Memo：

4/16/2022

注意到Rstata网站提供的空气质量数据为2014年开始的，yaoyao文章里是取了受访者受访前1年的空气质量数据取平均，涉及到2013年。

为了在现阶段简化初次数据清洗，先直接取现成的2014年和2017年年平均PM2.5，CO，NO2数据做空气质量指标。

而yaoyao的空气质量数据来自于一个公开的数据库：

We generated residence-specific ambient concentrations of six pollutants (PM2·5, PM10, SO2, NO2, O3, and CO) using a previously published database.21 Nationwide monitoring network observation is the gold standard for air pollution exposure assessment, but they were not available for participants who lived a long distance from the monitoring sites.

这个讲到了时间，数量，空间差异

公开数据库：

文献：<https://essd.copernicus.org/articles/13/529/2021/#section5>

2013年的日度数据：

<https://www.scidb.cn/en/detail?dataSetId=696756084735475712&dataSetType=personal>

2013年起的月度数据：

<https://www.scidb.cn/en/detail?dataSetId=712258947691577344&dataSetType=personal>

公开数据库说明：

 It provides the surface gridded fields of six conventional air pollutants (i.e. PM2.5，PM10，SO2，NO2，CO and O3) and the simulated surface fields of wind speed (u, v), pressure (psfc), relative humidity (RH) and temperature (temp) by WRF model. The spatial and temporal resolutions are respectively 15km and 1 hour. Currently, the time period of the dataset is from 2013 to 2019, and will be updated irregularly.

Data assimilation method: The dataset was produced by the chemical data assimilation system (ChemDAS) developed by IAP, CAS, which assimilates over 1000 surface air quality monitoring sites from CNEMC based on the ensemble Kalman filter (EnKF) and the Nested Air Quality Prediction Modeling system (NAQPMS). This method broke through the problems of instability, insufficient adjustment and negative assimilation effect in atmospheric chemistry data assimilation and develops the multi-air pollutant collaborative assimilation including the monitoring data automatic quality control methods, adaptive mode error estimation and other advanced algorithms. It has been published in Earth System Science Data, where detailed descriptions and validation of this dataset are available (https://doi.org/10.5194/essd-13-529-2021).

yaoyao下载的是月度数据。我们尝试效仿这种做法，已经下载了月度数据准备清洗和匹配栅栏。

4/17/2022

Frailty的定义来自A limit to frailty in very old, community-dwelling people: a secondary analysis of the Chinese longitudinal health and longevity study

<https://academic.oup.com/ageing/article/42/3/372/23917>

The FI was created following a standard procedure [18]. Symptoms, signs, disabilities and diseases were considered as health deficits, if they: were associated with the health status; had a prevalence >1%; generally increased with age, recognising that some age-related adverse conditions become less common in extreme old age due to survivor effects; did not saturate (i.e. were found to be present in most people at an early age, quantified here as a prevalence of 80%) and in their aggregate, represented several physiological systems [18]. Each variable was re-coded as 0 (absence of deficit), 1 ( presence of deficit) or missing.

The 1998–2005 CLHLS data set included 45 deficits, 29 of which had been used by another group to construct an FI [11,19]. We considered all 29 of these, plus 16 other variables that met the above criteria. Of these 45 variables, 4 were excluded due to saturation, 2 because they were not age related and 1 (‘Feel lonely and isolated’) because it was considered to reflect social vulnerability more than frailty. For each individual, the FI score was calculated by summing the deficits present and dividing by the 38 deficits considered (Supplementary data are available in Age and Ageing online, Appendix A).

|  |  |  |
| --- | --- | --- |
| Variables Used Here That Are Common To Gu et al's[18] Frailty Index | | Data Type |
| 1 | Self-reported health | Ordinal |
| 2 | Feel fearful or anxious | Ordinal |
| 3 | Feel useless with age | Ordinal |
| 4 | Bathing | Ordinal |
| 5 | Dressing | Ordinal |
| 6 | Toileting | Ordinal |
| 7 | Transferring | Ordinal |
| 8 | Continence | Ordinal |
| 9 | Feeding | Ordinal |
| 10 | Visual function | Ordinal |
| 11 | Rhythm of heart | Binary |
| 12 | Hand behind neck | Ordinal |
| 13 | Hand behind lower back | Ordinal |
| 14 | Able to stand up from sitting | Ordinal |
| 15 | Able to pick up a book from the floor | Ordinal |
| 16 | Number of times suffering from serious illness in the past two years | Ordinal |
| 17 | Hypertension | Binary |
| 18 | Diabetes | Binary |
| 19 | Heart disease | Binary |
| 20 | Stroke or CVD | Binary |
| 21 | Bronchitis, emphysema, pneumonia, asthma | Binary |
| 22 | Tuberculosis | Binary |
| 23 | Cancer | Binary |
| 24 | Gastric or duodenal ulcer | Binary |
| 25 | Parkinsons | Binary |
| 26 | Bedsore | Binary |
| 27 | Able to hear | Binary |
| 28 | Interviewer rated health | Ordinal |
| 29 | Look on the bright side of things | Ordinal |
| 30 | Keep my belongings neat and clean | Ordinal |
| 31 | Make own decisions | Ordinal |
| 32 | Housework at present | Ordinal |
| 33 | Able to use chopsticks to eat | Ordinal |
| 34 | Number of steps used to turn around a 360 degree turn without help | Interval |
| 35 | Cataract | Binary |
| 36 | Glaucoma | Binary |
| 37 | Other chronic disease | Categorical |
| 38 | Prostate Tumor | Binary |

