether lead in gasoline had been phased out for that country-year. ST-GPR was used to produce estimates of blood lead for all age groups, for both sexes, and for all GBD countries from 1970 to 2015. Next, to calculate blood lead over the lifetime of a given cohort, blood lead was assumed to grow linearly from 2.0 ug/dL in 1920 (see TMREL) to the value for that cohort in 1970. Using that database of blood lead over time and space, cohorts were constructed such that the lifetime blood lead could be expressed as a curve over each year of their life. The area under this curve was the cumulative blood lead index, which could be used to estimate bone lead in a given year with the aforementioned scalar.

Theoretical minimum-risk exposure level
The TMREL is taken from literature estimates of pre-industrial blood lead in humans. This value is estimated at 2.0 ug/dL. The decision was made that the TMREL of blood lead could not be 0 given the ambient sources of lead that would be impossible to eliminate.

Relative Risks
The blood lead relative risks were taken from a 2005 pooled analysis that was updated for GBD2010. The bone lead relative risks were taken from a 2008 meta-analysis that was updated for GBD2010. Neither of these effect sizes were modified for GBD2015.
References

Occupational Risk Factors

Flowchart

Risk factor estimation

Exposure

- Economically Active Population (EAP)
- % working in economic activity, by sex
- % working in economic activity, by sex (ILO)
- Spatio-temporal Gaussian process regression
- Exposure by risk, age, sex, year, and geography

Relative Risk

- Meta-analysis/Relative risk by economic activity
- Risk attributable to each risk by age, sex, and geography

Theoretical minimum-risk exposure level

- Published literature
- Theoretical minimum-risk exposure level (No exposure)

Final burden estimation

Deaths, YLLs, YLDs, DALYs attributable to each risk by age, sex, year, and geography

Metaphor: Database

Legend

- Input data
- Database
- Results
- Process

- Occupational Risk Factors
  - (except asbestos and injuries)

- Exposures by risk, age, sex, year, and geography

- Populations attributable to each risk by age, sex, year, and geography

- Study-level covariates
  - Log LDI
  - Education per capita
  - Urbanicity

- Total injuries by cause (GBD 2015)

- Occupational injuries by location/industry/year/sex/age

- Occupation-specific covariates
  - Log LDI
  - Education per capita
  - Urbanicity

- Spatio-temporal Gaussian process regression

- Relative Risk
  - Meta-analysis/Relative risk by economic activity
  - Risk attributable to each risk by age, sex, and geography

- Theoretical minimum-risk exposure level

- Published literature
- Theoretical minimum-risk exposure level (No exposure)
Input Data and Methodological Summary

Exposure
Case Definition

The following definitions were used for occupational risk factor exposures. All exposures were estimated only for ages 15+

<table>
<thead>
<tr>
<th>Occupational Risk Factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occupational Asbestos</td>
<td>Cumulative exposure to occupational asbestos using mesothelioma death rate as an analogue.</td>
</tr>
<tr>
<td>Occupational Asthmagens</td>
<td>Proportion of working population exposed to asthmagens based on distribution of the population in seven occupational groups</td>
</tr>
<tr>
<td>Occupational Carcinogens (arsenic, acid, benzene, beryllium, cadmium, chromium, diesel, formaldehyde, nickel, polycyclic aromatic hydrocarbons, second-hand smoke, silica, trichloroethylene)</td>
<td>Proportion of working population ever exposed to carcinogens in high or low exposures groups, based on distribution of the population in nine economic activity groups</td>
</tr>
<tr>
<td>Occupational Injuries</td>
<td>Proportion of fatal injuries attributed to occupational work in nine economic activities, based on fatal injury rates in those economic activities.</td>
</tr>
<tr>
<td>Occupational Ergonomic Factors</td>
<td>Proportion of working population exposed to lower back pain, based on distribution of the population in seven occupational groups.</td>
</tr>
</tbody>
</table>
Occupational Noise
Proportion of working population exposed to 85+ decibels of noise, based on distribution in nine economic activities.

Occupational Particulates
Proportion of working population exposed based on distribution in nine economic activities

Estimates of the proportion of population involved in economic activities and occupations were coded into the following categories:

<table>
<thead>
<tr>
<th>Economic Activities</th>
<th>Occupations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, hunting, forestry and fishing</td>
<td>Agriculture, animal husbandry, forestry workers, fishermen, and hunters</td>
</tr>
<tr>
<td>Mining and quarrying</td>
<td>Production, transport equipment operators and laborers, and related workers</td>
</tr>
<tr>
<td>Wholesale and retail trade, restaurants, and hotels</td>
<td>Professional, technical, and related workers</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>Sales workers</td>
</tr>
<tr>
<td>Electricity, gas, and water</td>
<td>Administrative and managerial workers</td>
</tr>
<tr>
<td>Transport, storage, and communication</td>
<td>Clerical and related workers</td>
</tr>
<tr>
<td>Construction</td>
<td>Service workers</td>
</tr>
<tr>
<td>Financing, insurance, real estate, and business services</td>
<td></td>
</tr>
<tr>
<td>Community, social, and personal services</td>
<td></td>
</tr>
</tbody>
</table>

Input data
Primary inputs were obtained from the ILO [1-4], using raw data on economic activity proportions, occupation proportions, fatal injury rates, and economically active population estimates. For different ISIC classifications, estimates were recoded to one of the nine economic activities or occupations. Subnational estimates for UK and China were added to the datasets for economic activities and occupations [5-6].

For occupational asbestos, primary inputs were obtained through GBD 2015 cause of death estimates and published studies. [7,13-14]

Modeling strategy
A spatial-temporal Gaussian process regression was used to generate estimates for all year/locations for the primary inputs (see app section 2). Parameters were chosen by maximizing out-of-sample cross-validation and minimizing RMSE. For economic activity and occupation proportions, estimates from ST-GPR were then re-scaled to sum to 1 across categories by dividing each estimate by the sum of all the estimates.

The following sections describe the modeling approaches for each occupational risk’s prevalence exposure.
Occupational carcinogens, occupational noise, occupational particulates
Prevalence of exposure to these risks was determined using the following equation:

\[
\text{Prevalence of Exposure}_{c,y,s,a,r,l} = \sum_{EA} \text{Proportion}_{EA,c,y} \times \text{EAP}_{c,y,s,a} \times \text{Exposure rate}_{EA,r,l,d}
\]

where:
- \( \text{EAP} \) = Economically active population
- \( \text{c} \) = country
- \( \text{r} \) = risk
- \( \text{EA} \) = economic activity
- \( \text{d} \) = duration
- \( \text{s} \) = sex
- \( \text{a} \) = age
- \( \text{l} \) = level of exposure
- \( \text{y} \) = year

Exposure rate was provided by expert group recommendations and literature [8-11] (see table 1). Duration was only considered for occupational carcinogens, through application of occupational turnover factors [12].

Occupational ergonomic factors and asthmagens
Prevalence of exposure to these risks was determined using the following equation:

\[
\text{Prevalence of Exposure}_{c,y,s,a,r} = \sum_{EA} \text{Proportion}_{OCC,c,y} \times \text{EAP}_{c,y,s,a}
\]

where:
- \( \text{EAP} \) = Economically active population
- \( \text{c} \) = country
- \( \text{r} \) = risk
- \( \text{OCC} \) = occupation
- \( \text{a} \) = age
- \( \text{s} \) = sex
- \( \text{y} \) = year

Occupational injuries
Occupational injury counts were estimated using the following equation:

\[
\text{Occupational fatal injuries}_{c,y,a,s} = \sum_{EA} \text{Injury rate}_{EA,c,y,a,s} \times \text{Population}_{c,y,a,s} \times \text{EAP}_{c,y,a,s} \times \text{Proportion}_{EA,c,y}
\]

where:
- \( \text{EAP} \) = Economically active population
- \( \text{c} \) = country
- \( \text{y} \) = year
- \( \text{EA} \) = economic activity
- \( \text{a} \) = age
- \( \text{s} \) = sex

Occupational asbestos
Prevalence of exposure to asbestos was estimated using the asbestos impact ratio (AIR), which is equivalent to the excess deaths due to mesothelioma observed in a population divided by excess deaths due to mesothelioma in a population heavily exposed to asbestos. Formally, this is defined using the following equation:
\[ AIR = \frac{Mort_{c,y,s} - N_{c,y,s}}{Mort'_{c,y,s} - N_{c,y,s}} \]

where:
- Mort = Mortality rate due to mesothelioma
- Mort' = Mortality rate due to mesothelioma in population highly exposed to asbestos
- N = Mortality rate due to mesothelioma in population not exposed to asbestos
- c = country
- y = year
- s = sex

Mortality rate due to mesothelioma was estimated from GBD 2015 causes of death [7]. Mortality rate due to mesothelioma in population not exposed to asbestos was calculated using the model in Lin et al. [13], while the mortality rate due to high exposure to asbestos was estimated in Goodman et al. [14]

**Theoretical minimum-risk exposure level**

For all occupational risks, with the exception of occupational asbestos, the theoretical minimum-risk exposure level was assumed to be no exposure to that risk.

**Relative risk**

Relative risks were obtained for all occupational risks by conducting a systematic review of published meta-analysis. The estimates used, as well as the associated studies, are reported by category group in appendix table 5.

**Population Attributable Fraction**

For all occupational risks, with the exception injuries outlined below, PAFs were calculated using the prevalences estimated above, using the PAF formula in appendix section 2.

**Occupational injuries PAF**

The PAF for occupational injuries was calculated using the following formula:

\[ PAF_{c,y,a,s} = \frac{Occupational\ fatal\ injuries_{c,y,a,s} - TMREL}{Fatal\ injuries_{c,y,a,s}} \]

where:
- c = country
- y = year
- a = age
- s = sex

Fatal injuries total was obtained from GBD 2015 causes of death [7].
Citations


Table 1 – Exposure rate by economic activity (per 100k workers)

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>Agriculture, Mining and Manufacturing</th>
<th>Electricity, Gas, and Water</th>
<th>Construction</th>
<th>Wholesale and Retail Trade and Restaurants and Hotels</th>
<th>Transport, Storage, and Communication</th>
<th>Financing, Insurance, Real Estate and Business Services</th>
<th>Community, Social and Personal Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arsenic</td>
<td>54</td>
<td>72</td>
<td>399</td>
<td>148</td>
<td>134</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>Asbestos</td>
<td>1,248</td>
<td>10248</td>
<td>589</td>
<td>1702</td>
<td>5203</td>
<td>292</td>
<td>684</td>
</tr>
<tr>
<td>Benzene</td>
<td>59</td>
<td>197</td>
<td>308</td>
<td>91</td>
<td>75</td>
<td>1037</td>
<td>520</td>
</tr>
<tr>
<td>Beryllium</td>
<td>-</td>
<td>55</td>
<td>207</td>
<td>70</td>
<td>4</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Cadmium</td>
<td>-</td>
<td>-</td>
<td>486</td>
<td>287</td>
<td>291</td>
<td>2</td>
<td>65</td>
</tr>
<tr>
<td>Chromium VI</td>
<td>-</td>
<td>346</td>
<td>2061</td>
<td>409</td>
<td>237</td>
<td>17</td>
<td>369</td>
</tr>
<tr>
<td>Diesel engine exhaust</td>
<td>646</td>
<td>21970</td>
<td>1192</td>
<td>3359</td>
<td>5816</td>
<td>485</td>
<td>13432</td>
</tr>
<tr>
<td>Second-hand smoke</td>
<td>2,082</td>
<td>163</td>
<td>5249</td>
<td>6172</td>
<td>4830</td>
<td>9278</td>
<td>6965</td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>186</td>
<td>255</td>
<td>2103</td>
<td>28</td>
<td>545</td>
<td>53</td>
<td>23</td>
</tr>
<tr>
<td>Nickel</td>
<td>-</td>
<td>2025</td>
<td>1663</td>
<td>352</td>
<td>47</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Polycyclic aromatic hydrocarbons</td>
<td>-</td>
<td>1021</td>
<td>1650</td>
<td>3066</td>
<td>1328</td>
<td>106</td>
<td>905</td>
</tr>
<tr>
<td>Silica</td>
<td>372</td>
<td>23049</td>
<td>2316</td>
<td>1415</td>
<td>18860</td>
<td>17</td>
<td>476</td>
</tr>
<tr>
<td>Sulfuric acid</td>
<td>-</td>
<td>366</td>
<td>1488</td>
<td>928</td>
<td>577</td>
<td>264</td>
<td>255</td>
</tr>
<tr>
<td>Noise, 90+ dB, high exposure</td>
<td>26100</td>
<td>57200</td>
<td>23300</td>
<td>27400</td>
<td>36200</td>
<td>100</td>
<td>18000</td>
</tr>
<tr>
<td>Noise, 85-90 dB, high exposure</td>
<td>16700</td>
<td>25400</td>
<td>32200</td>
<td>13800</td>
<td>21000</td>
<td>23100</td>
<td>28700</td>
</tr>
<tr>
<td>Noise, 90+ dB, low exposure</td>
<td>18000</td>
<td>39300</td>
<td>10600</td>
<td>20400</td>
<td>25100</td>
<td>0</td>
<td>7900</td>
</tr>
<tr>
<td>Noise, 85-90 dB, low exposure</td>
<td>14400</td>
<td>29400</td>
<td>24500</td>
<td>12300</td>
<td>19400</td>
<td>1800</td>
<td>20200</td>
</tr>
<tr>
<td>Particulates, developed, high exposure</td>
<td>10000</td>
<td>10000</td>
<td>10000</td>
<td>10000</td>
<td>10000</td>
<td>0</td>
<td>10000</td>
</tr>
<tr>
<td>Particulates, developed, low exposure</td>
<td>5000</td>
<td>7000</td>
<td>7000</td>
<td>5000</td>
<td>7000</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Particulates, developing, high exposure</td>
<td>10000</td>
<td>40000</td>
<td>40000</td>
<td>10000</td>
<td>40000</td>
<td>0</td>
<td>10000</td>
</tr>
<tr>
<td>Particulates, developing, low exposure</td>
<td>70000</td>
<td>40000</td>
<td>40000</td>
<td>70000</td>
<td>40000</td>
<td>10000</td>
<td>70000</td>
</tr>
<tr>
<td>Confidential. Do not cite or circulate.</td>
<td>77</td>
<td>77</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Suboptimal Breastfeeding Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure
Definition
Exposure to suboptimal breastfeeding is composed of two distinct categories: nonexclusive breastfeeding and discontinued breastfeeding. Non-exclusive breastfeeding is defined as the proportion of children under 6 months who are not exclusively breastfed. Those not exclusively breastfed are then parsed into three categories – predominate, partial, and no breastfeeding. Discontinued breastfeeding is defined as the proportion of children between 6 to 23 months who receive no breast milk.

Input data
The data used in this analysis consists mostly of processed micro data from surveys and tabulated data from scientific literature and reports. The data was primarily sourced from the micro data of surveys. The data updates were focused on the extraction of the larger surveys at the subnational level, especially for those subnational locations added into GBD 2015. Tabulated data was only used when micro data was not available.
Modeling
A complete time series from 1980 to 2015 for the prevalence of breastfeeding patterns for children 0 to 6 months and 6 to 23 months were generated. This was accomplished by carrying the processed micro and tabulated data through a three-step modeling process. First, a robust linear regression incorporating the covariates of log-transformed lag-distributed income, total fertility rate, and the mean years of education of women of reproductive age. This is followed by a spatial-temporal regression that uses the residuals of the predictions from the linear regression to perform a locally-weighted regression that provides a greater weighting factor to those nearer in space and time. The predicted residuals from this step are added to those created in the linear regression. The final of the three steps is the Gaussian Process Regression. This step incorporates the variance of the input data as well as that of the model predictions. It uses predictions from the spatial-temporal regression as the mean function and generates draws from a multinomial distribution, based on the data uncertainty in the prior, to generate the final prevalence estimates and their confidence intervals.

Relative risks
Relative risks used for suboptimal breastfeeding are generated based on two published meta-analyses. Non-exclusive breastfeeding exposure was paired with diarrhea and LRI as disease outcomes. Discontinued breastfeeding was paired with diarrhea only. The outcomes URI and otitis media that were included by analogy to LRI for this round of GBD.

In the case of developed regions, there is assumed to be no risk of diarrheal diseases. We have also applied a novel adjustment to the existing relative risks in order to make them representative to their larger GBD age groups (post neo-natal in the case of nonexclusive breastfeeding and 1 to 4 years in the case of discontinued breastfeeding.

Theoretical minimum-risk exposure level
For non-exclusive breastfeeding, those children that receive no source of nourishment other than breastmilk are considered to be at the lowest risk of any of the disease outcomes. For discontinued breastfeeding, we assume that children aged 6 to 23 months who receive any breastmilk as a source of nourishment to be at the lowest risk of disease outcome.
Input Data & Methodological Summary

Exposure

Case Definition
The exposure of childhood undernutrition was modeled by evaluating three anthropometric indicators which include underweight, wasting, and stunting. The definition of the three indicators are as follows:

**Childhood underweight:** Proportion of children aged 0 to 59 months in a given population who fall below 2 standard deviations (SD) of the WHO 2006 standard weight-for-age (wfa) curve. (1)

**Childhood stunting:** Proportion of children aged 0 to 59 months in a given population who fall below 2 standard deviations (SD) of the WHO 2006 height-for-age (hfa) curve.

