05/06/2021

The csvs in this folder comprise the final input coverage data for IHME's simulation science team's work modeling large scale food fortification for the BMGF nutrition team.

# 04\_01\_21\_ihme\_input\_data.csv

- This file contains the coverage data used for the model runs and results sent on 04/01/2021.

# 04\_19\_21\_final\_results\_input\_data.csv

- This file contains the coverage data used for the final model runs and results sent on 04/19/2021.

- industry oil: this vehicle uses data sent by BMGF industry partners, and shared with IHME via the BMGF nutrition team, regarding the percentage of the population consuming fortified and fortifiable oil.

- industry wheat: this vehicle uses data sent by BMGF industry partners, and shared with IHME via the BMGF nutrition team, regarding the percentage of the population consuming fortified and fortifiable wheat.

- salt fortified with iodine: we received this data from the BMGF nutrition team. Using these numbers, we estimated the impact of scaling folic acid fortification of salt from zero to the levels at which salt is currently fortified with iodine.

# oil\_data\_comparison.csv

After delivering model results on 04/01/2021, we underwent a series of model updates, using data sent by the BMGF nutrition team. On 04/16/2021, Jonathan notified us of numbers we had used for “oil fortified with vitamin a” in error, and a set of numbers he requested that we use instead. This csv contains these two sets of numbers, with citations for the numbers we had used before switching to the requested numbers.

* Value\_mean\_requested: the means Jonathan sent on 04/16/2021 for usage.
* Value\_mean\_used: the means IHME used prior to switching to value\_mean\_requested.
* Estimation\_status\_used: estimation\_status of the old data used.
* Source\_citation\_used: citation of the old data used.
* Source\_link\_used: source link for the old data used.
* CI\_source\_used: whether the confidence intervals of the old data IHME used were extracted or modelled.
* Data\_choice\_notes\_used: notes on how old data was chosen out of all data extracted.
* Source\_citation\_requested: citation for new data.

# Data dictionary notes

value\_description

* Values: “percent of population eating vehicle”, “percent of population eating industrially produced vehicle”, “percent of population eating fortified vehicle”
* Note that we use “industrially produced vehicle” as an analog for “fortifiable vehicle”.
* All of our models consider the impact of scaling up fortification from baseline (currently fortified) to the ceiling defined by what is fortifiable. The variable “percent of the population eating vehicle” was not directly used in our models, but rather used to model “percent of population eating industrially produced vehicle” and “percent of population eating fortified vehicle” when these variables were not available in the literature. We include the values we were able to find for “percent of the population eating vehicle” here for completeness.

u5\_applicable

* Boolean value: defines if used in models specific to under-five year olds.

wra\_applicable

* Boolean value: defines if used in models specific to women of reproductive age.

estimation\_status

* Defines how “value\_mean” was estimated.

|  |  |
| --- | --- |
| **Value** | **Interpretation** |
| na | Not modelled |
| regression | Modelled using a gradient-boosted machine learning regression model |
| multiplicative | Modelled using a multiplicative model |
| multiplicative + rejection sampling' | The variable was first modelled using a multiplicative model; then, due to logical inconsistency between "percent of population eating industrially produced vehicle" and "percent of population eating fortified vehicle", logical values within the confidence intervals were chosen with rejection sampling |
| regression + rejection sampling | The variable was first modelled using a gradient-boosted machine learning regression model; then, due to logical inconsistency between "percent of population eating industrially produced vehicle" and "percent of population eating fortified vehicle", logical values within the confidence intervals were chosen with rejection sampling |
| na + rejection sampling | The variable was not initially modelled (extracted); then, due to logical inconsistency between "percent of population eating industrially produced vehicle" and "percent of population eating fortified vehicle", logical values within the confidence intervals were chosen with rejection sampling |
| lack of evidence | We assumed that none of the vehicle is fortified due to a lack of evidence to the contrary |

source\_citation

* Citation for value\_mean.

source\_link

* Link for source of value\_mean.

CI\_source

* Defines whether the confidence intervals were extracted or modelled, if available.

data\_choice\_notes

* Notes how the given data was chosen if we extracted multiple data points for the corresponding value and discarded some of these.