Table 1 - Baseline Characteristics of CHARLS

Bruce Zhou

2024-11-26

Table of Contents

# import dataset

load("outputs/pooled\_11\_15\_hearing.RData")  
load("outputs/pooled\_11\_15\_covariates.RData")  
load("outputs/pooled\_11\_15\_PF.RData")  
  
### merge datafarmes ###  
pooled\_11\_15 <- left\_join(pooled\_11\_15, pooled\_11\_15\_Cov, by = c("id", "wave"))  
pooled\_11\_15 <- left\_join(pooled\_11\_15, pooled\_11\_15\_PF, by = c("id", "wave"))  
  
# limit the age  
pooled\_11\_15 <- pooled\_11\_15 %>%  
 filter(age >= 60) # 13338 participants excluded, 11620 left  
  
# stratify the visit  
pooled\_11\_15 <- pooled\_11\_15 %>%  
 arrange(id, wave) %>%  
 group\_by(id) %>%  
 mutate(visit = row\_number() - 1) %>%  
 ungroup()  
  
# keep only the baseline  
pooled\_11\_15 <- pooled\_11\_15 %>%  
 arrange(id, wave) %>%  
 group\_by(id) %>%  
 slice(1) %>%  
 ungroup()

# Type conversion

# [1] "id" "wave" "age" "sex" "education" "marital"   
# [7] "residence" "srhear" "hraid" "hr" "visit" "hhconsumpLevel"  
# [13] "mbmi" "hibpe" "diabe" "cancer" "lunge" "hearte"   
# [19] "stroke" "arthre" "dyslipe" "livere" "kidneye" "asthmae"   
# [25] "smoke" "drink" "gait" "cstand" "balance" "gaitScore"   
# [31] "standScore" "SPPB"  
  
pooled\_11\_15 <- pooled\_11\_15 %>%  
 mutate(hr = factor(hr, labels=c('Impaired Hearing', 'Normal Hearing'), levels=c(1,0))) %>%  
 mutate(sex = as.numeric(sex),  
 education = factor(education,levels = c(1,2,3), labels=c("Less than lower secondary","Upper secondary & vocational training","Tertiary")),  
 marital = as.numeric(marital),  
 residence = as.numeric(residence),  
 hhconsumpLevel = factor(hhconsumpLevel, labels = c("Q1 (< 2628)", "Q2 (2628 – 4528)", "Q3 (4528 - 8288)", "Q4 (> 8288)")),  
 mbmi = factor(mbmi, labels = c('Underweight (< 18.5 kg/m2)', 'Normal Weight (18.5 – 23 kg/m2)', 'Overweight (23 – 25 kg/m2)', 'Obesity (>= 25 kg/m2)')),  
 smoke = as.numeric(smoke),  
 drink = as.numeric(drink),  
 hibpe = as.numeric(hibpe),  
 diabe = as.numeric(diabe),  
 cancer = as.numeric(cancer),  
 lunge = as.numeric(lunge),  
 hearte = as.numeric(hearte),  
 stroke = as.numeric(stroke),  
 arthre = as.numeric(arthre),  
 dyslipe = as.numeric(dyslipe),  
 livere = as.numeric(livere),  
 kidneye = as.numeric(kidneye),  
 asthmae = as.numeric(asthmae),  
 wave = factor(wave, labels = c("Wave 2011", "Wave 2013", "Wave 2015")))

# Table creation

pooled\_11\_15 %>%  
 dplyr::select(hr, age, sex, education, marital,  
 residence, hhconsumpLevel, mbmi, smoke, drink,  
 hibpe, diabe, cancer, lunge, hearte,  
 stroke, arthre, dyslipe, livere,  
 kidneye, asthmae, wave) %>%  
 tbl\_summary(by = hr,   
 missing = "no",  
 # missing\_text = 'Missing',  
 label = list(age ~ "Age", sex ~ "Gender - female", education ~ "Education level",  
 marital ~ 'Marital status - alone', residence ~ 'Residence – rural',  
 hhconsumpLevel ~ 'Annual consumption per capita (RMB)', mbmi ~ 'Body mass index (BMI)',  
 smoke~'Cigarette consumption – ever smoked', drink~'Alcohol consumption – ever drank',  
 hibpe~'Hypertension', diabe~'Diabetes', cancer~'Cancer',  
 lunge~'Lung disease', hearte~'Heart problems', stroke~'Stroke',  
 arthre~'Arthritis', dyslipe~'Dyslipidemia', livere~'Liver disease',  
 kidneye~'Kidney disease', asthmae~'Asthma', wave ~ 'Enrollment wave'),  
 statistic = list(all\_continuous() ~ "{mean} ({sd})", all\_categorical() ~ "{n} ({p}%)", age ~ "{median} ({IQR})"),  
 digits = list(all\_categorical() ~ c(0,1),  
 age ~ c(0,0))) %>%  
 # add\_p(pvalue\_fun = ~ style\_pvalue(.x, digits = 2)) %>%  
 add\_overall() %>%  
 # add\_n() %>%  
 modify\_footnote(  
 all\_stat\_cols() ~ "Median (IQR) for age; N (%) for other categorical variables."  
 ) %>%  
 modify\_caption("\*\*Table 1. Baseline characteristics of study participants in CHARLS wave 2011 - 2015\*\*") %>%  
 modify\_spanning\_header(c("stat\_1", "stat\_2") ~ "\*\*Baseline Hearing Status\*\*") %>%  
 bold\_labels()