基本都可以覆盖，就是这些变量都要变成0-1变量，有些measure的分类很细，不知道哪些该归为1，哪些该归为0。目前都是尽可能往不衰弱方向code

但是这里面涉及到心理，还是存在measure不一致的问题。

! 心里问题，2018年问题和2014年问题的校对

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variables | 2014 |  | 2018 |  |
| Psy1 | 不论遇到什么事您是不是都能想得开？ | b21 | 不论遇到什么事您是不是都能想得开？ | b21 |
| Psy2 | 您是不是喜欢把东西弄得干净、整洁？ | b22 | 我喜欢把我的东西弄得干净、整洁 | b22 |
| Psy3 | 您是不是经常感到紧张、害怕 | b23 | 您是不是经常感到紧张、害怕 | b36 |
| Psy6 | 您是不是觉得越老越不中用？ | B26 | 我觉得越老越不中用 | b34 |
| Psy5 | 您自己的事情是不是自己说了算？ | B25 |  | 无 |

去掉“您自己的事情是不是自己说了算？“，最终留下37个

生成:

frailsum：37问加总

frailmissing：37问缺失

frailID：37问加总/（37-37问缺失）

frail\_cat：对frailID做六分位1-6。（但是分出来不是真正的人数比六等份，只是基本平均的分类）数字越小，衰弱情况越轻微。

4/18/2022

合并所有数据

1. 清除37问frailty中，至少有5道missing的
2. 保留城市有空气质量检测的数据
3. 保留14，18都有观测的数据

最终得到2,158 个样本。做成长数据，一共4316个观测值。

按照如下城市区分

treat=1 18个省

treat=0 5个省

表格

描述已自动生成

4/18/2022

### https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/

发现了另一种measure frailty的方式，感觉可以用来做robustness检验

### Criteria for Frailty

The modified Fried criteria was adopted to define the frailty status ([2](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B2), [12](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B12), [17](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B17)). Five domains including exhaustion, shrink, weakness, low mobility, and inactivity were evaluated using self-report data.

Exhaustion was defined if the participant answered “always,” “often,” or “sometimes” to either of the questions “I felt old and useless” or “I felt everything I did was an effort” ([17](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B17), [27](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B27)). Shrink was defined as BMI < 18.5 kg/m2 ([12](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B12), [17](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B17), [28](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B28)). Weakness was defined if the participant failed to lift a bag weighting 5 kg ([12](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B12), [28](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B28)). Low mobility was defined if the participant failed to walk for 1 km ([29](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B29)). Inactivity was defined if the participant did the following activities 1 time per week or less: housework, outside activity, gardening, keeping a pet, livestock breeding, playing cards or moh-jong, and social activity ([27](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8365226/#B27)).

Participant meeting 1–2 domains was defined as pre-frailty. Participant meeting ≥3 domains was defined as frailty. The prevalence of pre-frailty and frailty was defined according to the 2008 interview.

对了，其实我想起来经济学中有一个研究经济状况和健康的（忘记结局是啥了），用charles和clhls，结合国家养老保险政策的改变，用DID发了一个很不错的刊物。

我感觉健康结局是可以换的。不确定是否适合。

4/23/2022

加入受访者受访月所在区县的平均温度：temperature （无缺失）

加入受访月：month

加入饮用水质量：waterqual 1 "tap water" 2 "natural water" 3 "well water"

加入收入来源：incomesource (1=1 "Retirement pension") (2/5 = 2 "Family support") (6 = 3 "Social insurance") (7 = 4 "Working payment") (8=5 "others")

百度，坐标纠偏，火星坐标系，专成WGS 1984坐标系，在栅格图中点出。中国区域污染物模拟。LUR/ML/CM社区模拟模型。一张年均值浓度图。再导出成浓度图。不需要，R sql。R sqldf。每年年均值。