Some of the variables we use are slightly different from the ones used in Jiang et al., but they are not the cause of the discrepancies described below and are omitted from the below discussion.

Data cleaning

Below is a table with data cleaning numbers. The initial and final sample size (first and last row, respectively) refer to the numbers in the sample, while all other rows refer to the numbers of participants removed from the sample.

Step	Jiang et al.	Mur
Initial sample size	502,506	502,369
No hearing assessment	25,081	25,101
Dementia at baseline	283	217
Missing confounder data	39,348	93,005
Final sample size	437,704	384,183

Problems:

- Missing confounder data:
    1. The following variables have the most missing values in our sample: income (68,370), depressive symptoms (20,846), loneliness (7,638), education (5,027), physical activity (2,237), smoking (1,779), waist circumference (instead of BMI, 1,626), deprivation (592), social isolation (569), and alcohol use (421). There are just a handful missing cases for the history of the different disorders that were also included as covariates.
    2. Among the above variables, income (code 738), depressive symptoms (code 2050), loneliness (code 2020), smoking (code 20116), deprivation (code 22189[1]), social isolation (codes 709, 1031, 6160), and alcohol use (code 1558) are operationalised identically to the procedures by Jiang et al.
    3. Among the confounders common to my analysis and that by Jiang et al (point 2 above), the numbers of missing values in my sample closely correspond to the numbers of missing values (i.e., responded with “do not know” or with “prefer not to answer”) reported by UK Biobank (https://biobank.ndph.ox.ac.uk/showcase/). The missing values for variables in point 2 in my sample sum up to 100,215[2], whereas in UK Biobank they sum up to 109,584. This makes sense, because in the analysis, this is not the first data cleaning step, and one would expect less cases than in UKB. In our analysis, these missing values are found in 93,005 participants. However, in Jiang et al., all missing values, including other confounders not among those in point 2 above, sum up to cover only 39,348 participants. Why is this the case?
- Confusion regarding missing hearing aid data:
  Some people reported wearing hearing aids without reporting hearing difficulties. Jiang et al. do not explicitly write what they did with the participants that were not asked the question about hearing aid use (although they do write that not all participants were asked the question). But because the missing data were mostly in participants that reported no hearing

[1] Incorrectly referred to by Jiang et al. as code “189” (supplementary material).
[2] More than the number of individuals with missing values, because a single participant can have missing values in several variables.

loss, I am assuming that they re-coded them as not using hearing aids. I further assume that they also re-coded to “no hearing aid use” those that reported hearing aids despite reporting no hearing difficulties.

Results

The two samples (mine vs. Jiang et al.) are (apart from the different absolute numbers due to the above data-cleaning discrepancies) very similar in composition regarding all variables included in the modelling. Hearing loss and hearing aid use are also more common in men, positively correlated with waist circumference (BMI in Jiang et al.), age, cardiovascular disease, loneliness, and depressed mood, as in Jiang et al.

However, the main results are not the same. I divide the sample into two subsamples, (1) one containing participants without hearing loss or with hearing loss and without hearing aid use (n=373,210); and (2) one containing participants without hearing loss or with hearing loss and with hearing aid use (n =297,545). In each subsample I then run a Cox model, effectively testing the difference between the group with hearing loss (and no hearing aids in 1; and hearing aid use in 2) and the group without hearing loss. The results are below:

Model 1 (rate of dementia in aid non-users vs. controls)
simple adjustment: HR=1.11 (95% CI=1.05-1.19)
complex adjustment: HR=1.03 (95% CI=0.96-1.10)

Model 2 (rate of dementia in aid users vs. controls)
simple adjustment: HR=1.50 (95% CI=1.35-1.67)
complex adjustment: HR=1.27 (95% CI=1.14-1.42)

Thus, the dementia rate is higher in hearing aid users vs. controls than in non-hearing aid users vs. controls – the opposite result to that reported by Jiang et al.