**Childhood wasting:** Proportion of children aged 0 to 59 months in a given population who fall below 2 standard deviations (SD) of the WHO 2006 weight-for-height (wfh) curve.

Input data
There are two main inputs in the GBD 2015 undernutrition database—survey dataset and tabulated dataset. Survey dataset includes the standard multi-country or country-specific survey series such as: Reproductive and Health Surveys (RHS), Multiple Indicator Cluster Surveys (MICS), Demographic and Health Surveys (DHS), Living Standards Measurement Surveys (LSMS), China Health and Nutrition Survey (CHNS) and etc. In the absence of survey data we used tabulated data from survey reports or published literature that have been extracted at IHME, downloaded from external databases or obtained from personal communication with external collaborators. The last update for tabulated dataset was

---

1. [Olofin et al., 2013]
conducted for GBD 2010. Tabulated data include survey reports or published literature from databases from UNICEF(2), the United Nations (UN) Statistics Division (3), and the WHO Global Database on Child Growth and Malnutrition(4).

Tabulated data based on the National Center for Health Statistics (NCHS)/WHO international growth reference (the NCHS reference) (5) were converted into data based on the World Health Organization (WHO) Child Growth Standards (the WHO 2006 standard) using WHO algorithms (6). Estimates that were not representative of all children under the age of 5 were adjusted based on age groups.

**Modeling strategy**

**Exposure Estimate**

To generate a complete time series of prevalence of childhood underweight, wasting, and stunting, we employed a three-step ST-GPR modeling strategy that uses linear regression, spatiotemporal regression and Gaussian Process Regression (GPR) which is specified in the main text of this manuscript. Identical strategies and covariates were used for each undernutrition indicator. A variety of combinations of socioeconomic and environmental covariates in different transformation format were tested by running mixed-effect models with exposure data to decide the inclusion and exclusion. The final list of covariates included in the childhood undernutrition models are mean years of education of women of reproductive age, log transformed lagged-distributed income and total caloric availability (kcal per capita), which remained the same as GBD 2013. Uncertainty in the estimates was based on the data variance, then calculated through ST-GPR.

The final step of exposure estimate is to calculate the distribution of undernutrition prevalence across different levels of severity and age-sex groups. The levels of severity are defined as follows:

- **Severe**: individuals less than 3SD below the median (<-3SD);
- **Moderate**: individuals between 3SD and 2SD below the median (-3SD to -2SD);
- **Mild**: individuals between 2SD and 1SD below the median (-2SD to -1SD).

In GBD 2013, prevalence of undernutrition in each of severity categories was predicted by applying a linear regression model of the prevalence of undernutrition in each of severity categories against the prevalence of undernutrition below -2SD of the reference median at global level using microdata from 179 DHS surveys. We assumed no difference in the prevalence of undernutrition at any severity level across age and sex among children under 5.

This strategy has experienced a major change in GBD 2015. We estimated the prevalence of undernutrition by GBD age-sex groups, assuming the distribution of undernutrition of different severity categories are difference across age and sex among children under 5. Using available microdata, we first created a pooled global database that consisted of binary indicators of undernutrition by GBD age-sex groups at individual level. Then we ran a logit regression model to predict the proportion of
undernutrition outcome in most-detailed severity category (e.g. <-3SD) among the broader severity category (e.g. <-2SD) against the effects of age group and sex. We also took into account the covariance of the proportions among different age-sex groups by using variance-covariance matrix. Last, we applied the proportions by GBD age-sex group generated above onto our GPR estimates.

Theoretical minimum-risk exposure level

Theoretical minimum risk exposure levels (TMREL) for underweight, stunting, and wasting where all children under the age of 5 are above -1SD of the WHO 2006 standard weight-for-age, height-for-age, and weight-for-height curves respectively.

Relative risks

Relative risks (RRs) of risk-outcome pairs were extracted based on a study that conducted a pooled cohort analysis (7), which remained the same as GBD 2013. The final list of outcomes paired with childhood undernutrition risks included lower respiratory infections (LRI), diarrhea, measles, and protein energy malnutrition (PEM). Originally in GBD 2013, URI and otitis media were considered as analogies for LRI considering the similar pathological pathways they share. However, they were dropped from analysis in GBD 2015 due to the lack of evidence on the causal relationships with undernutrition risks. We also attributed 100% of PEM to childhood wasting and underweight but not stunting. A literature search was conducted for GBD 2015 searching for meta-analysis on the association of risk-outcome pairs published after January 1st, 2013, no updated results was found.

The RRs were adjusted using an optimization algorithm we developed at IHME for GBD 2013 that takes into account covariance between the three undernutrition indicators.
Iron Deficiency Risk Factors Appendix

Flowchart

Iron deficiency

Input Data and Methodological Summary

Exposure

Definition
To estimate anemia in GBD 2015, we employed the same method used in GBD 2013 and largely similar to GBD 2010. Our analytic strategy began with calculation of an anemia envelope – a determination of mean hemoglobin, as well as sum total of anemia prevalence, by severity for each country, age group, and both sexes for each year from 1990 through 2015. The envelope approach avoids double counting while capturing potentially different disease profiles within each population group. We defined a population group as a specific geography, sex, age-group, and year.

Input data
Iron-deficiency anemia (IDA) estimates include acute and chronic hemorrhagic states for which supplementation may be helpful, but poor nutritional intake is not the only underlying problem. A few causes in this category – hookworm, schistosomiasis, upper gastrointestinal bleeding, and gynecologic diseases – were considered separately from IDA because there was enough data from GBD prevalence estimation processes to do so. Distribution of anemia burden to IDA only after assignment to “known” causes avoided double counting of these cases.  

1. Confidential. Do not cite or circulate.

2. Confidential. Do not cite or circulate.

83
Modeling strategy
We estimated the mean hemoglobin in g/dL among pregnant women aged 15 to 49 years of age and the implied mean hemoglobin among pregnant women in the absence of iron deficiency anemia, as the risk exposure for maternal iron deficiency anemia.

Theoretical minimum-risk exposure level
The implied mean hemoglobin in the absence of iron deficiency anemia is the theoretical minimum risk exposure level. This was calculated by adding the iron deficiency shift back onto the observed hemoglobin concentration for each demographic. For example, if the observed hemoglobin concentration among 30-34 year old pregnant women in Ethiopia was 132.9 g/L, and the shift was 1.6 g/L in that demographic, then the counteractual was 134.5 g/L. The GBD 2015 anemia modeling strategy provides details on how the iron deficiency shifts were calculated.

Relative risk
We attribute 100% of iron-deficiency anemia to iron deficiency. The other outcomes are maternal hemorrhage and maternal sepsis and other maternal infections. Sources of evidence for these relative risks are unchanged from GBD 2013.

References
Vitamin A Deficiency Capstone Appendix

Flowchart

Exposure

Definition
For GBD 2015, vitamin A deficiency is defined as serum retinol <70 µmol/L. We examined vitamin A deficiency as a risk factor in children aged 6 months to 5 years.

Input data
For GBD 2015, we used data from the WHO Vitamin and Mineral Nutrition Information System, Demographic and Health Surveys, and studies identified through literature review. A systematic review was conducted for GBD 2013.

The PubMed search terms were: ((vitamin A deficiency>Title/Abstract) AND prevalence>Title/Abstract)) AND (“2009”[Date – Publication] : “2013”[Date – Publication])

The exclusion criteria were:
1. Studies that were not population-based, e.g., hospital or clinic-based studies
2. Studies that did not provide primary data on epidemiological parameters, e.g. commentaries
3. Review articles

Input Data & Methodological Summary

Exposure

Definition
For GBD 2015, vitamin A deficiency is defined as serum retinol <70 µmol/L. We examined vitamin A deficiency as a risk factor in children aged 6 months to 5 years.

Input data
For GBD 2015, we used data from the WHO Vitamin and Mineral Nutrition Information System, Demographic and Health Surveys, and studies identified through literature review. A systematic review was conducted for GBD 2013.

The PubMed search terms were: ((vitamin A deficiency>Title/Abstract) AND prevalence>Title/Abstract)) AND (“2009”[Date – Publication] : “2013”[Date – Publication])

The exclusion criteria were:
1. Studies that were not population-based, e.g., hospital or clinic-based studies
2. Studies that did not provide primary data on epidemiological parameters, e.g. commentaries
3. Review articles
4. Case series
5. Self-reported cases

Updates to systematic reviews are performed on an ongoing schedule across all GBD causes and risk factors, an update for vitamin A deficiency will be performed in the next 1-2 iterations.

Modeling Strategy

We used DisMod MR-2.0 to model prevalence of vitamin A deficiency. We used a study level covariate to indicate national and subnational observations, where nationally representative studies were set as the reference category. We used vitamin A supplementation coverage and malnutrition proportion as location-level covariates to inform variation over year and geography, especially in location-years with no or sparse data. We have made no substantive changes in the modeling strategy from GBD 2013.

Theoretical minimum-risk exposure level

The theoretical minimum risk exposure is that the prevalence of vitamin A deficiency is zero.

Relative risks

The relative risks have not changed from GBD 2013.
**Zinc Deficiency Capstone Appendix**

**Flowchart**

**Input Data & Modelling Strategy**

**Exposure**

**Case Definition**

Exposure to zinc deficiency is a measured total absorbed zinc which is a function of both zinc and phytate consumption.

**Input data**

The Food and Agriculture Organization’s (FAO) Food Balance Sheets are used to determine the total absorbed zinc per person for each country-year that they publish.

**Modeling strategy**

For GBD 2015, first, available zinc and phytate in each country-year were calculated using FAO’s Food Balance Sheets. The availability of each of these nutrients was determined using composition indices provided by our expert group. We extract phytate as well as zinc due to its functioning to inhibit zinc absorption. Then, using an equation defined by literature, the average total absorbed zinc was estimated based on the ratio of zinc to phytate in available foods. A normal distribution with a standard deviation of .25 was assumed to estimate the proportion of each population that would fall below the recommended zinc intake. Then, a complete time series from 1980 to 2015 for the proportion zinc deficient children 1 to 5 years was generated. This was accomplished by the FAO-based data through a three-step modeling process. First, a robust linear regression incorporating the covariates of log-transformed lag-distributed income as well as the proportion of malnourished individuals for each country-year. This is followed by a spatiotemporal regression that uses the residuals of the predictions from the linear regression to perform a locally-weighted regression that provides a greater weighting factor to those nearer in space and time. The predicted residuals from this step are added to those created in the linear regression. The final of the three steps is the Gaussian Process Regression. This step incorporates the variance of the input data as well as that of the model predictions. It uses predictions from the spatial-temporal regression as the
mean function and generates draws from a multinomial distribution, based on the data uncertainty in the prior, to generate the final prevalence estimates and their confidence intervals.

**Theoretical minimum-risk exposure level**
The theoretical minimum-risk exposure level for proportion zinc deficient is zero percent deficient.

**Relative risks**
Relative risks used for zinc deficiency is based on the results of clinical trials that measured the effect of zinc supplementation that were adjusted for background zinc estimates that come from the GBD estimation process.
Input Data & Methodological Summary

Exposure

Case definition

We used the Smoking Impact Ratio (SIR) for modeling burden attributable to smoking for cancers, chronic obstructive pulmonary disease (COPD), interstitial lung disease, other chronic respiratory diseases, and pneumoconiosis. SIR is the population lung cancer mortality in excess of lung cancer mortality among never-smokers, relative to excess lung-cancer mortality observed in a known reference group of smokers. Currently, SIR is adjusted to account for differences in baseline never-smoker lung cancer mortality across geography, age, and sex, but not for differences across time.

We used 5-year lagged smoking prevalence, for modeling burden attributable to smoking for cardiovascular diseases, TB, diabetes, lower respiratory infections, asthma, cataracts, macular degeneration, fractures, rheumatoid arthritis, and peptic ulcer disease. Smoking is a dichotomous exposure defined as current daily use of smoked tobacco.

A full list of outcomes included in GBD 2015 and their exposure definition is available in the table below.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atrial fibrillation and flutter</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Aortic aneurysm</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Hypertensive heart disease</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Disease</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Other cardiovascular and circulatory diseases</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Hemorrhagic stroke</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Ischemic stroke</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Diabetes</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Lower respiratory infections</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Asthma</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Peptic ulcer disease*</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Rheumatoid arthritis*</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Cataract*</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Macular degeneration*</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Hip fracture*</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Non-hip fracture*</td>
<td>5-year lagged smoking prevalence</td>
</tr>
<tr>
<td>Bladder cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Colon and rectum cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Esophageal cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Kidney cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Leukemia</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Liver cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Tracheal, bronchus, and lung cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Lip and oral cavity cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Nasopharynx cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Stomach cancer</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Larynx cancer*</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Interstitial lung disease and pulmonary sarcoidosis</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Other chronic respiratory diseases</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
<tr>
<td>Pneumoconiosis</td>
<td>Smoking Impact Ratio (SIR)</td>
</tr>
</tbody>
</table>

* New outcome in GBD 2015

**Input data**

Consistent with GBD 2013, we used nationally representative survey data to estimate smoking prevalence. Survey and report data identified in the Global Health Data Exchange (GHDx), the WHO InfoBase, and the International Smoking Statistics (ISS) Database.

**Inclusion Criteria**

- Nationally representative
- Report current use of any of the following frequency-type combinations:
  - Daily use of smoked tobacco
  - Any use (both daily and occasional) of smoked tobacco
  - Daily use of cigarettes
  - Any use (both daily and occasional) of cigarettes
  - Daily use of any tobacco (both smoked and smokeless)
  - Any use (both daily and occasional) of any tobacco (both smoked and smokeless)
  - Daily use of any tobacco excluding cigarettes
• Report data within the time period of January 1, 1980 – December 31, 2015 for any geography estimated in the GBD framework
• Smoking prevalence reported among individuals ages 10+

Global Health Data Exchange (GHDx)
Sources were identified through a systematic search of the GHDx.

- Search Terms (Keywords): Tobacco Use
- Time Period: January 1, 1980 – December 31, 2015
- Data Type: Survey OR Report
- Search Date: February 16, 2016

Out of 3,912 sources identified in the GHDx, 2,818 sources were included.

WHO InfoBase and International Smoking Statistics (ISS) Database
An effort was made to replace database-derived estimates used in GBD 2013 with original extractions from primary data sources. In GBD 2013, [851] sources were derived from the WHO InfoBase or the ISS Database. In GBD 2015, we replaced [257] sources with extractions from primary data sources and continued to use [594] sources from the WHO InfoBase (n=[281]) and the ISS Database (n=[313]).

Outliers
Throughout the modeling process, data were assessed for bias and outliers were flagged. A data point was flagged as a candidate outlier if it was not consistent with the majority of other data points in a country with respect to level, age-pattern, sex-pattern, or temporal trend. In data-scarce countries, data points were also compared to data from other countries in a region. Candidate outliers were scrutinized for potential sources of bias and were ultimately excluded if the point or source was deemed to not be representative.

Modeling strategy
Data Extraction
When possible, we extracted individual smoking status for all available frequency-type categories (listed above) from person-level microdata and collapsed these data to produce prevalence estimates in the standard GBD 5-year age-sex groups. If microdata were unavailable we extracted the most granular age-sex groups available from survey reports. Any available measures of uncertainty were extracted, including standard error, confidence or uncertainty intervals, and sample size.

Data Preparation: Crosswalking
Regressions to crosswalk other frequency-type categories to the gold-standard definition of daily use of smoked tobacco were estimated in the form:

\[ p_{\text{daily-smoked},k} = \beta_1 p_{i,k} + \epsilon_k \]

where \( p_{\text{daily-smoked},k} \) is the prevalence of daily smoking reported in survey \( k \), and \( p_{i,k} \) is the prevalence of an alternative frequency-type combination \( i \) also reported in survey \( k \). Consistent with previous GBD smoking crosswalks, the intercept was omitted from the regression. The estimated regression coefficient \( \beta_1 \) was used to crosswalk alternative frequency-type categories to the gold-standard daily smoking definition in
sources only providing the alternative category. Predication error at the data-point level was used to propagate uncertainty and was calculated using the following equation:

\[ \text{PE}_k = \sigma^2 + X_k^2 \text{var}(\hat{\beta}) \]

Compared to the separate frequency and type crosswalks used in GBD 2013, the combined frequency-type crosswalk used in GBD 2015 represents an improvement because patterns in frequency that may vary by type and patterns in type that may vary by frequency are captured.

**Data Preparation: Age and Sex Splitting**

Report data provided in age groups wider than the standard GBD 5-year age groups or as both sexes combined were split using the approach used in Ng et al. Briefly, age-sex patterns were identified using sources with data on multiple age-sex groups and these patterns were applied to split aggregated report data. Uncertainty in the age-sex split was propagated by multiplying the standard error of the data (including the predication error of the crosswalk) by the square root of the number of splits performed.

**Modeling: Linear Model**

After data preparation, the dataset consisted of prevalence estimates of daily smoked tobacco use in standard GBD country-year-age-sex groups. The mean function used in ST-GPR was estimated using the following hierarchical mixed-effects linear regression, run separately by sex:

\[
\logit(p_{c,a,t}) = \beta_0 + \beta_1 \text{CPC}_{c,t} + \sum_{k=2}^{16} \beta_k I_{A[a]} + \alpha_s + \alpha_r + \alpha_c + \epsilon_{c,a,t}
\]

where CPC\(_{c,t}\) is the annual tobacco consumption per capita covariate, \(I_{A[a]}\) is a dummy variable indicating specific age group A that the prevalence point \(p_{c,a,t}\) is capturing, and \(\alpha_s, \alpha_r, \) and \(\alpha_c\) are super region, region, and country-specific random effects.

