## **Appendix**

## Algorithm for the allocation of the information on education

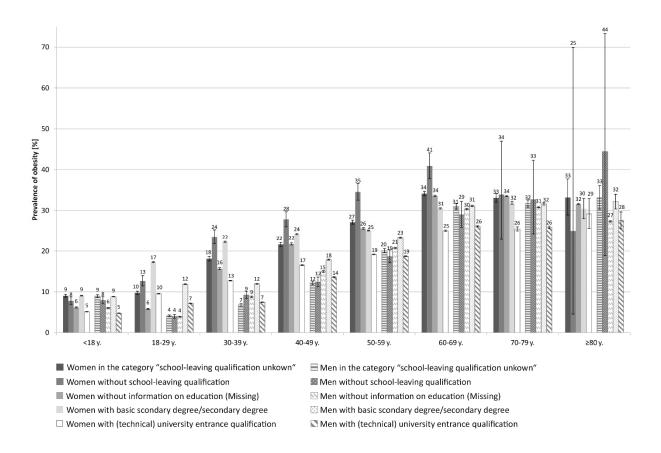
The following steps were performed to allocate the educational status:

- [1] Assumption of school education based on vocational qualification. Persons with a university degree are assumed to possess the (technical) university entrance qualification, irrespective of the information on their educational background.
- [2] Identification of the highest educational status based on the classification before or after 2011. At this point, the search for information on education continues even if the information "no formal degree" is identified in this step.
- [3] Assumption of school education based on occupational status (e.g., persons with highly complex jobs are classified as having the (technical) university entrance qualification).
- [4] Transfer of the highest known information on education of the linked principal insurance holder. If a person has more than one linked principal insurance holder over time, the one with the highest educational status is selected.
- [5] Transfer of the highest educational status of a linked family member whose educational status is known. If a person has more than one linked family member over time, the one with the highest educational status is selected.

Procedure for approximating the county level

The basis for using the German Index of Socioeconomic Deprivation was the definition of the index published by the Robert Koch Institute in its revised version from 2018 and based on data from 2014. However, the county codes in GePaRD reflect the 2017 county level. There was a merger of two counties between the 2014 and 2017 area statuses. In order to apply the GISD to the GePaRD data despite this discrepancy, the index for the newly created county was calculated on a population-weighted basis from the indices of the two previous counties.

Figure 7: Prevalence of obesity in GePaRD for 2017 stratified by age, sex, and individually assigned educational level, including the category school-leaving qualification "unknown" (based on occupation key) and including individuals with no information on education at all



Attachment 1 to: Asendorf M, Reinold J, Schink T, Kollhorst B, Haug U. Abbildbarkeit des sozioökonomischen Status in der pharmakoepidemiologischen Forschungsdatenbank GePaRD: Beschreibung und Anwendung am Beispiel des Zusammenhangs mit Adipositas [Assessing the socioeconomic status in the German Pharmacoepidemiological Research Database (GePaRD): Description and exemplary application using the association with obesity]. GMS Med Inform Biom Epidemiol. 2022;18(1):Doc02. DOI: 10.3205/mibe000235

Figure 8: Prevalence of obesity in GePaRD for 2017 stratified by age, sex, and information on education missing vs. available

