Neolithic life tables de

Providing and presenting data on human age at death.

Christoph Rinne*

Version 2.1 (2024-11-20)

Contents

1 Introduction

The dataset contains mortality data of humans taken from the literature. Therefore, the data is heterogeneous in many respects, especially with regard to the quality of the anthropological information. However, the data is simplified and classified to fit into a limited number of columns, e.g. a general archaeological classification.

I collected the data to get a general impression of the differences in mortality during the Neolithic in different communities in relation to archaeological periods and regions within Germany. I did not intend to build up a comprehensive database of all mortality data, nor to create a well-founded collection of information from physical anthropology. Due to the easy availability of the data, I also integrated data from an Early Neolithic cemetery in Austria.

Important note: Most anthropologists will reject extrapolations, especially of life expectancy, based on this data. The attribute "(archaeological) culture" serves as a label for a number of features related to the way of life, e.g. houses, settlement structure, economy and burial practices. The attribute "period" is based on the previous entity. Both must be placed in an archaeological framework and a current absolute dating, depending on the research question. Please bear in mind that collective burials can be used over a long period of time.

The data is provided as a sqlite database in a very simple form, not normalised or as implemented relations. This documentation is provided as an R-markdown with a resulting pdf to give a first overview of the data in the database. At the same time, a simple workflow for data processing in R with the R package [mortAAR] (https://cran.r-project.org/web/packages/mortAAR/index.html) for calculating life tables is described. The literature used and cited is provided as a Bibtex file.

The data collection is available via the LandMan portal of CRC 1266[1]. Subproject: "Regional and Local Patterns of 3rd Millennium Transformations of Social and Economic Practices in the Central German Mountain Range (D2)" https://gepris.dfg.de/gepris/projekt/316739879]. In addition, I have decided to set up this Github repository to make updates traceable and further additions to the database more easily accessible https://github.com/chrinne/lifetables4archeologists.

*Kiel University, crinne@ufg.uni-kiel.de

[1]"Scales of Transformation - Human-Environmental Interaction in Prehistoric and Archaic Societies." Deutsche Forschungsgemeinschaft (DFG) - project number 290391021 https://gepris.dfg.de/gepris/projekt/290391021

2 Data presentation

All code chunks will be visible, this is part of the reproducibility. The documentation uses some R packages:

```
knitr::opts_chunk$set(echo = TRUE, include = TRUE)
require(pacman) || install.packages("pacman")
```

```
## Lade nötiges Paket: pacman
```

```
## [1] TRUE
```

```
pacman::p_load(dplyr, ggplot2, knitr, mortAAR, RSQLite)
```

Please set the working directory to the folder with the data, e.g.:

```
setwd('d:/data/folder/')
```

Set up the database connection.

```
db01 <- dbConnect(RSQLite::SQLite(), dbname = "./de_lifetable_neol.sqlite")
```

2.1 Data structure

The data is provided in four tables: one table with cited literature, two tables for individual and tabled mortality data and one table with cooridinates for the nearby places. The metadata of the data is saved in an additional table.

```
select name  as 'Name' from sqlite_master where type = 'table' order by 1;
```

Table 1: 5 records

Name
citations
lifedata_ind
lifedata_tbl
metadata
places

Import the data into the R environment.

```
tables<- dbGetQuery(db01, "select name from sqlite_master
                    where type = 'table' order by 1;")[,1]
for (tbl in tables){
  assign(tbl, dbReadTable(db01, tbl))
}
```

Most sqlite columns are set to TXT due to type affinity in sqlite, one exception is 'count'. You might need to adjust this.

2.1.1 lifetable_ind and lifetable_tbl

The structure of the mortality data tables for individual and tabled data is identical for the easy combination (*rbind()*). The column names are self-explanatory in general.

```
paste(colnames(get(tables[2])), collapse = ", ")
```

[1] "place, site, grave, sex.orig, sex, from, to, ageclass.orig, count, culture, period, literature, comment"

The column 'count' is always 1 for individual data and can provide a float for tabled data due to interpolation in the original life tables. The column 'sex' provides a simplified English version of the original sex determination in the column 'sex.orig'. The columns 'from' and 'to' are derived from the original age classification in 'ageclass.orig' to fulfil the needs of mortAAR.