**Modeling: Spatio-Temporal Gaussian Process Regression (ST-GPR)**

The estimated mean function was then propagated through the ST-GPR framework to obtain 1,000 draws of smoking prevalence estimates for each location, year, age, and sex. Parameter selection for the ST-GPR hyper-parameters were selected through out-of-sample cross-validation using the strategy described elsewhere in this appendix.

**Smoking Impact Ratio Estimation**

We have made no substantive changes in the SIR estimation strategy from GBD 2013. The only change in input data for estimating never-smoker lung-cancer mortality was to update data from the China Kadoorie Biobank prospective cohort to include follow-up through 2014. Country-year-age-sex specific lung cancer mortality rates are derived from GBD 2015 Cause of Death estimation and detailed in that Capstone’s appendix. The formula for calculating SIR is:

\[
\text{SIR} = \frac{C_{L_C} - N_{L_C}^*}{S_{L_C}^* - N_{L_C}^*} \times \frac{N_{L_C}^*}{N_{L_C}}
\]
Theoretical minimum-risk exposure level
The theoretical minimum-risk exposure level is that no one in the population smokes tobacco; that is, the smoking impact ratio is zero and smoking prevalence is zero.

Relative risk
We have made no substantive updates to relative risks for outcomes included in GBD 2013. The following outcomes using 5-year lagged smoking prevalence as the exposure were added in GBD 2015: peptic ulcer disease, rheumatoid arthritis, cataracts, macular degeneration, hip fracture, and non-hip fracture. Larynx cancer was the only new outcome added using SIR as the exposure. Relative risks for rheumatoid arthritis, cataracts, and macular degeneration were derived from recent published meta-analyses. We performed our own meta-analyses of prospective cohort studies to derive relative risks for peptic ulcer disease, hip fracture, and non-hip fracture. We used Kontis et al.’s re-analysis of CPS-II smokers for the relative risk of larynx cancer.
Second-hand Smoke Capstone Appendix

Flowchart

Input Data and Methodological Summary

Exposure

Case Definition
We measure exposure to any tobacco smoke inside the home among non-smokers. Ex-smokers and occasional smokers are considered non-smokers for the purposes of this analysis. Exposure was evaluated for both children and adults.

Input data
We included surveys that had at least one question about smoking status and also asked about whether respondents live with any smokers or whether their spouse smokes. For children we also used surveys that asked about parental smoking. Some main sources include Global Adult Tobacco Survey (GATS), Global Youth Tobacco Survey (GYTS), DHS, NHANES, BRFSS, Eurobarometer, etc. Updates to systematic reviews are performed on an ongoing schedule across all GBD causes and risk factors, an update for second-hand smoke will be performed in the next 1-2 iterations.

Many new surveys were added for GBD 2015, which were identified and accessed using GHDx. We cross-referenced with available sources used for smoking in order to evaluate whether these sources were also useful for second-hand smoke. Some of the big new survey series that were added included the National Adult Tobacco Survey and National Youth Tobacco survey series from the U.S., VIGITEL and Risk Factor Chronic Disease Surveillance data from Brazil, and the Chronic Disease Risk Factor Surveillance from China. All new Global Youth Tobacco Surveys (GYTS), Global Adult Tobacco Surveys (GATS), Global school-based student health surveys (GSHS) and Eurobarometer were added as well, in addition to other one-off surveys that evaluated second-hand smoke in the household.
Modeling strategy

We used the traditional PAF equation to estimate burden based on exposure and relative risks. Prevalence of secondhand smoke exposure among nonsmokers is modeled in Dismod-MR and all crosswalks/adjustments are done both within and outside of DisMod to account for alternative case definitions.

In GBD 2015, a new modelling change we implemented was to crosswalk surveys asking about spousal smoking or parental smoking (depending on adults versus children) to our gold standard of any exposure to second-hand smoke in the household by anyone. A sizable group of the DHS surveys do not ask directly about smoke exposure in the household, and thus exposure is ascertained indirectly through looking at the smoking status of each partner in the couple’s module to see if there is a “mixed-status” relationship in which one partner is exposed to the other’s smoke.

Another adjustment that we made prior to DisMod was for the act of smoking. In some surveys, such as the Global Youth Tobacco Survey, the survey only asks whether their parent smokes, not whether the child being interviewed is actively exposed to smoke on a regular basis (which we define as at least once a week). Thus, in addition to adjusting for spouse/parent versus anybody, we also adjusted for whether the survey asked the person whether they were directly exposed to smoke or just whether people smoked who lived in their home. The two-by-two table below helps illustrate the different potential combinations of alternate definitions that we adjusted for.

<table>
<thead>
<tr>
<th>Act of smoking</th>
<th>Spouse/Parent</th>
<th>Anybody</th>
</tr>
</thead>
<tbody>
<tr>
<td>Act</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Non-act</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

We used a mixed effects regression to crosswalk these alternative definitions, with interactions between anybody smoking and sex, fixed effects on act of smoking, and nested random effects at the super-region, region and country level. Previously, this crosswalk was done in DisMod.

Once we had crosswalked these alternative definitions, we modeled second-hand smoke prevalence as a single parameter prevalence model in DisMod-MR. Another modelling change that we made in GBD 2015 was to run separate models for male and female secondhand smoke exposure, with children included in the female model. This decision was made because the sex effect being estimated with the combined gender model was underestimating the sizably higher impact of second-hand smoke on women as compared to men. Thus, we decided to model them separately.

In the female model, we used with age mesh points at 0 5 10 15 18 20 30 40 50 60 80 & 100, while in the male model we used age mesh points at 0 15 18 20 30 40 50 60 80 & 100. The difference in age mesh points was due to the fact that all children were modeled as female due to similar rates of exposure, while the male model was limited to adult males greater than 15.

We use the age-standardized smoking prevalence among females as a country-level covariate in the male model, and the age-standardized smoking prevalence among males as a country-level covariate in the female model. This was a modelling change from GBD 2013, in which we only had one second-hand
smoke model and used the age-standardized smoking prevalence rate among men. In addition, we used one study level fixed effects to account for the different case definitions in our dataset:

- **Study level fixed effects on integrand value (x-cov)**
  - Prevalence figure includes exposure to tobacco smoke outdoors as well as indoors

- **Study level fixed effect on integrand variance (z-cov)**
  - Study asked about exposure to second-hand smoke at home and/or work (rather than exposure inside the home only)
  - Study was not nationally representative

All raw input CSA data points had a measure of uncertainty going into DisMod – standard error, confidence interval or effective sample size – and the uncertainty around final estimates also takes into account uncertainty from study-level covariate fixed effects on variance, as well as geographic random effects.

**Theoretical minimum-risk exposure level**
The theoretical minimum-risk exposure level for second-hand smoke is zero exposure among non-smokers to second-hand smoke in the home.

**Relative risks**
For children under 5 years of age, we estimate the burden of lower respiratory infections (LRI) and otitis media attributable to second-hand smoke exposure. For adults greater or equal to 25 years of age we estimate the burden of lung cancer, ischemic heart disease, cerebrovascular disease and lower respiratory infections (LRI) attributable to second-hand smoke exposure. For GBD 2010 all of these pooled relative risks came from published meta-analyses, but for GBD 2015 we used country-specific relative risks that were created using integrated exposure response curves (IER). The relative risk for otitis media still comes from a published meta-analysis, as opposed to the IER approach.
Alcohol Capstone Appendix

Flowchart

Input Data and Methodological Summary

Exposure

Case definition

The impact of alcohol consumption on morbidity and mortality can be largely described by two separate but related dimensions. The 1st dimension is the individual level drinking and consists of four indicators;

1. Current drinkers, defined as the proportion of individuals who have consumed at least one alcoholic beverage (or some approximation) in the last 12 months.
2. Former drinkers, defined as the proportion of individuals who have ever consumed an alcoholic beverage, but not in the last 12 months.
3. Lifetime abstainers, defined as the proportion of individuals who have never consumed an alcoholic beverage.
4. Alcohol consumption (in grams per day), defined as grams of alcohol consumed by current drinkers, per day, over a 12 month period.

The 2nd dimension of alcohol consumption relates to the pattern of drinking and consists of two indicators;

5. Binge drinkers, defined as the proportion of drinkers who have had a binge event in the past 12 months. A binge event was defined as consuming 60 grams of alcohol (approximately five drinks or more) in a single occasion for males and 48 grams of alcohol in a single occasion for females.
6. Binge times, defined as the proportion of drinking events that are binge amongst binge drinkers i.e. the proportion of days that a binger has a binge event.
**Input data**

For GBD 2013, a systematic review of the literature was conducted to capture population survey data on all six alcohol use indicators. In summary, the search was conducted in three stages involving electronic searches of the peer-reviewed literature via PubMed, the grey literature and, expert consultation. Updates to systematic reviews via PubMed are performed on an ongoing schedule across all GBD causes and risk factors, an update for alcohol use will be performed in the next 1-2 iterations. For GBD 2015, stages two and three of the literature review were conducted, prioritizing countries for which subnational estimates were generated. 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. Data-sources were included if they captured a sample representative of the geographic location under study and contained variables that could be used to formulate any of the six alcohol use indicators. Relevant survey variables from each data-source were documented in a Microsoft Excel codebook and extracted using STATA 13.1. A total of 629 potential data-sources were available in GhDx across countries with subnational locations, out of which 127 data-sources (66,108 data-points) were included across all six indicators.

To generate estimates of alcohol consumption in grams per day, data from population surveys were used in combination with estimates of per capita consumption from the Food and Agriculture Organization (FAO) [1] and the Global Information System on Alcohol and Health (GISAH database [2]) Per capita consumption is an aggregate measure of recorded, unrecorded, and tourist per capita consumption of alcohol (UNWTO database [3]) derived from sales, production, and other economic statistics. While population-based surveys provide accurate estimates of the prevalence of lifetime abstainers, former drinkers and current drinkers, they typically underestimate real alcohol consumption levels. As a result, the all-age, both-sex per capita consumption figures from the FAO and GISAH are considered 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 modeled from the population survey data and the overall volume of consumption from FAO and GISAH to determine the total amount of alcohol consumed by country.

To generate estimates of alcohol consumption in liter per capita, raw inputs were obtained from FAOSTAT [1] and WHO GISAH database [2]. To provide more stable time trends in the model, FAO sales data was transformed to a lagged 5-year average. FAO data was used when WHO data wasn’t available. Otherwise, FAO and WHO data was adjusted (crosswalked) by running a mixed effect model on the log average of the data with indicators for the FAO and WHO data series, as well as random effects on super region, region, country, and time. Each data point was adjusted by the predicted betas on super-region and region.

\[
\text{Log Average Data} = D + (\text{Super Region} | D, \text{Region} | D, \text{Country} | D, \text{Year} | D)
\]

\[
\text{Transformed data} = \text{data} \times e^{\hat{\beta}_1 + \hat{\beta}_s}
\]

Where \( D \) = Indicator variable for data source
To generate uncertainty, a Lowess model was run on the adjusted data and the standard deviation between the difference of the Lowess smoothed model and the adjusted data points was used for data points missing uncertainty.

Unrecorded consumption was incorporated into the alcohol LPC data using estimates provided by the WHO [4]. WHO estimates were only reported for the years 1990, 2005, and 2010 so for missing years, estimates were interpolated. For years outside this range, unrecorded estimates were carried forward or backwards from the closest year. Unrecorded consumption estimates were reported in liters per capita so estimates were added to adjusted data points to account for unrecorded consumption.

Tourism data was obtained through the UNWTO [4]. A crosswalk was applied across different tourist categories, similar to the one used for FAO and WHO data, to estimate tourist proportions for a given country. Tourism consumption was incorporated after modeling unadjusted alcohol LPC as outlined below.

**Modeling strategy**

DisMod-MR 2.1 was used to estimate country-, year-, age- and sex-specific proportions of current drinkers, former drinkers, lifetime abstainers, binge drinkers, and binge times; and alcohol consumption as a continuous variable in grams per day. We have made no substantive changes in the modeling strategy from GBD 2013. We ran single-parameter models for each alcohol use indicator and included a combination of location- and study-level covariates in each model. An alcohol liters per capita location-level covariate was used for all six indicators to assist in the predictive power of the models. Additionally, study-level covariates were used to accommodate for known sources of variability in the raw data. In the current drinkers, former drinkers, binge drinkers and binge times models, we included two covariates which adjusted estimates derived in the past week and past month towards those derived in the past year respectively. Estimates derived in the past year were considered to be the gold standard given the previously outlined definition for each indicator.

In the alcohol consumption model, we included a separate study-level covariate flagging data points derived from The World Health Organization’s World Health Surveys (WHS) conducted across multiple countries. There was considerable variability in estimates derived from the WHS which may have been influenced by methodological differences in how alcohol use was captured. This study-level covariate looked for unsystematic bias between data-points and added more uncertainty onto those from the WHS. If other data-points causing higher or lower modelled output were identified during the modelling process for a given indicator, the plausibility of these data points was assessed and the study methodology reviewed. Data points with methodological limitations, for instance those derived from survey items not entirely representative of the alcohol use indicators required, with small sample sizes, or derived from samples not entirely representative of the general population were excluded.

A spatial-temporal Gaussian process regression was used to model total alcohol in liters per capita (see appendix, section 2). Parameters and a random effect model for the prior were chosen using out-of-sample cross validation. This produced estimates of alcohol LPC for a complete time series for the years 1980-2015 by country.

Alcohol LPC was adjusted for each country hosting tourists using the following equations:
Alcohol LPC\textsubscript{H} = Unadjusted Alcohol LPC\textsubscript{H} + Alcohol LPC\textsubscript{Consumption abroad} − Alcohol LPC\textsubscript{Tourist consumption}

Alcohol LPC\textsubscript{Consumption abroad} =
\[\sum_{V} \text{Proportion of tourists}_{HV} \times \text{Unadjusted Alcohol LPC}_{H} \times \frac{\text{Average length of stay}_{HV} \times \text{Tourist Population}_{V}}{\text{Population}_{H}}\]

Alcohol LPC\textsubscript{Tourist consumption} =
\[\sum_{V} \text{Proportion of tourists}_{V} \times \text{Unadjusted Alcohol LPC}_{V} \times \frac{\text{Average length of stay}_{V} \times \text{Tourist Population}_{H}}{\text{Population}_{H}}\]

Where H = Host country, V = Visiting country

Or, in other words, alcohol LPC was adjusted by adding in the per capita rate of consumption abroad and subtracting the per capita rate of tourist consumption domestically.

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 modeled in Dismod for percentage of current drinkers within a location/year/sex/age, as well as consumption trends modeled in Dismod g/day by location/year/sex/age, using the following equations.

\[\text{Proportion of total consumption}_{LYSA} = \frac{\text{Alcohol g/day}_{LYSA} \times \text{Population}_{LYSA} \times \% \text{Current drinkers}_{LYSA}}{\sum_{s,a} \text{Alcohol g/day}_{LYSA} \times \text{Population}_{LYSA} \times \% \text{Current drinkers}_{LYSA}}\]

\[\text{Alcohol LPC}_{LYSA} = \frac{\text{Alcohol LPC}_{LY} \times \text{Population}_{LY} \times \text{Proportion of total consumption}_{LYSA}}{\text{Population}_{LYSA}}\]

Where L = location, Y = Year, S = Sex, A = Age

A similar scalar was applied so that total subnational consumption equaled national consumption.

**Theoretical minimum-risk exposure level**

For alcohol use, the theoretical minimum-risk exposure level (TMREL) was assumed to be no alcohol use, i.e. 0 g/day of alcohol consumption. This diverges from the definition of other theoretical minimum-risk exposure level of risks because, for some alcohol-use relative risks, there’s a preventative effect for low levels of consumption. However, due to the modeling of alcohol relative risks outlined below, it was found that 0 g/day provided the most consistency between the definition of alcohol-use TMREL and other GBD risk’s TMREL. This is an area of improvement for future GBD iterations. Current research suggests that the preventative effect noted in studies may be due to issues in estimating abstainer populations. [5-7] If this is the case, a TMREL of 0 would still be valid.
Relative Risks

Relative risks were derived for each GBD cause by mapping functions to the dose-response relationships found in meta-analysis. [11-22] Due to data availability, for high levels of consumption, uncertainty in the relative risk functions increases greatly. To minimize the uncertainty of these measures, relative risks were estimated up to the 90\textsuperscript{th} percentile of exposures in men (85 g/day) and the 95\textsuperscript{th} percentile of exposures in women (60 g/day). For exposures beyond this, the associated relative risk was carried forward from these chosen percentile exposure levels. Though a dose-response relationship is evident at higher levels of exposure, the shape of the relative risk function is highly uncertain for higher levels of exposure both due to a lack of observations at these exposure levels, as well as confounding variables affecting estimation of the relative risk of these populations. Thusly, our relative risk estimates are likely an underestimate for the top 10\% of male exposures and 5\% of female exposures. For exact relative risks used, see appendix section 4.