## 434 observations missing `hr` have been removed. To include these observations, use `forcats::fct\_na\_value\_to\_level()` on `hr` column before passing to `tbl\_summary()`.

## Table printed with `knitr::kable()`, not {gt}. Learn why at  
## https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html  
## To suppress this message, include `message = FALSE` in the code chunk header.

**Table 1. Baseline characteristics of study participants in CHARLS wave 2011 - 2015**

| **Characteristic** | **Overall**, N = 11,186 | **Impaired Hearing**, N = 7,099 | **Normal Hearing**, N = 4,087 |
| --- | --- | --- | --- |
| **Age** | 64 (9) | 64 (10) | 63 (8) |
| **Gender - female** | 5,592 (50.0%) | 3,542 (49.9%) | 2,050 (50.2%) |
| **Education level** |  |  |  |
| Less than lower secondary | 10,312 (92.2%) | 6,644 (93.6%) | 3,668 (89.7%) |
| Upper secondary & vocational training | 642 (5.7%) | 334 (4.7%) | 308 (7.5%) |
| Tertiary | 229 (2.0%) | 118 (1.7%) | 111 (2.7%) |
| **Marital status - alone** | 2,121 (19.0%) | 1,419 (20.0%) | 702 (17.2%) |
| **Residence – rural** | 6,508 (58.2%) | 4,318 (60.8%) | 2,190 (53.6%) |
| **Annual consumption per capita (RMB)** |  |  |  |
| Q1 (< 2628) | 2,062 (23.9%) | 1,366 (25.2%) | 696 (21.6%) |
| Q2 (2628 – 4528) | 2,103 (24.3%) | 1,367 (25.2%) | 736 (22.8%) |
| Q3 (4528 - 8288) | 2,200 (25.5%) | 1,337 (24.7%) | 863 (26.7%) |
| Q4 (> 8288) | 2,278 (26.4%) | 1,345 (24.8%) | 933 (28.9%) |
| **Body mass index (BMI)** |  |  |  |
| Underweight (< 18.5 kg/m2) | 763 (8.7%) | 521 (9.2%) | 242 (7.8%) |
| Normal Weight (18.5 – 23 kg/m2) | 3,718 (42.4%) | 2,424 (42.9%) | 1,294 (41.6%) |
| Overweight (23 – 25 kg/m2) | 1,740 (19.9%) | 1,101 (19.5%) | 639 (20.5%) |
| Obesity (>= 25 kg/m2) | 2,544 (29.0%) | 1,606 (28.4%) | 938 (30.1%) |
| **Cigarette consumption – ever smoked** | 4,781 (42.8%) | 3,077 (43.4%) | 1,704 (41.7%) |
| **Alcohol consumption – ever drank** | 4,711 (42.2%) | 3,025 (42.7%) | 1,686 (41.4%) |
| **Hypertension** | 3,569 (33.7%) | 2,340 (34.8%) | 1,229 (31.7%) |
| **Diabetes** | 870 (8.3%) | 575 (8.6%) | 295 (7.7%) |
| **Cancer** | 119 (1.1%) | 81 (1.2%) | 38 (1.0%) |
| **Lung disease** | 1,415 (13.4%) | 998 (14.9%) | 417 (10.8%) |
| **Heart problems** | 1,767 (16.7%) | 1,217 (18.2%) | 550 (14.2%) |
| **Stroke** | 418 (3.9%) | 282 (4.2%) | 136 (3.5%) |
| **Arthritis** | 3,926 (37.0%) | 2,724 (40.5%) | 1,202 (31.0%) |
| **Dyslipidemia** | 1,337 (12.9%) | 902 (13.7%) | 435 (11.4%) |
| **Liver disease** | 431 (4.1%) | 306 (4.6%) | 125 (3.2%) |
| **Kidney disease** | 697 (6.6%) | 509 (7.6%) | 188 (4.9%) |
| **Asthma** | 632 (6.0%) | 467 (7.0%) | 165 (4.3%) |
| **Enrollment wave** |  |  |  |
| Wave 2011 | 7,228 (64.6%) | 4,580 (64.5%) | 2,648 (64.8%) |
| Wave 2013 | 2,004 (17.9%) | 1,234 (17.4%) | 770 (18.8%) |
| Wave 2015 | 1,954 (17.5%) | 1,285 (18.1%) | 669 (16.4%) |