GRAM Model: Data Needs for Adaptations to Different Settings

This document lists data inputs for the GRAM model that require external sourcing. Internal model definitions are excluded; model parameters are defined in gram\_01setup.R.

The model is currently set up to use data from the United States. If you are using the model in a different setting, you will need to adjust the data inputs accordingly, including defining a cohort that represents the population of interest. There are two ways to do this:

1. **Read in an individual-level dataset** that represents your population of interest (e.g., census microdata). Ensure that the dataset includes all attributes defined in the model. GRAM uses this approach, and for a new setting would use a dataset like this to simulate the population. This is the preferred option if such individual-level data are available.
2. **Define distributions for each individual attribute** (e.g., age, sex, education, etc.) for your population of interest. This is less preferred than approach #1. GRAM will create a synthetic population by randomly assigning attributes to individuals based on the distributions. This information might be easier to source in some settings. However, the dependencies between attributes (e.g., age and sex) will need to be accounted for, which will require revisions to the model and may not always be possible if data on associations between attributes are not available.

**Required Data to Portray a Selected Setting**

*Note: Reasonable approximations may suffice for discussion.*

* Age-specific mortality (total and ideally for non-dementia-related deaths) (we have this for South Africa)
* Prevalence or incidence of MCI and dementia by age (can find for Joburg or from U of Cape Town)
* Excess mortality (relative) risk among those with MCI/dementia (ask at Cape Town)
* Test (BHA, CDR-SB, or other) validity and performance metrics (schedule time with Elena Tsoy to discuss options – she works on adapting the BHA for other countries. Cyprian will talk with Kirti)
* Rate of progression of cognitive impairment (we expect to have to do some estimations here, data are hard to find)
* Risk factors for dementia and their distribution in the population (Cyprian did lit review for Africa for most dominant risk factors – education, poverty, infections, diabetes, obesity)
* Health state utility by dementia status (DALYs? QALYs?) – can look at Tufts registry
* Cost of care and treatment by dementia status
* Wage rates (general and for healthcare) – supportive care + productivity loss

**Ideal External Data Inputs for GRAM**

| Input Name | Description | Suggested Data Sources | | Notes |
| --- | --- | --- | --- | --- |
| AGE\_start\_mean, AGE\_start\_sd | Mean and SD of starting age for simulated individuals | NHANES, HRS, ADNI, census | Relevant only if modeling a prevalent cohort. Default model starts at age 50. | |
| p.SEX\_start\_male, p.SEX\_start\_female | Proportion of males and females in the starting population | NHANES, HRS, census |  | |
| p.EDU\_start | Proportion with each education level (college+, high school, < high school) | NHANES, HRS, census |  | |
| p.RACEETH\_start | Proportion with each race/ethnicity (White, Black, Hispanic) | NHANES, HRS, census |  | |
| p.INCOME\_start | Proportion with each income level (low, medium, high) | NHANES, HRS, census, government economic reports | Model cutoffs: low = <$9,000/year; medium = $9,000–$36,000/year; high = >$36,000/year. Define equivalent cutoffs for your setting. | |
| p.APOE4\_start | Proportion of APOE4 carriers vs. non-carriers | ADNI, NACC, published genetic studies |  | |
| p.SEV\_start | Proportion with each severity at baseline (MCI, mild, mod, severe) | Clinical cohorts, ADNI, NACC | Only required if starting with a prevalent cohort. Otherwise, defaults to NA (i.e., no one has impairment) | |
| MEDBUR\_start | Distribution of comorbidities at baseline | NHANES, HRS, health system data |  | |
| m.cogcon\_spon, m.cogcon\_elic | Matrices for spontaneous/elicited cognitive concerns representing the probability of receiving a cognitive test based on age and true cognitive status | Peer-reviewed literature, original cohort data | Only required if simulating a cognitive testing intervention. Otherwise, defaults to 1 (i.e., everyone receives a cognitive test). | |
| m.lifetable | Age-specific non-dementia mortality probabilities | CDC/NCHS life tables, WHO, published tables |  | |
| m.hr\_mci | Age-specific annual MCI incidence rates | Peer-reviewed epidemiological studies, ADNI, NACC |  | |
| Health state utilities | Health state utilities for cognitive impairment states (e.g., QALY values) | Literature reviews, meta-analyses, clinical studies |  | |
| Costs | Annual costs by health state and treatment | Peer-reviewed cost studies, CMS, health system data |  | |
| Cognitive test cutoffs | CDR and other test thresholds | Peer-reviewed literature, clinical guidelines | GRAM tracks severity of cognitive impairment based on the CDR-SB measure. The cutoffs are necessary to determine categorical severity of impairment. Default values are based on O’Bryant et al. (2008). | |
| Sensitivity/specificity | Sensitivity/specificity of cognitive tests | Peer-reviewed literature, validation studies | Only required if simulating a cognitive testing intervention. | |
| Treatment parameters | Efficacy, eligibility, duration, risk ratios | Clinical trials, meta-analyses, regulatory reports | Only required if simulating a treatment intervention. | |

**Abbreviations:** - NHANES: National Health and Nutrition Examination Survey - HRS: Health and Retirement Study - ADNI: Alzheimer’s Disease Neuroimaging Initiative - NACC: National Alzheimer’s Coordinating Center - CMS: Centers for Medicare & Medicaid Services - CDC/NCHS: Centers for Disease Control/National Center for Health Statistics