Population Attributable Fraction

For chronic conditions, PAF was defined as

\[ PAF(x) = \frac{P_A + P_F + \int_0^{150} P(x) \cdot RR_F(x) \, dx}{P_A + P_F + \int_0^{150} P(x) \cdot RR_C(x) \, dx} \]

where:

- \( x \) = alcohol consumption in g/day
- \( P_A \) = Prevalence of lifetime abstainers
- \( P_F \) = Prevalence of former drinkers
- \( P(x) \) = Prevalence of alcohol consumption
- \( RR_F \) = Relative risk of former drinkers
- \( RR_C(x) \) = Relative risk function for drinkers

A thousand draws were taken of PAFs to generate uncertainty. The gamma distribution was used to estimate individual level variation within drinking populations [8-9]. Binge drinkers were not taken into account for chronic causes since the pattern of drinking has not been found to be an indicator of most outcomes [10].

For non-chronic conditions, such as injuries, binge drinking was accounted for in the model since patterns of drinking is significant.

\[ PAF(x) = \frac{P_A + P_F + P_C + P_{C+B} + RR_{C+B}(x) - 1}{P_A + P_F + P_C + P_{C+B} + RR_{C+B}(x)} \]

\[ RR_{C+B}(x) = P_D \cdot P_{D+B} \cdot (RR_{crude}(x) - 1) + 1 \]

where:

- \( P_{C+B} \) = Prevalence of current drinkers who binge
- \( RR_{C+B} \) = Relative risk of current drinkers who binge
- \( P_D \) = Proportion of a day that is a binge event
RR\text{crude} = \text{Relative risk for a given mean level of consumption} \quad \text{P}_{D+B} = \text{Proportion of all days where a binge event occurs}

The estimated PAF draws were then used to estimate YLL, YLDs, and DALYs, as per the other risk factors (see appendix section 2).

References

1. Food and Agriculture Organization of the United Nations. FAOSTAT Statistics Database.


11. Roerecke M, Rehm J. Alcohol consumption and the risk for morbidity and mortality of ischemic heart disease - A systemic review and meta-analysis. Toronto, Canada: Centre for Addiction and Mental Health; 2011


Injecting Drug Use Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case definition
Injecting drug users (IDU) are at high risk from blood-borne infections, including human immunodeficiency virus (HIV) and Hepatitis B and C viruses (HBV and HCV, respectively), through the use of shared needles and injection equipment. In GBD 2010, based on the available epidemiological literature and the availability of exposure estimates\(^2,3\) we measure the burden of disease attributable to HIV, HBV and HCV due to injecting drug use. An injecting drug user was defined as a current or recent user aged 15-64 years old.

Input data
The major burden of mortality from viral hepatitis is due to cirrhosis and liver cancer resulting from chronic hepatitis infection. Cirrhosis mortality was modelled with vital registration data using CODEm. Etiologic proportion models, estimated using DisMod-MR 2.0, were used to split the overarching cirrhosis mortality estimates into cases of cirrhosis attributable to hepatitis B, hepatitis C, alcohol, and other causes.\(^1,4\)

Liver cancer mortality was modelled using cancer registry data. The incidence numbers were transformed into mortality estimates using mortality to incidence ratios. The mortality estimates from cancer registries were then combined with vital registration system data as input data into CODEm, which produced the final mortality estimates for liver cancer. As with cirrhosis mortality, etiologic proportions for liver cancer due to hepatitis B, and C, alcohol, and other causes were generated using DisMod-MR 2.0.
To estimate the burden of HIV cases attributable to IDU, we extracted data on the proportion of notified HIV cases by transmission route – sexual intercourse, injecting drug use, commercial sex work and other - from a number of agencies that conduct surveillance of HIV across the globe.(6-13) This produced 728 data points from 81 countries.

The prevalence of current injecting drug use was estimated using data and estimates from a review conducted by the Reference Group to the UN on HIV and injecting drug use(15). This review used a multistage process of systematic review adhering to international guidelines. It involved multiple stages of peer and expert review, with searches of the peer-reviewed literature in addition to an extensive review of online grey literature databases in the drug and alcohol and HIV fields. Additional data on the age and sex distribution of injecting drug use were sourced for this modelling exercise.

In order to generate a pooled incidence rate/absolute relative risk for viral hepatitis among people who inject drugs, we conducted a meta-analysis of longitudinal epidemiological studies that reported a hepatitis B (16-20) or hepatitis C(16-31) incidence rate among PWID. We calculated confidence intervals for the incidence rate (where no CI was reported) from a Poisson distribution around the number of cases.

We excluded studies that focused on non-representative subgroups, such as recent injectors or adolescents or because hepatitis incidence is far higher in those groups than for all people who inject drugs (e.g.(32)). We did not vary incidence among active injectors according the availability of blood borne virus prevention strategies (e.g. NSPs, opioid substitution therapy) because too few studies have examined different levels of incidence according to variable coverage, and we were not able to estimate coverage by country over time. In any case, in most countries, coverage of virus prevention strategies remains very low among people who inject drugs,(33) and would have been negligible in most countries until recent years.

Modeling strategy
As part of the GBD 2013 study, we measured the burden of hepatitis B and hepatitis C (including attributable cirrhosis and liver cancer) and HIV at the country, regional, and global level for each age-sex group for the years 1990 to 2013. For HIV, hepatitis B and hepatitis C, disease-specific natural history models were used to estimate deaths and YLDs, because the three-state model in DisMod-MR 2.0 (susceptible, cases, dead) did not capture the complexity of the disease processes.

Mortality estimation
Mortality due to overall acute hepatitis was modelled with vital registration data using the Cause of Death Ensemble Modelling tool (CODEm), an analytical tool that tests the predictive power of hundreds of models to estimate trends in causes of death.(5) Due to poor coverage of cause of death data for each of the acute hepatitis varieties, four natural history models for hepatitis B and C were used to estimate mortality by deriving incidence from measurements of seroprevalence and then multiplying incidence by case fatality to estimate the number of deaths. These four models were then squeezed so as to fit the parent cause of death model.

We estimated HIV mortality using a modified UNAIDS Spectrum model.(2) This is a compartmental HIV progression model estimates age-specific incidence, prevalence and death rates using methods described elsewhere.(2) This modelling approach was adapted according to epidemic type, including concentrated and generalized epidemics. For concentrated epidemics, the Spectrum models were corrected for misclassification of HIV deaths and then calibrated to align with vital registration data. For generalised HIV
epidemics, we minimised a loss function to select epidemic curves that were most consistent with the prevalence and all-cause mortality data.(2)

*Estimation of Years Lived with Disability*

For non-fatal estimation, we estimated the incidence of hepatitis B and C using seroprevalence data in DisMod-MR 2.0. For both hepatitis B and C, we use data on the seroprevalence of the hepatitis surface antigen (a marker of chronic infection in hepatitis B and a marker of ever-infection in hepatitis C), excess mortality, and remission, to estimate incidence of both hepatitis infections. Incidence of cirrhosis was also estimated in DisMod using cirrhosis hospital data and cause-specific mortality rate (CSMR) data.

Incidence of liver cancer was derived by dividing mortality by the mortality to incidence ratios, which were then used to predict liver cancer survival. Finally, we estimated prevalence as a function of incidence and survival by splitting prevalence into four phases. Each phase had different disability weights, which were used to generate YLDs for that phase.

Finally, incidence of HIV was also estimated using the UNAIDS Spectrum modelling approach described above in the mortality estimation section.

*Burden of HIV attributable to injecting drug use*

We then estimated the proportion of HIV cases attributable to three transmission categories (sex, IDU and other) for all country-time periods using DisMod-MR 2.0. The only covariate used in the model was one that added variance to the data points derived from data sources that attributed a portion of HIV cases to “unknown” transmission sources. We scaled the proportions from each of the three transmission models (sex, IDU and other) to ensure that they fit the total HIV transmission envelope by country, year, age and sex.

*Burden of hepatitis B and hepatitis C attributable to injecting drug use*

To estimate the relative contribution of IDU to hepatitis B and C disease burden at the country, regional and global level, we used a cohort method. We re-calibrated individuals according to history of injecting drug use, and their accumulated risk of incident hepatitis B and C due to IDU. We made use of data on prevalence of current injecting drug use, pooled in DisMod-MR 2.0; a meta-analysis of incidence rates of hepatitis B and hepatitis C among people who inject drugs; and estimates of population-level incidence of hepatitis B and C between 1990 and 2013. We used back extrapolations to estimate incidence before 1990. These steps are detailed below.

To estimate the lifetime risk of being infected with hepatitis B or C, we undertook a cohort analysis for each country, year, age, and sex category and estimated the probability of an individual having been infected in each preceding year. One of the main inputs to this cohort method was the probability of having injected drugs in a specific age cohort in a given calendar year. For example, for a cohort of 40-year-olds in 2015, the relevant probability in 2005 is the estimated prevalence of injecting drug use among 30-year-olds.

In addition to a global time series of estimated prevalence of injecting drug use, we also used the incidence of hepatitis B or C and the sero-conversion rate of hepatitis B and hepatitis C among people who inject drugs for each age-sex-country-year from 1960 to 2013 by 5-year age groups.

1. Incidence rate of Hepatitis B and C in the general population
We modelled the annual incidence rate of hepatitis B and hepatitis C using sero-prevalence data in DisMod-MR 2.0. We assumed a low remission (mean 0.015 and standard error 0.0075)(14) in the hepatitis B model to reflect the small proportion of cases who spontaneously clear the infection. We assumed zero remission for hepatitis C.

2. Prevalence of ever-injecting drug use
DisMod-MR 2.0 was used to estimate the prevalence of injecting drug use with year as a covariate to estimate the trends over time. DisMod makes an average estimate of the change in drug use over the time period from 1990-2015 and we took draws from a normal distribution of the coefficient to project IDU prevalence backward in time to 1960 from baseline level in 1990.

3. Pooled seroconversion hazard of hepatitis C and hepatitis B among people who ever injected drugs
This pooled sero-conversion hazard for both hepatitis C and hepatitis B was derived from a meta-analysis of longitudinal epidemiologic studies described above in the input data section.

**Theoretical minimum-risk exposure level**
The theoretical minimum-risk exposure level is defined as zero exposure to injecting drug use.

**Relative risks**
For drug use, there were not substantial changes made to the effect sizes from GBD 2013. We used a pooled absolute risk of Hepatitis C and Hepatitis B among those who have ever used injecting drugs.

In addition to assessing IDU as a risk factor for blood-borne infections, the broader category of mental and substance use disorders is assessed as risk factors for suicide. The suicide burden attributable to mental and substance use disorders is estimated by comparing the current health status with a theoretical-minimum-risk exposure defined as the counterfactual status of the absence of mental and substance use disorders (Ferrari, Norman et al 2014).

**References**


Dietary Risks Capstone Appendix

Flowchart

Input data & Methodological summary

Exposure

Case definition

For GBD 2015, risk factors associated with diet include: diet low in fruits, vegetables, whole grains, nuts and seeds, fiber, seafood omega-3 fatty acids, polyunsaturated fatty acids, calcium; and diet high in red meat, processed meat, sugar sweetened beverages, trans fatty acids, and sodium. Exposure to diet low in fruits is defined as average daily consumption of less than 250 grams per day of fruits (fresh, frozen, cooked, canned, or dried, excluding fruit juices and salted or pickled fruits). Exposure to diet low in vegetables is defined as average daily consumption of less than 420 grams per day of vegetables (fresh, frozen, cooked, canned or dried vegetables including legumes but excluding salted or pickled vegetables, juices, nuts and seeds, and starchy vegetables such as potatoes or corn). Exposure to diet low in whole grains is defined as average daily consumption of less than 125 grams per day of whole grains (bran, germ, and endosperm in their natural proportion) from breakfast cereals, bread, rice, pasta, biscuits, muffins, tortillas, pancakes and other sources. Exposure to diet low in nuts and seeds is defined as average daily consumption of less than 20 grams per day of nuts and seeds. Exposure to diet low in milk is defined as average daily consumption of less than 435 grams per day of milk including non-fat, low-fat, and full-fat milk, excluding soy milk and other plant derivatives. Exposure to diet low in calcium is defined as average daily consumption of less than 1.25 grams per day of calcium from all sources, including milk,
yogurt, and cheese. Exposure to diet low in fiber is defined as average daily consumption of less 23 grams per day of than fiber from all sources including fruits, vegetables, grains, legumes and pulses. Exposure to diet low in seafood omega-3 fatty acids is defined as average daily consumption of less than 250 milligrams per day of eicosapentaenoic acid and docosahexaenoic acid. Exposure to diet low in polyunsaturated fatty acids is defined as average daily consumption of less than 11% of total energy intake from polyunsaturated fatty acids as a replacement for high intake of saturated fatty acids (> 7% of total energy intake). Exposure to diet high in red meat is defined as average daily consumption of greater than 23 grams per day of red meat (beef, pork, lamb, and goat but excluding poultry, fish, eggs, and all processed meats). Exposure to diet high in processed meat is defined as average daily consumption of greater than 2 grams of meat preserved by smoking, curing, salting, or addition of chemical preservatives. Exposure to diet high in sugar sweetened beverages is defined as average daily consumption of greater than 2.5 grams per day of beverages with ≥50 kcal per 226.8 gram serving, including carbonated beverages, sodas, energy drinks, fruit drinks, but excluding 100% fruit and vegetable juices. Exposure to diet high in trans fatty acids is defined as average daily consumption of greater than 0.5% of trans fat from all sources, mainly partially hydrogenated vegetable oils and ruminant products. Exposure to diet high in sodium is defined as average 24 hour urinary sodium greater than 3 grams per day.

Input data

We used dietary data from multiple sources including nationally and sub-nationally representative nutrition surveys, household budget surveys, and United Nations FAO Food Balance Sheets and Supply and Utilization Accounts. Additionally, for sodium and trans fatty acids, we used data on 24-hour urinary sodium and availability of partially hydrogenated vegetable oil in packaged foods, respectively. All dietary data (other than sodium and sugar-sweetened beverages) were standardized to 2000 kcal/day. We modelled missing country-year data from FAO using a space-time Gaussian process regression and lag-distributed country income as the covariate. For each dietary factor, we estimated the global age pattern of consumption based on nutrition surveys (i.e., 24-hour diet recall) and applied that age pattern to the FAO data. Substantive changes in input data compared to GBD 2013 are as follows: (a) using data from United Nations Supply and Utilization Accounts to estimate the intakes of fiber, calcium, seafood omega-3 fatty acids, polyunsaturated fatty acids, and saturated fatty acids; (b) using data from United Nations FAO Food Balance Sheets to estimate the intake of fruits; (c) excluding data from United Nations FAO Food Balance Sheets in estimating the whole grain intake.

Modeling strategy

We used DisMod-MR 2.1 to estimate the intake of each dietary factor by age, sex, country, and year. In GBD 2015, for all dietary factors other than sodium, we considered data from 24-hour diet recall as the gold standard, and cross-walked other methods of assessment to the gold standard method. For sodium, the 24-hour urinary sodium was considered as the gold standard. To estimate the 24-hour urinary sodium based on dietary sodium, we performed a crosswalk between these two types of data in a subset of countries with sodium data from both urinary and dietary surveys.
Table 1 summarizes the study-level and country-level covariates used in modeling of each dietary factor.

Table 1. Covariates used in modeling of each dietary factor.

<table>
<thead>
<tr>
<th>Diet factor</th>
<th>Sex</th>
<th>Suboptimal metric</th>
<th>Nationally Representativeness</th>
<th>Data from FFQ¹</th>
<th>Data from HBS²</th>
<th>Data from FAO</th>
<th>Country level covariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diet low in fruits</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet low in vegetables</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet low in whole grains</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet low in nuts and seeds</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet low in milk</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet high in red meat</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>National availability of red meat (grams/person/day)</td>
</tr>
<tr>
<td>Diet high in processed meat</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>National availability of pig meat (% of energy/person/day)</td>
</tr>
<tr>
<td>Diet high in sugar-sweetened beverages</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>National availability of sugar (Kcal/person/day)</td>
</tr>
<tr>
<td>Diet low in fiber</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet suboptimal in calcium</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet low in seafood omega-3 fatty acids</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>Landlocked nation (Yes,/No)</td>
</tr>
<tr>
<td>Diet low in polyunsaturated fatty acids</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet high in trans fatty acids</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>National availability of hydrogenated oil (% of energy/person/day)</td>
</tr>
<tr>
<td>Diet high in sodium</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
</tbody>
</table>

¹ Food Frequency Questionnaire
² Household Budge Survey

To characterize the distribution of each dietary factor at population level, we used the following equation to model the relationship between the standard deviation and mean of intake in nationally representative nutrition surveys using multiple 24-hour diet recalls:

\[
\ln (\text{Standard deviation}) = \beta_0 + \beta_1 \times \ln (\text{Mean}) + \beta_{\text{risk}} \times I_{\text{risk}}
\]
Then we applied the coefficients of this regression to the outputs of DisMod-MR 2.1 to calculate the standard deviation of intake by age, sex, year, and country.

**Theoretical minimum-risk exposure level**

In GBD 2015, to estimate the TMREL for each dietary factor, we first calculated the level of intake associated with the lowest risk of mortality from each disease endpoint based on the studies included in the meta-analyses of the dietary relative risks. Then, we calculated the TMREL as the weighted average of these numbers using the global number of deaths from each of outcome as the weight (Table 2).

<table>
<thead>
<tr>
<th>Dietary Factor</th>
<th>GBD 2013</th>
<th>GBD 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fruits</td>
<td>200-400 gr/day</td>
<td>200-300 gr/day</td>
</tr>
<tr>
<td>Vegetables</td>
<td>350-450 gr/day</td>
<td>340-500 gr/day</td>
</tr>
<tr>
<td>Whole grains</td>
<td>100-150 gr/day</td>
<td>100-150 gr/day</td>
</tr>
<tr>
<td>Nuts</td>
<td>12-20 gr/day</td>
<td>16-25 gr/day</td>
</tr>
<tr>
<td>Red meats</td>
<td>11.4-17.1 gr/day</td>
<td>18-27 gr/day</td>
</tr>
<tr>
<td>Processed meats</td>
<td>0-14.3 gr/day</td>
<td>0-4 gr/day</td>
</tr>
<tr>
<td>Milk</td>
<td>425-475 gr/day</td>
<td>350-520 gr/day</td>
</tr>
<tr>
<td>Sugar sweetened beverages</td>
<td>0-64.3 gr/day</td>
<td>0-5 gr/day</td>
</tr>
<tr>
<td>Polyunsaturated fatty acids</td>
<td>10-15% of total daily energy</td>
<td>9-13% of total daily energy</td>
</tr>
<tr>
<td>Seafood omega-3 fatty acids</td>
<td>200-300 mg/day</td>
<td>200-300 mg/day</td>
</tr>
<tr>
<td>Trans fatty acids</td>
<td>0-0.8% of total daily energy</td>
<td>0-1%E</td>
</tr>
<tr>
<td>Dietary fiber</td>
<td>28-32 gr/day</td>
<td>19-28 gr/day</td>
</tr>
<tr>
<td>Dietary calcium</td>
<td>1.0-1.3 gr/day</td>
<td>1-1.5 gr/day</td>
</tr>
</tbody>
</table>

**Relative Risk**

We obtained the relative risk of each disease endpoint per serving of the dietary components from the most recent dose-response meta-analyses of prospective observational studies, and where available randomized controlled trials. In GBD 2015, we specifically updated the relative risks for the following risk outcome pairs: diet low in fruits-ischemic heart disease; diet low in fruits-ischemic stroke; diet low in fruits-hemorrhagic stroke; diet low in vegetables-ischemic heart disease; diet low in vegetables -ischemic
stroke; diet low in vegetables-hemorrhagic stroke; diet low in whole grains-ischemic heart disease; diet low in whole grains-ischemic stroke; diet low in whole grains- hemorrhagic stroke; and diet low in fiber-ischemic heart disease. We also included diabetes as an outcome for diet low in fruits based on the evidence from a most recent meta-analysis of prospective observational studies. Considering the well-established age trend of the relative risks of metabolic risk factors for cardiovascular disease and diabetes, we conducted a literature review to identify the most important metabolic mediators for each dietary factor and used the age trend of the relative risk of that mediator(s) and the disease endpoint to estimate the age-specific relative risk for each dietary factors (Table 3).

Table 3. Metabolic mediators used to determine the age trend of the effect of dietary factors on cardiometabolic outcomes.

<table>
<thead>
<tr>
<th>Body Mass Index</th>
<th>Total Serum Cholesterol</th>
<th>Fasting Plasma Glucose</th>
<th>Systolic Blood Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diet low in fruits</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Diet low in vegetables</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Diet low in whole grains</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Diet low in nuts and seeds</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Diet high in red meats</td>
<td>●</td>
<td>-</td>
<td>●</td>
</tr>
<tr>
<td>Diet high in processed meats</td>
<td>●</td>
<td>-</td>
<td>●</td>
</tr>
<tr>
<td>Diet low in fiber</td>
<td>-</td>
<td>●</td>
<td>-</td>
</tr>
<tr>
<td>Diet low in seafood omega-3 fatty acids</td>
<td>●</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Diet low in polyunsaturated fatty acids</td>
<td>-</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Diet high in trans fatty acids</td>
<td>●</td>
<td>●</td>
<td>-</td>
</tr>
</tbody>
</table>
Childhood Sexual Abuse Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case Definition
The case definition for childhood sexual abuse (CSA) is ever having had the experience of any contact abuse (i.e. fondling and other sexual touching) or intercourse when aged 15 years or younger, and the perpetrator or partner was older than the victim.

Input data
Currently, we use self-reported survey data to measure CSA prevalence, not data from Child Protection Services (CPS) or other crime data. The reliability and comprehensiveness of CPS and crime statistics varies too much geographically to warrant including it.

An updated systematic review of CSA prevalence literature was conducted for sources published between January 2011 and September 2015. The following search terms were used:

((("health surveys"[MeSH Terms] AND prevalence[Title/Abstract]) OR ("sentinel surveillance"[MeSH Terms] AND prevalence[Title/Abstract]) OR ("prevalence"[Title/Abstract] AND cross sectional studies[MeSH Terms])) AND (("child abuse"[MeSH Terms] OR "child abuse, sexual"[MeSH Terms]) OR ("sex offenses"[MeSH Terms] OR "child abuse, sexual"[MeSH Terms]) OR (child*[Title/Abstract] AND sexual[Title/Abstract] AND abuse[Title/Abstract]))) NOT ("comment"[Publication Type] OR "letter"[Publication Type] OR "editorial"[Publication Type]))
Intimate Partner Violence Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case Definition

The case definition for intimate partner violence (IPV) is ever experienced one or more acts of physical and/or sexual violence by a current or former intimate partner since the age of 15 years. Estimated in females only because IPV is more common in females and there is more evidence quantifying the associated risk for health outcomes.

- Physical violence is defined as: being slapped or having something thrown at you that could hurt you, being pushed or shoved, being hit with a fist or something else that could hurt, being kicked, dragged, or beaten up, being choked or burnt on purpose, and/or being threatened with or actually having a gun, knife, or other weapon used on you.
- Sexual violence is defined as: being physically forced to have intercourse when you did not want to, having sexual intercourse because you were afraid of what your partner might do, and/or being forced to do something that you found humiliating or degrading (the definition of humiliating and degrading may vary across studies depending on the regional and cultural setting).
- Intimate partner is defined as: a partner to whom you are married or with whom you cohabit. In countries where people date, dating partners will also be considered (a partner with whom you have an intimate (sexual) relationship with but are not married to or cohabiting).
Input data
A systematic review of the intimate partner violence prevalence literature was conducted in Pubmed for anything published between November 2014 and February 2016. The following search terms were used to conduct the systematic review:

(("health surveys"[MeSH Terms] AND prevalence[Title/Abstract]) OR ("sentinel surveillance"[MeSH Terms] AND prevalence[Title/Abstract]) OR ("prevalence"[Title/Abstract] AND cross sectional studies[MeSH Terms])) AND (abuse, sexual[MeSH Terms] OR domestic violence[MeSH Terms] OR abuse, partner[MeSH Terms] OR abuse, spousal[MeSH Terms] OR rape[MeSH Terms]) NOT ("comment"[Publication Type] OR "letter"[Publication Type] OR "editorial"[Publication Type])

This query produced 92 results, and of these, 33 data points were extracted for 13 different countries. In addition to literature, we supplemented this data with surveys tagged with “intimate partner violence” in the GHDx. Some of the big survey series that were updated or newly added include: all new Demographic and Health surveys, the National Youth Risk Behavior Survey, the Gender, Alcohol and Culture International Study (GENACIS), the CDC Reproductive Health Surveys, Mexican National Addiction Survey, USA Collaborative Psychiatric Epidemiology Surveys, and the Brazil National Alcohol and Drug Survey.

We get the proportion of solved homicides that were perpetrated by an intimate partner from crime statistics and police reports. For GBD 2013, the main source of these crime statistics and police reports came from an IPV-homicide systematic review in the Lancet in 2013.

In GBD 2015, an updated systematic review was done for IPV homicide sources in PubMed through April 2016. The query used for this Pubmed search was:

((IPV[All Fields] OR ("intimate partner violence"[MeSH Terms] OR ("intimate"[All Fields] AND "partner"[All Fields] AND "violence"[All Fields]) OR "intimate partner violence"[All Fields])) AND ("homicide"[MeSH Terms] OR "homicide"[All Fields]) OR femicide[All Fields])) AND ("2013/01/01"[PDAT] : "3000/12/31"[PDAT])

These literature sources were supplemented with sources from the GHDx that were tagged with Intimate partner violence AND Homicide.

Modeling strategy
For GBD 2015, we use three distinct approaches to estimate burden attributable to IPV, including 1) the traditional exposure and relative risk to PAF method for depression, suicide and abortion; 2) the direct PAF approach for estimating the proportion of homicides that are perpetrated by an intimate partner; and 3) a cumulative risk approach for estimating the burden of HIV/AIDS attributable to IPV.

Estimating attributable burden to IPV for depression, suicide and abortion
Before upload to DisMod, we first adjust data with variable recall periods (previous 12 months versus lifetime), type of violence (sexual, physical, or both) and severity (severe only versus all levels). To convert data to our gold standard definition of ever having experienced any IPV, we use data from the WHO multi-country violence against women surveys to construct crosswalk regressions. The dependent variable in each of these regression was ever any IPV (gold standard), while the key independent variable was one of the 11 alternative metrics of IPV that were represented in our dataset:
1. Physical IPV in the past 12 months
2. Sexual IPV in the past 12 months
3. Severe IPV in the past 12 months
4. Severe physical IPV in the past 12 months
5. Severe sexual IPV in the past 12 months
6. Any IPV (physical and/or sexual) in the past 12 months
7. Ever any physical IPV
8. Ever any sexual IPV
9. Ever any severe IPV
10. Ever severe physical IPV
11. Ever severe sexual IPV

For alternate metrics 1-6 there is likely to be a relationship between current exposure and age. For these metrics we included a series of age dummies:

\[
\text{logit}(GSait) = \beta + \beta_1 \text{logit}(ALTait) + \beta_2 i + \varepsilon
\]

For alternate metrics 7-11, we ran the following regression:

\[
\text{logit}(GSit) = \beta_0 + \beta_1 \text{logit}(ALTit) + \varepsilon
\]

where GS refers to the gold standard metric of IPV prevalence, ALT is the alternate metric of IPV prevalence, \(i_a\) refers to the complete set of age-group indicators, \(a\) refers to an age-group, \(i\) refers to a country, and \(t\) refers to year. We included age-group indicators in the first six regressions because we expected the prevalence of recent IPV to vary by age. Using the intercepts, coefficients, and variance-covariance matrix from each of these eleven regressions, we were able to convert all of the alternate metrics of IPV prevalence in our dataset to estimates of “ever any IPV”. We eliminated observations based on alternate metrics of IPV which came from studies that also provided estimates of IPV based on the gold standard definition (i.e. duplicates).

After applying crosswalks to the alternate metrics of IPV in the manner described above, we made an additional adjustment to the subset of our data that was based on only ever-partnered, currently partnered women currently married women or ever married women. To adjust these values so that they reflected IPV prevalence in the entire female population, regardless of partnered status, we multiplied estimates from these studies by the age-specific fraction of women who had ever been partnered.

An updated time series was generated in GBD 2015 using MICS and DHS data in a single parameter DisMod model to reflect the most recent data on proportion of women that have ever been partnered. This revised time series was used to adjust values for surveys with restricted partner status to reflect the prevalence among all women in the population.

After these pre-DisMod crosswalks and adjustments, a single-parameter prevalence model was run in DisMod with age mesh points at 0 14 15 20 30 40 50 60 80 & 100. A study-level covariate fixed effect (x-cov) was used to adjust data points where the survey question used to calculate prevalence only asked about violence perpetrated by the woman’s spouse. A study-level fixed effect on integrand variance (z-
cov) to indicate whether a study was nationally representative or not was used to account for the heterogeneity introduced by studies that are not generalizable to the entire population.

We tried using alcohol liters per capita, prevalence of binge drinking, and prevalence of male binge drinking in the GBD 2015 model as national-level fixed effects, but they were not significant so they were ultimately dropped.

**Direct PAF for female homicides**

The burden of homicides attributable to intimate partner violence is modeled as a direct PAF.

Input data all fed into a single-parameter proportion DisMod model, which has age mesh points at 0 10 20 45 & 100. The model has a study-level covariate fixed effect on integrand value (x-cov) for sources just including police reported homicides. We also included a study-level fixed effect on integrant variance (z-cov) to indicate whether a study was nationally representative or not.

In GBD 2015, we added prevalence of binge drinking to the model as a country-level covariate.

**Cumulative risk approach for PAF of HIV/AIDS due to IPV**

The third and final modelling approach that we used to assess burden attributable to intimate partner violence was a cumulative risk approach to measure the burden of HIV/AIDS attributable to IPV.

The approach itself remained the same in GBD 2015, but included updated intimate partner violence exposure numbers from the DisMod model described above, as well as revised HIV incidence numbers.

From the literature we have information on the incidence rate ratio (IRR) of HIV incidence from two cohort studies (Jewkes et al, Lancet 2010 & Kouyoumdjian, et al AIDS 2013). As we measure burden based on deaths and prevalence, we need to be able to quantify attributable fractions on prevalence and death rather than incidence. To get a PAF on prevalence we need to consider the history of exposure to IPV and the accumulated associated risk of incident HIV due to IPV, relative to the overall risk of HIV at the population level. The ratio of cumulative IPV-attributable HIV incidence to total HIV incidence is an approximation of the relevant PAF on HIV prevalence and we will assume this PAF can also applied to mortality.

\[
\frac{Cumulative\ HIV\ incidence\ due\ to\ IPV}{Cumulative\ HIV\ incidence\ overall} = \frac{1 - \prod_{a=0}^{n} (1 - PAF_{ay} \times I_{ay})}{1 - \prod_{a=0}^{n} (1 - I_{ay})}
\]

Where:

1 = annual incidence rate of HIV

\(a\) = age (15-84)

\(y\) = year (1980-2013)

\[PAF_{HIV\ incidence} = \frac{[Prevalence\ of\ IPV]_{ay} \times (IRR-1)}{[Prevalence\ of\ IPV]_{ay} \times (IRR-1)+1}\]
Theoretical minimum-risk exposure level
The theoretical minimum-risk exposure level is zero exposure to intimate partner violence, as defined above.

Relative risks
We estimate burden attributable to IPV for abortion, depression, suicide, interpersonal violence (i.e. homicide) and HIV incidence. We have added HIV as an outcome for GBD 2013 in response to bolstered causal evidence from a second prospective study published in 2013 (Kouyoumdjian, 2013). We use a pooled incidence rate ratio (IRR) of 1.59 (95% CI 1.3-1.94) from a meta-analysis of the two available prospective studies as of date.

The relative risks for depression and suicide come from a systematic review of longitudinal studies assessing intimate partner violence and incident depressive symptoms and suicide attempts. For the relative risk for IPV-abortion, we ran a custom meta-analysis in GBD 2013 that we continued to use in GBD 2015. An important methodological note with IPV-abortion is that we must apply the pooled relative risk for abortion to the current prevalence of IPV (in the previous 12 months), rather than lifetime prevalence. This is because the relevant exposure for abortion would be recent IPV, and because the case definition for all but one of the RR component studies was physical or sexual IPV in the past year.
Unsafe Sex Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case Definition
For GBD 2015, unsafe sex is defined as the risk of disease due to sexual transmission.

Input data
To be used in our models, sources must report HIV cases attributable to various modes of transmission. We cannot use data on the prevalence of HIV in the population in general or among specific populations like drug users or CSW. We screened all UNAIDS country progress reports and searched government epidemiological surveillance records for these data. The primary data sources we used were UNAIDS, the European CDC, and the US CDC.

For GBD 2015, we extracted all new European CDC, UNAIDS, and US CDC reports that had been published since the previous iteration of GBD. We also extracted state-level HIV surveillance reports where available. These were found through the US CDC: National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Division of HIV/AIDS Prevention’s website.

Barring the time for a full systematic review, these ECDC, US CDC and UNAIDS reports are the main sources for breakdown of HIV transmission. Updates to systematic reviews are performed on an ongoing schedule across all GBD causes and risk factors, an update for unsafe sex will be performed in the next 1-2 iterations.

We excluded all extractions where the “other” category for HIV transmissions accounted for greater than 25 percent of all cases. We believe that this indicates issues with the reporting system used to report HIV cases in that location.
Modeling strategy
There were no substantial changes in the modelling approach for unsafe sex from GBD 2013. We model the proportion of HIV cases attributable to unsafe sex for GBD. To do this we collect and clean data, run three DisMod models (HIV attributable to sex, HIV attributable to IDU, HIV attributable to other routes of transmission), squeeze the results of the three DisMod models, and prepare PAFs. Additionally, we run a fourth DisMod model that is the proportion of sexually transmitted HIV cases that are due to commercial sex work (CSW) (defined as in sex workers vs in the 2nd or 3rd contacts of sex workers), that is used in the intimate partner violence calculations.

We attribute burden to unsafe sex for all ages 0-100, for both sexes, and all GBD locations for years 1990 to 2015.

All of the DisMod models include a study-level covariate fixed effect on integrand variance (x-cov) for sources that include unknown cases in their “other” category. We assumed that the inclusion of unknown cases in the other category would impact the uncertainty around the point estimates. We used age mesh points 0 and 100 for all models, since almost all of the data was for large age ranges, rather than age-specific. No study-level x-covs or country level covariates were included in the models.

All raw input unsafe sex data points had a measure of uncertainty going into DisMod – standard error, confidence interval or effective sample size – and the uncertainty around final estimates also takes into account uncertainty from study-level covariate fixed effects on variance, as well as geographic random effects.

After the 3 main HIV transmission models (sexual, IDU, other) are run, the results of all 3 must be squeezed so they sum to 100% for a given country-year-age-sex group.

Theoretical minimum-risk exposure level
The theoretical minimum level used for unsafe sex is the absence of disease transmission due to sexual contact.

Population attributable fraction calculations
The outcomes associated with unsafe sex that we report on include HIV, cervical cancer, and all sexually transmitted diseases (STDs) except for those in neonates from vertical transmission, including HIV, ophthalmia neonatorum and neonatal syphilis.

Based on evidence in the literature, we attribute 100% of cervical cancer to unsafe sex. These sources state that HPV infection is necessary for cervical cancer to develop and that HPV is spread through sexual contact. The proportion of STDs attributable to unsafe sex is also 100%.

For HIV, the results from the single parameter proportion DisMod model for HIV transmission due to sex are used directly as the population attributable fraction.
Low Physical Activity Capstone Appendix

Flowchart

Input Data and Methodological Summary

Exposure

Case Definition
We measure physical activity performed by adults greater than or equal to 25 years of age, for durations of at least ten minutes at a time, across all domains of life (leisure/recreation, work/household and transport). We use frequency, duration and intensity of activity to calculate total metabolic equivalent-minutes per week. MET (Metabolic Equivalent) is the ratio of the working metabolic rate to the resting metabolic rate. One MET is equivalent to 1 kcal/kg/hour and is equal to the energy cost of sitting quietly. A MET is also defined as the oxygen uptake in ml/kg/min with one MET equal to the oxygen cost of sitting quietly, around 3.5 ml/kl/min.

Input data
We included surveys of the general adult population that captured self-reported physical activity in all domains of life (leisure/recreation, work/household and transport), where random sampling was used.

Data were primarily derived from two standardized questionnaires: The Global Physical Activity Questionnaire (GPAQ) and the International Physical Activity Questionnaire (IPAQ), although we included any other survey instrument that asked about intensity, frequency and duration of physical activities performed across all activity domains.

Due to a lack of a consistent relationship on the individual level between activity performed in each domain and total activity, we were not able to use studies that included only recreational/leisure activities.

Physical activity level is categorized by total MET-minutes per week using four categories based on rounded values closest to the quartiles of the global distribution of total MET-minutes/week. The lower limit for the Level 1 category (600 MET-min/week) is the recommended minimum amount of physical
activity to get any health benefit. We used four categories with higher thresholds rather than the GPAQ and IPAQ recommended 3 categories to better capture any additional protective effects from higher activity levels.

- Level 0: < 600 MET-min/week (inactive)
- Level 1: 600-3999 MET-min/week (low-active)
- Level 2: 4000-7,999 MET-min/week (moderately-active)
- Level 3: ≤ 8,000 MET-min/week (highly active)

The GHDx was used to locate all surveys that use the GPAQ or IPAQ questionnaire. Although there were many other surveys that focused specifically on leisure activity, we were unable to use these sources because they did not comprise all three domains (work, transport and leisure). In addition, we excluded any surveys that did not report frequency, duration, and intensity of activity.

**Modeling strategy**

**Pre-DisMod crosswalks**

We conducted two crosswalks prior to DisMod to adjust the raw data to our “gold standard” definition. In GBD 2010, our gold standard definition was GPAQ, but for GBD 2013 and into 2015, we shifted to IPAQ due to concern that GPAQ was not accurately capturing “domestic” (house/yard) activities and was thus greatly underestimating activity level. In an empirical comparison between the World Health Survey (IPAQ) and the WHO Study on global AGEing and adult health (SAGE) (GPAQ) showed significantly lower activity levels assessed using GPAQ as compared to IPAQ for females in low income countries.

We calculated an adjustment factor to apply to GPAQ surveys for females only (since the difference between questionnaire activity level estimates were not significantly different for men). A regression was fitted on data from nationally representative surveys that used either GPAQ or IPAQ for each activity category, where the dependent variable was the logit of the proportion in the relevant activity level and the main independent variable was an interaction between super region and survey (1=GPAQ, 0=IPAQ), with fixed effects for age categories and a country level random effect.

We also adjusted non-nationally-representative urban and rural data points. We constructed an urbanicity covariate that is equal to 1 for urban data points, 0 for rural data points and the proportion urban for the country for nationally representative data points. The dependent variable was the logit of the proportion in the relevant activity level and the main independent variable is an interaction between sex and urbanicity, with fixed effects for age categories and a country level random effect.

A new adjustment that we implemented in GBD 2015 prior to DisMod was to shift data points from the Behavioral Risk Factor Surveillance System (BRFSS), which asks about recreation/leisure, transport and household chores, but does not ask about activity on the job. We used data from the National Health and Nutrition Examination Survey (NHANES) to create age-sex level average proportions in each activity category, which we used to adjust the state-level US BRFSS estimates.

**DisMod modeling**
Once the raw data had been adjusted to meet our gold standard definition of physical activity, we modeled activity as a single parameter proportion model in DisMod. We estimated the proportion of each country/year/age/sex subpopulation in each of the above four activity levels using six separate Dismod models. We use six models rather than four to accommodate the different MET-minute/week cutoffs presented in tabulated data sources where individual unit record data was not available. Since the accepted threshold/definition for inactivity is consistently <600 MET-minutes/week, the vast majority of tabulated data was broken down into proportion inactive (model A) and proportion low, moderate or highly active (model B).

<table>
<thead>
<tr>
<th>Label</th>
<th>MET-min/week</th>
<th>Name of sequelae in online visualization tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>A inactive</td>
<td>&lt;600</td>
<td>Physical inactivity and low physical activity, inactive</td>
</tr>
<tr>
<td>B low/moderately/highly active</td>
<td>≥600</td>
<td>Physical inactivity and low physical activity, low/moderately/highly active</td>
</tr>
<tr>
<td>C low active</td>
<td>600-3999</td>
<td>Physical inactivity and low physical activity, low active</td>
</tr>
<tr>
<td>D moderately/highly active</td>
<td>&gt;4000</td>
<td>Physical inactivity and low physical activity, moderately/highly active</td>
</tr>
<tr>
<td>E moderately active</td>
<td>4000-7999</td>
<td>Physical inactivity and low physical activity, moderately active</td>
</tr>
<tr>
<td>F highly active</td>
<td>≥8,000</td>
<td>Physical inactivity and low physical activity, highly active</td>
</tr>
</tbody>
</table>

These models have mesh points at 0 15 25 35 45 55 65 100, and a study-level fixed effect on integrand variance (Z-cov) for whether a study was nationally representative or not, to account for the heterogeneity introduced by studies that are not generalizable to the entire population. They also have a national level fixed effects on mean BMI.

After DisMod, we rescale these 6 models so that the proportions sum to one. Since we have the most data for models A and B, we rescale the sum of the proportion in each category to be equal to one. Next we rescale the sum of model C and D to be equal to the rescaled value from model B. Then we rescale the sum of models E and F to be equal to the rescaled value from model D. After these three rescales we are left with a proportion for each of the four categories that all sum to 1.

We have made few substantive changes in the modeling strategy from GBD 2013.

**Theoretical minimum-risk exposure level**

The theoretical minimum-risk exposure level for physical inactivity is 600 MET-min per week, which is the WHO recommended level for physical activity for health.³

**Relative risks**

We estimate burden attributable to physical inactivity for breast cancer, colon cancer, diabetes, ischemic heart disease and stroke. A systematic review of relevant epidemiological literature was conducted for each health outcome up to February 27, 2016. Due to considerable heterogeneity in the literature with respect to physical activity metrics and domain(s) covered, a methodologically intensive strategy was required to standardize the relative risk units to match those of the exposures.⁴
In addition to updates to the literature review, the main other change that has been implemented since GBD 2013 is that we previously used a non-increasing prior in the Bayesian meta-regression to estimate the association between PA and each of the five outcomes. For GBD 2015, we no longer use any prior in the Bayesian meta-regression analysis.

References


High Fasting Plasma Glucose Capstone Appendix

Flowchart

High fasting plasma glucose

Input Data & Methodological Summary

Exposure

Case Definition
We measure fasting plasma glucose as a continuous exposure in units of mmol/L and define diabetes according to the American Diabetes Association (ADA) and World Health Organization (WHO) diagnostic guidelines as FPG > 7.0 mmol/L and/or currently taking diabetes.1,2

Input Data
Consistent with GBD 2013, we utilized data on mean systolic blood pressure from literature and from household survey microdata and reports (e.g. STEPS, NHANES). Please see appendix for a full list of included sources. In GBD 2013, a systematic review of the literature was completed to capture population survey data on mean systolic blood pressure. For GBD 2015, we updated the systematic review using the same strategy, drawing from the GHDx and Medline via PubMed. In total, we have utilized 717 sources corresponding to 24,926 unique data points.

Global Health Data Exchange Database
We systematically searched the Global Health Data Exchange (GHDx) for multi-country survey programs, national surveys, and longitudinal studies which provide measured individual level data on systolic blood
pressure. The search was completed for systolic blood pressure, fasting plasma glucose, and blood cholesterol simultaneously, as many sources studying the other metabolic risks will often report mean fasting plasma glucose or diabetes prevalence.

**Search Terms (Keywords):** Blood pressure OR Blood glucose OR Glucose tests OR Cholesterol OR Cholesterol tests OR Hypercholesterolemia

**Data Type:** Survey OR Report

**Search date:** 2/6/2016

**Literature Review**
We systematically searched PubMed for articles published between 15 July 2009 and 31 December 2015 which provided national or subnational estimates of mean systolic blood pressure. As above, the literature review was completed for systolic blood pressure, fasting plasma glucose, and blood cholesterol simultaneously for the reasons previously stated.

Search terms:

```plaintext
```

**Search date:** 1/26/2016

**Expert Groups**
To capture any remaining sources not identified in the GHDx or in PubMed, we looked to other leaders in the field to ensure our datasets were as comprehensive as possible. These included the IDF Atlas Database and a recent publication on diabetes from the NCD Risk Factor Collaboration. 3,4

**Inclusion Criteria**
Studies were included if they were population-based and measured glucose using a blood test (as FPG, HbA1c). We accepted data on diabetes prevalence only if the study performed an objective blood measurement and/or individuals reported self-report of taking anti-diabetic medication. Studies that included self-report of diabetes were excluded. We assumed the data is representative of the location if the geography was not related to the diseases (a mining area) and if it is not an outlier compared to other data in the country or region.

**Outliers**
Data was utilized in the modeling process unless an assessment of data showed that the data is biased. A data point was considered to be an outlier candidate if the level is not consistent with other (sources)
country data, or - if there are no other data points - not consistent with other country in the region. A candidate outlier source was scrutinized and validated and the data point was excluded if the quality of study did not warrant a valid estimate because of selection (specific populations), different definitions, other biases, or if the study did not provide methodological details for evaluation.

Data Extraction
Where possible, individual level data on fasting plasma glucose was extracted from survey microdata and these were collapsed across demographic groupings to produce mean estimates in the standard GBD 5-year age-sex groups. If microdata were unavailable, information from survey reports or from literature were extracted along with any available measure of uncertainty including standard error, uncertainty intervals, and sample size.

Survey reports and literature often only report information on diabetes prevalence in the population studied. If the study was otherwise representative, we extracted data on the prevalence of diabetes and, using all available data with both estimates of mean fasting plasma glucose and prevalence of diabetes, crosswalked this to estimates of mean fasting plasma glucose.

Crosswalk from Prevalence of Diabetes and HbA1c
We used a mixed-effects regression to crosswalk estimates of diabetes prevalence to the mean fasting plasma glucose of a given population. A separate regression was run for a given diagnostic criteria using the form:

\[
\log(FPG_{c,a,s,t,k}) = \beta_0 + \beta_1 \logit(p_{c,a,s,t,k}) + \beta_2 \text{male} + \sum_{k=10}^{21} \beta_h I_{A[a]} + \alpha_s + \epsilon_{c,a,s,t,k}
\]

Where \(FPG_{c,a,s,t,k}\) is the outcome of interest—the mean fasting plasma glucose of a given country-, age-, sex-, time-, from survey \(k\); \(p_{c,a,s,t,k}\) is the prevalence of diabetes for a given definition or the mean HbA1c level; \(I_{A[a]}\) is a dummy variable indicating a specific age group \(A\); and \(\alpha_s\) is a super-region specific random effect.

Age and Sex Splitting
Prior to modeling, data provided in age groups wider than the GBD 5-year age groups were split using the approach outlined in Ng et al.\(^5\) Briefly, age-sex patterns were identified using sources of data with multiple age-sex groups and these patterns were applied to split aggregated report data. Uncertainty in the age-sex split was propagated by multiplying the standard error of the data performed by the square root of the number of splits performed.

Modeling
Exposure estimates were produced from 1980 to 2015 for each national and subnational location, sex, and for each 5-year age group starting from 25+. As in GBD 2013, we used a Spatio-Temporal Gaussian Process Regression (ST-GPR) framework to model the mean systolic blood pressure at the location-, year-, age-, sex- level. Updates to the ST-GR modeling framework for GBD 2015 are detailed in the appendix.

The FPG mean function was estimated using a mixed-effects linear regression, run separately by sex:
\[
\logit(F\text{PG}_{c,a,t}) = \beta_0 + \beta_1 SD\text{S}_{c,t} + \beta_2 P_{\text{overweight}_{c,a,t}} + \beta_2 \log(sugar_{c,t}) + \sum_{k=2}^{16} \beta_k I_{A[a]} + \alpha_s + \alpha_r + \alpha_c + \varepsilon_{c,a,t}
\]

where SDS\text{S}_{c,t} is socio-demographic status (SDS), an index metric that includes a measure of education and income, \( P_{\text{overweight}_{c,a,t}} \) is the prevalence of overweight, sugar\text{S}_{c,t} is the diet adjusted mean consumption of sugar in grams per capita per day, \( I_{A[a]} \) is a dummy variable for a fixed effect on a given 5-year age group, and \( \alpha_s, \alpha_r, \alpha_c \) are random effects at the super-region, region, and country level, respectively.

The estimates were then propagated through the ST-GPR framework to obtain 1000 draws for each location, year, age, and sex.

**Theoretical minimum-risk exposure level**

As in GBD 2013, the theoretical minimum risk exposure level for fasting plasma glucose is between 4.9 and 5.3 mmol/l (uniformly distributed) with a standard deviation 0.3mmol/l. This SD is the lowest reported in population data, after correction for the effects of one-time measurement. We used the same TMREL at all ages because FPG does not rise sharply with age in populations with low blood glucose.

**Relative risks**

We used Dismod-MR 2.1 to pool effect sizes from included studies and generate a dose-response curve for each of the outcomes associated with high fasting plasma glucose. The tool enabled us to incorporate random effects across studies and include data with different age ranges. RRs were used universally for all countries and the meta-regression only helped to pool the three major sources and produce RRs with uncertainty and covariance across ages taking into account the uncertainty of the data points.

As in GBD 2013, RRs for IHD, ischemic, and hemorrhagic stroke are obtained from meta-regressions of pooled epidemiological studies: the Asia Pacific Cohort Studies Collaboration (APCSC), the Prospective Studies Collaboration (PSC), and the Emerging Risk Factor Collaboration (ERFC).6 These studies have shown that relative risks associated with high fasting plasma glucose decline with the log (RR) having an approximately linear relationship with age, approaching a value of 1 between the ages 100 and 110. Thus we estimated age-specific RRs of using DisMod-MR 2.1 with log (RR) as the dependent variable and median age at event as the independent variable with an intercept at age 110. Morbidity and mortality directly caused by diabetes was considered directly attributable to FPG.

In GBD 2015, we have added peripheral vascular disease as an outcome of diabetes using evidence from the CALIBER study, a recent health record linkage cohort study from the UK. In addition, we have updated the relative risks for tuberculosis as an outcome of diabetes using evidence from recent health record linkage studies from the UK, Australia, and Taiwan, as well as other prospective cohort studies. Please see the citation list for a full list of studies utilized.
References


Input Data & Methodological Summary

Exposure

Case Definition
We measure systolic blood pressure as a continuous exposure in units of mmHg and define hypertension according to the World Health Organization (WHO) standard definition as total blood cholesterol > 6.2 mmol/L and/or currently on cholesterol-lowering medication.¹

Input Data
Consistent with GBD 2013, we utilized data on mean systolic blood pressure from literature and from household survey microdata and reports (e.g. STEPS, NHANES). Please see the appendix for a full list of included sources. In GBD 2013, a systematic review of the literature was completed to capture population survey data on mean systolic blood pressure. For GBD 2015, we updated the systematic review using the same strategy, drawing from the GHDx and Medline via PubMed. In total, we have utilized 537 sources corresponding to 36,727 unique data points.

Global Health Data Exchange Database
We systematically searched the Global Health Data Exchange (GHDx) for multi-country survey programs, national surveys, and longitudinal studies which provide measured individual level data on systolic blood pressure. The search was completed for systolic blood pressure, fasting plasma glucose, and blood cholesterol simultaneously, as many sources studying the other metabolic risks will often report mean cholesterol level or hypercholesterolemia prevalence.
Search Terms (Keywords): Blood pressure OR Blood glucose OR Glucose tests OR Cholesterol OR Cholesterol tests OR Hypercholesterolemia
Data Type: Survey OR Report
Search date: 2/6/2016

Literature Review
We systematically searched PubMed for articles published between 15 July 2009 and 31 December 2015 which provided national or subnational estimates of mean systolic blood pressure. As above, the literature review was completed for systolic blood pressure, fasting plasma glucose, and blood cholesterol simultaneously for the reasons previously stated.

Search terms:

Search date: 1/26/2016

Inclusion Criteria
Studies were included if they were population-based and measured total blood cholesterol using a blood test. We accepted data on hypercholesterolemia only if the study performed an objective blood measurement and/or individuals reported taking cholesterol-lowering medication. Studies that included self-report of high cholesterol were excluded. We assumed the data is representative of the location if the geography was not related to the diseases (a mining area) and if it is not an outlier compared to other data in the country or region.

Outliers
Data was utilized in the modeling process unless an assessment of data showed that the data is biased. A data point was considered to be an outlier candidate if the level is not consistent with other (sources) country data, or - if there are no other data points - not consistent with other country in the region. A candidate outlier source was scrutinized and validated and the data point was excluded if the quality of study did not warrant a valid estimate because of selection (specific populations), different definitions, other biases, or if the study did not provide methodological details for evaluation.

Data Extraction
Where possible, individual level data on blood pressure estimates were extracted from survey microdata and these were collapsed across individuals and collapsed across demographic groupings to produce mean estimates in the standard GBD 5-year age-sex groups. If microdata were unavailable, information from survey reports or from literature were extracted along with any available measure of uncertainty including standard error, uncertainty intervals, and sample size.
Survey reports and literature often only report information about the prevalence of hypercholesterolemia in the population studied. If the study was otherwise representative, we extracted data on the prevalence of hypercholesterolemia and, using all available data with both estimates of mean total cholesterol and prevalence of hypercholesterolemia, crosswalked this to estimates of mean cholesterol levels.

**Crosswalk from Prevalence of Hypercholesterolemia**

We used a mixed-effects regression to crosswalk estimates of hypercholesterolemia prevalence to the mean total cholesterol of a given population. A separate regression was run for a given diagnostic criteria using the form:

\[
\log(TC_{c,a,s,t,k}) = \beta_0 + \beta_1 \logit(p_{c,a,s,t,k}) + \beta_2 \text{male} + \sum_{k=10}^{21} \beta_h I_{A[a]} + \alpha_s + \epsilon_{c,a,s,t,k}
\]

Where \(TC_{c,a,s,t,k}\) is the outcome of interest— the mean total cholesterol of a given country-, age-, sex-, time-, from survey \(k\); \(p_{c,a,s,t,k}\) is the prevalence of hypercholesterolemia for a given definition; \(I_{A[a]}\) is a dummy variable indicating a specific age group \(A\); and \(\alpha_s\) is a super-region specific random effect.

**Age and Sex Splitting**

Prior to modeling, data provided in age groups wider than the GBD 5-year age groups were split using the approach outlined in Ng et al.\(^2\) Briefly, age-sex patterns were identified using sources of data with multiple age-sex groups and these patterns were applied to split aggregated report data. Uncertainty in the age-sex split was propagated by multiplying the standard error of the data performed by the square root of the number of splits performed.

**Modeling**

Exposure estimates were produced from 1980 to 2015 for each national and subnational location, sex, and for each 5-year age group starting from 25+. As in GBD 2013, we used a Spatio-Temporal Gaussian Process Regression (ST-GPR) framework to model the mean total blood cholesterol at the location-, year-, age-, sex- level. Updates to the ST-GR modeling framework for GBD 2015 are detailed in the appendix.

The total cholesterol mean function was estimated using a mixed-effects linear regression, run separately by sex:

\[
\logit(TC_{c,a,t}) = \beta_0 + \beta_1 SDS_{c,t} + \beta_2 p_{overweight_{c,a,t}} + \beta_3 sat_{fats_{c,a,t}} + \beta_4 PUFA_{c,a,t} + \sum_{k=2}^{16} \beta_k I_{A[a]} + \alpha_s + \alpha_r + \alpha_c + \epsilon_{c,a,t}
\]

where \(SDS_{c,t}\) is socio-demographic status (SDS), an index metric that includes a measure of education and income, \(p_{overweight_{c,a,t}}\) is the prevalence of overweight, \(sat_{fats_{c,a,t}}\) is the diet adjusted mean intake of saturated fats per capita per day, \(PUFA_{c,a,t}\) is the diet adjusted mean intake of PUFA per
capita per day, $I_{A[a]}$ is a dummy variable for a fixed effect on a given 5-year age group, and $\alpha_s, \alpha_r, \alpha_c$ are random effects at the super-region, region, and country level, respectively.

The estimates were then propagated through the ST-GPR framework to obtain 1000 draws for each location, year, age, and sex.

**Theoretical minimum-risk exposure level**

For GBD 2015, we altered the TMREL for total cholesterol in light of new evidence from statin trials at low levels of cholesterol; a recent meta-analysis found that cardiovascular outcomes could be improved even at low levels of LDL-cholesterol, below 1.3 mmol/l. We used the strong correlation between LDL-cholesterol and total cholesterol to map the proposed LDL-cholesterol TMREL of 0.7-1.3 mmol/l to a TMREL for total cholesterol of 2.8-3.4 mmol/l.

**Relative Risks**

We used Dismod-MR 2.1 to pool effect sizes from included studies and generate a dose-response curve for each of the outcomes associated with high total cholesterol. The tool enabled us to incorporate random effects across studies and include data with different age ranges. RRs were used universally for all countries and the meta-regression only helped to pool the three major sources and produce RRs with uncertainty and covariance across ages taking into account the uncertainty of the data points.

As in GBD 2013, RRs for IHD and ischemic stroke are obtained from meta-regressions of pooled epidemiological studies: the Asia Pacific Cohort Studies Collaboration (APCSC) and the Prospective Studies Collaboration (PSC). RR for IHD were modeled using with log (RR) as the dependent variable and median age at event as the independent variable with an age intercept (RR equals 1) at age 110. For total cholesterol and ischemic stroke, a similar approach was used, except that there was no age intercept at age 110, due to the fact that there was no statistically significant relationship between total cholesterol and stroke after age 70 with a mean RR less than one. We assumed that there is not a protective effect of high cholesterol and therefore did not include an RR for ages 80+.

**References**


High Systolic Blood Pressure Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure
Case Definition
We measure systolic blood pressure as a continuous exposure in units of mmHg and define hypertension according to the World Health Organization (WHO) standard definition as SBP > 140 mmHg and/or DBP > 90 mmHg and/or currently taking anti-hypertensive medication.¹

Input Data
Consistent with GBD 2013, we utilized data on mean systolic blood pressure from literature and from household survey microdata and reports (e.g. STEPS, NHANES). Please see the appendix for a full list of included sources. In GBD 2013, a systematic review of the literature was completed to capture population survey data on mean systolic blood pressure. For GBD 2015, we updated the systematic review using the same strategy, drawing from the GHDx and Medline via PubMed. In total, we have utilized 844 sources corresponding to 36,727 unique data points.

Global Health Data Exchange Database
We systematically searched the Global Health Data Exchange (GHDx) for multi-country survey programs, national surveys, and longitudinal studies which provide measured individual level data on systolic blood pressure. The search was completed for systolic blood pressure, fasting plasma glucose, and blood cholesterol simultaneously, as many sources studying the other metabolic risks will often report mean blood pressure or hypertension prevalence.
**Search Terms (Keywords):** Blood pressure OR Blood glucose OR Glucose tests OR Cholesterol OR Cholesterol tests OR Hypercholesterolemia  
**Data Type:** Survey OR Report  
**Search date:** 2/6/2016

**Literature Review**

We systematically searched PubMed for articles published between 15 July 2009 and 31 December 2015 which provided national or subnational estimates of mean systolic blood pressure. As above, the literature review was completed for systolic blood pressure, fasting plasma glucose, and blood cholesterol simultaneously for the reasons previously stated.

Search terms:

```plaintext
```

**Search date:** 1/26/2016

**Inclusion Criteria**

Studies were included if they were population-based and measured SBP using a sphygmomanometer (either manual or electronic). Almost all studies reported an average of repeated measurements of SBP done in a visit. We assumed the data is representative of the location if the geography was not related to the diseases (a mining area) and if it is not an outlier compared to other data in the country or region.

**Outliers**

Data was utilized in the modeling process unless an assessment of data showed that the data is biased. A data point was considered to be an outlier candidate if the level is not consistent with other (sources) country data, or - if there are no other data points - not consistent with other country in the region. A candidate outlier source was scrutinized and validated and the data point was excluded if the quality of study did not warrant a valid estimate because of selection (specific populations), different definitions, other biases, or if the study did not provide methodological details for evaluation.

**Data Extraction**

Where possible, individual level data on blood pressure estimates were extracted from survey microdata and these were collapsed across individuals (if multiple measurements were taken for a given individual) and collapsed across demographic groupings to produce mean estimates in the standard GBD 5-year age-sex groups. If microdata were unavailable, information from survey reports or from literature were extracted along with any available measure of uncertainty including standard error, uncertainty intervals, and sample size.

138
Survey reports and literature often only report information about the prevalence of hypertension in the population studied. If the study was otherwise representative, we extracted data on the prevalence of hypertension and, using all available data with both estimates of mean SBP and prevalence of hypertension, crosswalked this to estimates of mean SBP.

**Crosswalk from Prevalence of Hypertension**

We used a mixed-effects regression to crosswalk estimates of hypertension prevalence to the mean SBP of a given population. A separate regression was run for a given diagnostic criteria using the form:

$$\log(SBP_{c,a,s,t,k}) = \beta_0 + \beta_1 \logit(p_{c,a,s,t,k}) + \beta_2 \text{male} + \sum_{k=10}^{21} \beta_h I_{A[a]} + \alpha_s + \epsilon_{c,a,s,t,k}$$

Where $SBP_{c,a,s,t,k}$ is the outcome of interest—the mean SBP of a given country-, age-, sex-, time-, from survey $k$; $p_{c,a,s,t,k}$ is the prevalence of hypertension for a given definition; $I_{A[a]}$ is a dummy variable indicating a specific age group $A$; and $\alpha_s$ is a super-region specific random effect.

**Age and Sex Splitting**

Prior to modeling, data provided in age groups wider than the GBD 5-year age groups were split using the approach outlined in Ng et al. Briefly, age-sex patterns were identified using sources of data with multiple age-sex groups and these patterns were applied to split aggregated report data. Uncertainty in the age-sex split was propagated by multiplying the standard error of the data performed by the square root of the number of splits performed.

**Modeling**

Exposure estimates were produced from 1980 to 2015 for each national and subnational location, sex, and for each 5-year age group starting from 25+. As in GBD 2013, we used a Spatio-Temporal Gaussian Process Regression (ST-GPR) framework to model the mean systolic blood pressure at the location-, year-, age-, sex- level. Updates to the ST-GR modeling framework for GBD 2015 are detailed in the appendix.

The SBP mean function was estimated using a mixed-effects linear regression, run separately by sex:

$$\logit(SBP_{c,a,t}) = \beta_0 + \beta_1 SDS_{c,t} + \beta_2 p_{\text{overweight,}c,a,t} + \sum_{k=2}^{16} \beta_k I_{A[a]} + \alpha_s + \alpha_r + \alpha_c + \epsilon_{c,a,t}$$

where $SDS_{c,t}$ is socio-demographic status (SDS), an index metric that includes a measure of education and income, $p_{\text{overweight,}c,a,t}$ is the prevalence of overweight, $I_{A[a]}$ is a dummy variable for a fixed effect on a given 5-year age group, and $\alpha_s$, $\alpha_r$, $\alpha_c$ are random effects at the super-region, region, and country level, respectively.

The estimates were then propagated through the ST-GPR framework to obtain 1000 draws for each location, year, age, and sex. Parameter selection for the ST-GPR hyper-parameters were selected through cross-validation using the strategy described in the appendix.
Estimate of Standard Deviation
The standard deviation of SBP was estimated for every age, sex, country, and year by estimating the relationship between the mean of SBP and the standard deviation in available studies. To account for in-person variation, person-level microdata were extracted as means across multiple measurements if possible. To further account for regression dilution bias, we estimated the proportion of the variance of SBP accounted for by measurement error and temporal and inter-individual variation and corrected survey estimates of the standard deviation based on an analysis of multiple cohort studies in the United States, China, Indonesia, South Africa, and Brazil.

Theoretical minimum-risk exposure level
We estimated that the TMREL of SBP ranges from 110 to 115 mm Hg based on pooled prospective cohort studies that show risk of mortality increases for SBP above that level. Our selection of a TMREL of 110-115 mmHg is consistent with the GBD study approach of estimating all attributable health loss that could be prevented even if current interventions do not exist that can achieve such a change in exposure level, for example a tobacco smoking prevalence of zero percent. Recent randomized clinical trial results, including the Systolic Blood Pressure Intervention Trial (SPRINT) and the Heart Outcomes Prevention Evaluation (HOPE-3), show that lifestyle modification early in life is likely to be a major component for lowering SBP to near this level given the variable range of benefit observed in these studies when blood pressure was lowered with anti-hypertensive medications alone. To include the uncertainty in the TMREL, we took a random draw from the uniform distribution of the interval between 110 mm and 115 mm Hg each time the population attributable burden was calculated.

Relative risks
As with GBD 2013, RRs for chronic kidney disease are from the Renal Risk Collaboration meta-analysis of 2.7 million individuals in 106 cohorts. For other outcomes, we used data from two pooled epidemiological studies: the Asia Pacific Cohort Studies Collaboration (APCSC) and the Prospective Studies Collaboration (PSC). In GBD 2015, we have added additional estimates of RR for cardiovascular outcomes from the CALIBER study, a health-record linkage cohort study from the UK.

For cardiovascular disease, epidemiological studies have shown that the RR associated with SBP declines with age, with the log (RR) having an approximately linear relationship with age and reaching a value of 1 between the ages of 100 and 120. RRs were reported per 10 mm Hg increase in SBP above TMREL value (115 mm Hg) as in the equation below:

\[ RR_x = RR^{(x-TMREL)} \]

We used Dismod-MR 2.1 to pool effect sizes from included studies and generate a dose-response curve for each of the outcomes associated with high SBP. The tool enabled us to incorporate random effects across studies and include data with different age ranges. RRs were used universally for all countries and the meta-regression only helped to pool the three major sources and produce RRs with uncertainty and covariance across ages taking into account the uncertainty of the data points.
References

1 Bangalore S, Gong Y, Cooper-DeHoff RM, Pepine CJ, Messerli FH. 2014 Eighth Joint National Committee panel recommendation for blood pressure targets revisited: results from the INVEST study. *J Am Coll Cardiol* 2014; 64: 784–93.


Input Data & Methodological summary

Exposure

Case definition

Exposure to high body mass index (BMI) is defined using metrics related to national and subnational estimates of BMI. If a person has a BMI of 25 kg/m² or greater, he/she is considered at risk for a range of diseases including cardiovascular diseases, musculoskeletal disorders, and cancers.

Input Data

We used data from multi-country survey programs, national surveys, and longitudinal studies which were available in the Global Health Data Exchange (GHDx) and provided either self-report or measured data on height and weight. A complete description of the data seeking and update process for the GHDx is provided elsewhere (See Section 2 of appendix). Additionally, to include articles published after our search period for GBD 2013, we systematically searched Medline for studies published between 1 January 2014 and 31 December 2015 providing national or subnational estimates of BMI, overweight, or obesity. The search was conducted on 26 January 2016 using the following search terms: ((("Body Mass Index"[Mesh] OR "Overweight"[Mesh] OR "Obesity"[Mesh]) AND "Geographic Locations"[Mesh] NOT "United States"[Mesh]) AND ("humans"[Mesh] AND "adult"[MeSH]) AND ("Data Collection"[Mesh] OR "Health Services Research"[Mesh] OR "Population Surveillance"[Mesh] OR "Vital statistics"[Mesh] OR "Population"[Mesh] OR "Epidemiology"[Mesh] OR "surve*"[TiAb]) NOT Comment[ptyp] NOT Case

Confidential. Do not cite or circulate.
Reports[ptyp] NOT "hospital"[TiAb]. Of the 2,036 articles identified through Medline search, 162 articles met inclusion criteria and selected for data extraction.

Data Preparation

We adjusted self-reported data for overweight prevalence, obesity prevalence, and mean BMI using the following hierarchical mixed-effects regression models, fit using maximum likelihood separately by sex:

\[
\begin{align*}
\text{logit(overweight)}_{c,a,t} &= \beta_0 + \beta_1 \text{measured} + \sum_{k=2}^{15} \beta_k I_{A[a]} + \sum_{l=16}^{41} \beta_l I_{A[a]} I_{M[m]} + \alpha_s + \alpha_r + \alpha_c + \alpha_t + \epsilon_{c,a,t} \\
\text{logit(obesity)}_{c,a,t} &= \beta_0 + \beta_1 \text{measured} + \sum_{k=2}^{15} \beta_k I_{A[a]} + \sum_{l=16}^{41} \beta_l I_{A[a]} I_{M[m]} + \alpha_s + \alpha_r + \alpha_c + \alpha_t + \epsilon_{c,a,t} \\
\log(BMI)_{c,a,t} &= \beta_0 + \beta_1 \text{measured} + \sum_{k=2}^{15} \beta_k I_{A[a]} + \sum_{l=16}^{41} \beta_l I_{A[a]} I_{M[m]} + \alpha_s + \alpha_r + \alpha_c + \alpha_t + \epsilon_{c,a,t}
\end{align*}
\]

Models included fixed effects on measurement (binary, either measured (1) or self-report (0)) and age group, an interaction between measurement and age group, random intercepts at each level of the geographic hierarchy and by three time periods (1980-1991, 1992-2003, 2004-2015), and random slopes on measurement at each level of the geographic hierarchy and by the three time periods. Random effects at the country and time-period level were used to fit the model, but were taken as noise and were not used in adjustment. We propagated the uncertainty in the model by adding the variance of each of the regression coefficients to the data variance in delta-transformed space.

After adjusting for self-report bias any report or literature data provided in age groups wider than the standard GBD 5-year age groups or as both sexes combined were split using the approach used by Ng et al. Briefly, age-sex patterns were identified using sources with data on multiple age-sex groups and these patterns were applied to split aggregated report data. Uncertainty in the age-sex split was propagated by multiplying the standard error of the data by the square root of the number of splits performed.

Throughout the modeling process, data were assessed for bias and outliers were flagged. At the individual level, BMI <10 or BMI>70 were considered to be outliers, based on biological plausibility. At the population level, data points were flagged as candidate outliers if they were not consistent with the majority of other data points in a country with respect to level, age-pattern, sex-pattern, or temporal trend. In data-scarce countries, data points were also compared to data from other countries in a region. Candidate outliers were scrutinized for potential sources of bias and were ultimately excluded if the point or source was deemed to not be representative.

Modeling strategy

We used Spatio-Temporal Gaussian Process Regression (ST-GPR) to estimate the prevalence of overweight and obesity. Consistent with the approach used in GBD2013, the mean functions used in ST-GPR were estimated using the following linear regressions, run separately by sex:

\[
\text{logit}(ow_{c,a,t}) = \beta_0 + \beta_1 \text{energy}_{c,t} + \beta_2 \text{lat}_{c} + \beta_3 \text{urbanicity}_{c,t} + \sum_{k=4}^{16} \beta_k I_{A[a]} + \epsilon_{c,a,t}
\]
\[ \logit \left( \frac{\text{ob}}{\text{ow}}_{c,a,t} \right) = \beta_0 + \beta_1 \text{energy}_{c,t} + \beta_2 \text{lat}_c + \beta_3 \text{urbanicity}_{c,t} + \sum_{k=4}^{16} \beta_k I_{A[a]} + \epsilon_{c,a,t} \]

where \(\text{energy}_{c,t}\) is a 10-year lag distributed energy intake per capita in country \(c\) at year \(t\), \(\text{lat}_c\) is the absolute latitude of country \(c\), \(\text{urbanicity}_{c,t}\) is the proportion of people living in urban areas in country \(c\) in time \(t\), and \(I_{A[a]}\) is an indicator variable for specific age group \(A\) that the overweight prevalence point \((\text{ow}_{c,a,t})\) or obese as a proportion of overweight point \((\text{ob/ow})_{c,a,t}\) is capturing. The estimated mean functions were then propagated through the ST-GPR framework to obtain 1,000 draws of overweight prevalence estimates and obesity as a proportion of overweight estimates. Based on the results of out-of-sample cross validation, we used different space-weight parameters for locations with low data coverage (less than 15 years covered by data in at least one age-sex group) versus locations with high data coverage (more than 15 years covered by data in at least one age-sex group).

To estimate the mean BMI for each country, age, sex, and time period estimated in GBD, we first used the following equation, fit using data from sources containing estimates of all three indicators, to characterize the relationship between overweight, obesity, and mean BMI:

\[ \text{BMI}_{i,c,a,t} = \beta_0 + \beta_1 \text{overweight}_{i,c,a,t} + \beta_2 \text{obesity}_{i,c,a,t} + \beta_3 \text{sex} + \sum_{k=4}^{17} \beta_k I_{A[a]} + \alpha_s + \alpha_r + \alpha_c + \epsilon_{c,a,t} \]

Then, we applied the regression coefficients to the 1,000 draws of overweight prevalence and obesity prevalence produced through ST-GPR to estimate 1,000 draws of mean BMI for each country, year, age, and sex. This allowed overweight prevalence, obesity prevalence, and mean BMI to be correlated at the draw level, a methodological improvement compared to GBD2013. The estimated mean BMI, overweight prevalence, and obesity prevalence were used to compute the parameters of a beta distribution for BMI in each location, year, age, and sex. The details of this approach have been described by Ng et al. elsewhere. We updated the constraints on the minimum and maximum of the distribution based on biological plausibility.

**Theoretical Minimum-risk Exposure Level**

The TMREL of BMI was determined based on the BMI level that was associated with the lowest risk of all-cause mortality in prospective cohort studies. In GBD 2015, based on the findings of the most recent pooled analysis of prospective cohorts, we changed the TMREL of BMI from 21-23 to 20-25 kg/m².

**Relative Risk**

The relative risk of change in BMI for each disease endpoint was obtained from meta-analyses, and where available, pooled analyses of prospective observational studies. For most outcomes, we have made no substantive updates to relative risks. We dropped the following outcomes, previously included in GBD 2013, due to a lack of conclusive evidence supporting a causal relationship: other cardiovascular diseases, atrial fibrillation and flutter, cardiomyopathy and myocarditis, peripheral vascular disease, and endocarditis. We updated relative risks for osteoarthritis of the hip and osteoarthritis of the knee using recently published meta-analyses.
Bone Mineral Density Capstone Appendix

Flowchart

Input Data & Methodological Summary

Exposure

Case Definition
Bone mineral density (BMD) is a continuous variable measured by dual-x-ray-absorptiometry (DXA) at the femoral neck (FN) and is presenting in g/cm² after standardizing for the brand of densitometer. The burden attributed to low bone mineral density is estimated for adults greater than or equal to 20 years of age.

For estimating burden, we need to estimate:

- Exposure: Mean and standard deviation of standardized BMD according to the brand of densitometer (sBMD) for each country and all subnational levels for which we do GBD estimation.
- Risk of fractures in people exposed to low BMD relative to people who have BMD equal or greater than the TMREL. We consider fatal outcomes for hip and vertebral fractures and non-fatal outcomes for hip and other osteoporosis-associated non-hip fractures. These osteoporotic non-hip fractures include fractures of vertebrae, clavicle, scapula, humerus, skull, sternum, rib, face bone, radius or ulna, femur, patella, tibia, fibula, ankle, pelvis, vertebral and other extremities.

Input data
For GBD 2015, a systematic review was conducted to update the GBD 2013 dataset. Inclusion criteria that informed the search included:

- Representative, population-based surveys
- Reporting of quantitative BMD
  - measured by DXA
• performed at the femoral neck region
• measured in grams/centimeters squared

Mean BMD for was occasionally reported in stratified groups (e.g. by fracture status) so that a total sample BMD mean was not available. In these cases, the stratified means were aggregated to obtain a total mean BMD per study group.

See the search query below for the exact terms used to conduct the systematic review. In GBD 2015, 144 new data points were added, from the following super-regions. The table below indicates the geographic spread of these values.

<table>
<thead>
<tr>
<th>Super region</th>
<th>GBD 2015 new data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Europe, Eastern Europe, and Central Asia</td>
<td>2</td>
</tr>
<tr>
<td>High-income</td>
<td>93</td>
</tr>
<tr>
<td>North Africa and Middle East</td>
<td>11</td>
</tr>
<tr>
<td>South Asia</td>
<td>30</td>
</tr>
<tr>
<td>Southeast Asia, East Asia, and Oceania</td>
<td>8</td>
</tr>
</tbody>
</table>

**Modeling strategy**

We model mean BMD in DisMod-MR. Mean and standard deviation are correlated for BMD. We used a mixed effects model to predict coefficient of variation (standard deviation over mean BMD) with fixed effects on health system access and Ln_LDI (Lag Distributed Income) and random effects on super-region, region and country.

We model continuous mean BMD using a single parameter model for ages 20 to 100, both sexes, and all GBD locations for years 1990 to 2015. The model has age mesh points at 0 10 25 30 40 50 60 70 80 90 & 100, a time window of 10 years for fitting data, and a minimum coefficient of variation of 0.4 for global, 0.2 super region and 0.1 for the region level.

The country covariates of BMI, smoking (different variables), alcohol consumption and milk consumption did not have a significant effect on BMD and some even had a significant effect in the opposite direction to what we know about the pathophysiology. Therefore, we excluded them from our final model.

Some of the data points from the newly added data were outliered during the modeling process.

On both the fatal and non-fatal side, there are various modelling steps that must happen after DisMod modelling of exposure. First, we must calculate the proportion of deaths that are due to fracture. This proportion of death caused by fracture is the envelope that we use to attribute death to bone mineral density. In order to do this, we first used evidence to create a list of all fractures that can be fatal (most are not fatal). Hip fracture and some non-hip fractures (spinal cord and sternum) are considered potentially fatal fractures.

Then, we use available hospital data to estimate what proportion of deaths are due to fracture. However, most hospital data are based just on “E-code”, meaning it is reported that that person died because of a car accident but we do not know the “N-code”, or the nature of the injury (internal hemorrhage, fracture, etc.). First, we restrict to the cases that are dual-coded with both an “E-code” and an “N-code”. Second, many cases have multiple forms of trauma, and therefore, we must apply a severity hierarchy to the fatal...
hospital data to decipher what proportion of the deaths are due to one of the fracture types and not to a more severe fatal trauma. Then, once each observation has just one E-code and N-code we collapse over E-code to find the number of deaths attributable to fracture versus non-fracture injuries. We apply this fracture to the YLL of each of these types of fracture injury.

On the non-fatal side, we first restrict to a list of causes such as transportation injuries, falls, homicide, and disasters that cause non-fatal fractures. Then, we use an E to N-code matrix generated from dual-coded (E-code/N-code) patient level data in order to calculate the proportion of each E-code that results in a certain N-code. This proportion is then applied to the YLDs. The hip and non-hip fracture population attributable fractions are then applied to get the YLD population attributable fractions.

Theoretical minimum-risk exposure level
The theoretical minimum of risk exposure level or TMREL is the age-sex specific 90th percentile of BMD of the NHANES III study as the reference population.

Relative risks
For relative risks, we use a systematic review of bone mineral density that was also used in GBD 2013. Relative risks must be reported per standard deviation or per unit bone mass density in order for us to use the data. Many studies report relative risk based on a z-score or the relative risks in the osteoporotic group versus the non-osteoporotic group; neither of these relative risks are usable.

Updates to systematic reviews are performed on an ongoing schedule across all GBD causes and risk factors, an update for bone mineral density will be performed in the next 1-2 iterations.

The table below illustrates the GBD 2015 search queries.
<table>
<thead>
<tr>
<th>Search</th>
<th>Query</th>
<th>Items found</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>#11</td>
<td>Search (#8 AND #10) Filters: Humans</td>
<td>326</td>
<td>12:37:09</td>
</tr>
<tr>
<td>#8</td>
<td>Search (#7 AND #6) Filters: Humans</td>
<td>622</td>
<td>12:33:16</td>
</tr>
<tr>
<td>#7</td>
<td>Search (&quot;Absorptiometry, Photon&quot;[Mesh] OR &quot;dual-energy x-ray absorptiometry&quot; OR &quot;dual energy x-ray absorptiometry&quot;) Filters: Humans</td>
<td>21368</td>
<td>12:32:34</td>
</tr>
<tr>
<td>#6</td>
<td>Search (#5 AND (&quot;2010&quot;[Date-Publication] : &quot;3000&quot;[Date-Publication])) Filters: Humans</td>
<td>1387</td>
<td>12:30:26</td>
</tr>
<tr>
<td>#5</td>
<td>Search ((#1 OR #2) AND #3) Filters: Humans</td>
<td>3702</td>
<td>12:29:47</td>
</tr>
<tr>
<td>#4</td>
<td>Search ((#1 OR #2) AND #3)</td>
<td>4015</td>
<td>12:29:33</td>
</tr>
<tr>
<td>#3</td>
<td>Search (((&quot;bone mineral density&quot;[title/abstract] OR &quot;bone mineral densities&quot;[title/abstract]) OR &quot;Bone Density&quot;[Mesh]) AND (mean[title/abstract] OR average[title/abstract]))</td>
<td>12892</td>
<td>12:29:00</td>
</tr>
</tbody>
</table>

Confidential. Do not cite or circulate.
Glomerular Filtration Rate Capstone Appendix

Flowchart

Low glomerular filtration rate

Input Data & Methodological Summary

Exposure
Case Definition
For GBD 2015, the reduced glomerular filtration rate (GFR) risk factor, exposure is defined as the three categories of reduced renal function included in the Global Burden of Disease Study (GBD): chronic kidney disease (CKD) stage 3 (GFR of 30-60ml/min/1.73m$^2$), stage 4 (GFR of 15-30ml/min/1.73m$^2$), and stage 5 (GFR <15ml/min/1.73m$^2$, not yet on renal replacement therapy). These exposure categories were each modeled for the GBD 2015 YLD capstone manuscript, and the modeling approach is described in detail there.

Input data
For GBD 2010, a systematic review of the prevalence of CKD throughout the world was conducted. This search was updated for GBD 2013. For GBD 2015, this literature search was repeated using PubMed search terms: ((chronic kidney disease[Title/Abstract]) AND prevalence[Title/Abstract]) AND ('2012/01/01'[Date - Publication] : '3000'[Date - Publication]) (humans).

<table>
<thead>
<tr>
<th>Disease</th>
<th>Number of sources</th>
<th>Super-regions with Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>CKD Stage III</td>
<td>64</td>
<td>All Seven super-regions</td>
</tr>
<tr>
<td>CKD Stage IV</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>CKD Stage V</td>
<td>43</td>
<td></td>
</tr>
</tbody>
</table>
Included surveys were identified by querying the IHME global health data index for any surveys including the term “glomerular filtration rate”. Five total surveys were included of the seventy-six survey results that were identified by the search. Exclusion criteria included surveys that were not population-representative.

**Modeling strategy**

The reduced GFR modeling strategy involved determining the population-attributable burden of cardiovascular outcomes of ischemic heart disease, stroke, peripheral arterial disease, and musculoskeletal outcome gout, to reduced GFR (Equation 1). This was achieved by determining the relative risk of these outcomes based on CKD stage. The CKD stages exposure was obtained from the GBD 2015 analysis, which includes stage-specific prevalence estimates at the country level across twenty age-groups for both genders. CKD stage models included country-level covariates diabetes mellitus and systolic blood pressure. The data informing the model included a cross-walk adjusting data points estimated using the CKD-Epi equation to the MDRD equation, which is our gold standard CKD estimating equation for CKD stages 3-5 for GBD 2015.

The relative risks were calculated by the Chronic Kidney Disease Prognosis Consortium, a consortium composed of population-level cohorts with prospective data collection from several countries (details below). Cardiovascular and gout population prevalences at the country level were obtained from the GBD 2015 Study for the same geographic, time-period, and age-groups as detailed above.

**Theoretical minimum-risk exposure level**

The theoretical minimum risk is a diagnosis of CKD stages 3, 4, or 5 as an eGFR<60ml/min/1.73m² has been demonstrated in the literature to be the GFR below which increased cardiovascular and gout events occur secondary to reduced GFR. (1-10)

**Relative risk**

A two-stage pooled meta-analysis was used to calculate relative risks for ischemic heart disease and stroke. The relative risk of ischemic heart disease and stroke was first determined within each cohort, and then a pooled analysis of cohort-level relative risks was performed using a random effects modeling approach. Uncertainty intervals overlapped in a separate analysis of the relative risk of fatal and nonfatal cardiovascular events from GFR exposure. Thus we decided to use the relative risks from the combined analysis for fatal and nonfatal cardiovascular outcomes. The relative risk for peripheral vascular disease by stage of reduced renal function was determined from the Atherosclerotic Risk in the Communities (ARIC) cohort.(11) Gout relative risk was determined by meta-analysis of a literature review performed for GBD 2013. Search terms included “gout” and “chronic kidney disease”. Exclusion criteria for search results included special populations, reversal of exposure and outcome categories, unclear exposure category definition. This search resulted in four articles.

The relative risks have not changed between GBD 2013 and GBD 2015 analyses.
Population Attributable Fraction

We calculated the cardiovascular and gout fatal and nonfatal burden attributable to the categorical exposure of low GFR stages using the following equation:

\[ PAF = \frac{\sum_{i=1}^{n} P_i (RR_i - 1)}{\sum_{i=1}^{n} P_i (RR_i - 1) + 1} \]

**Equation 1.** PAF based on categorical exposure

where \( RR_i \) is the relative risk for exposure level \( i \), \( P_i \) is the proportion of the population in that exposure category, and \( n \) is the number of exposure categories.\(^{12}\) \( P \) is obtained from GBD 2015 CKD stage estimates, and \( n \) refers to the three CKD stages.

**References**


Section 4. Supplemental Appendix Materials and Detailed Results for Risk Factors

Appendix figures and tables are provided in separate files.