2.1.2 places

The table 'places' provides coordinates (WGS84) of nearby places to allow spatial differentiation. The 'id' is related to 'places' in the lifetables.

```
paste(colnames(get(tables[5])), collapse = ", ")
```

[1] "id, name, latitude, longitude"

2.2 Content of places

The nearby places are distributed from Hildesheim south of Hannover to the Alps and from the Rhine to Straubing in Southeast Bavaria. The site of Kleinhadersdorf is an outlier in Eastern Austria. You can notice a significant agglomeration of places and thus of nearby sites around the Harz region.

```
plot(places$longitude, places$latitude, cex = 0.2,
     xlab = "Longitude", ylab = "Latitude")
text(places$longitude, places$latitude, places$name, cex = 0.6)
```

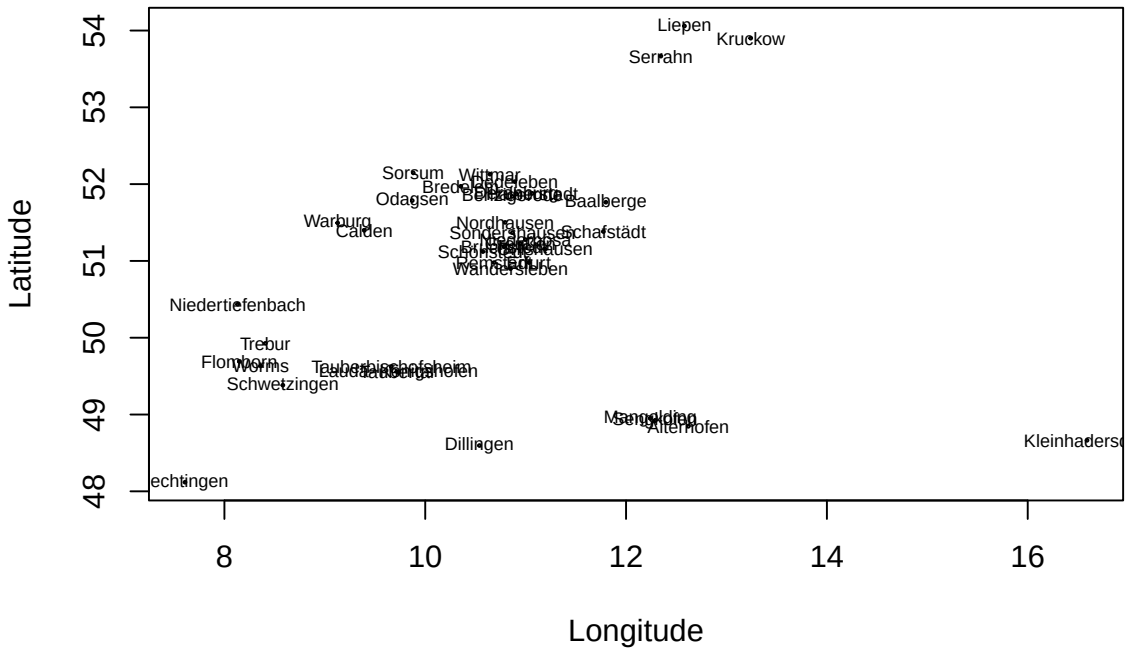


Figure 1: Plot of nearby places.

```
paste(places[,2], collapse = ", ")
```

[1] "Aiterhofen, Baalberge, Benzigerode, Bredelem, Bruchstedt, Calden, Dedeleben, Derenburg, Dillingen, Erfurt, Flomborn, Greußen, Halberstadt, Jechtingen, Kleinhadersdorf, Mangolding, Niederbösa, Niedertiefenbach, Nordhausen, Odagsen, Orlishausen, Remstädt, Schafstädt, Schwetzingen, Schönstedt, Sengkofen, Sondershausen, Sorsum, Tauberbischofsheim, Taubertal, Trebur, Wandersleben, Warburg, Wittmar, Worms, Lauda-Königshofen, Kruckow, Liepen, Serrahn"

In some cases, the tabular data is derived from regional summaries, e.g. for the Late Neolithic Corded Ware in the Tauber Valley or the Middle Neolithic Baalberge burials from a region, localised with the eponymous site.

2.3 Content of life tables

The general overview is comprised of two tables for individual and tabled data.

```
ld <- rbind(lifedata_ind, lifedata_tbl)
ld$from <- as.numeric(ld$from)
ld$to <- as.numeric(ld$to)
```

There are cases where I first entered the tabulated data and then the individual data. Both datasets have advantages and disadvantages, which is why both are offered. A further reason for multiple recorded sites is the re-determination or the recording on the basis of different bones, e.g. in collective graves. Unfortunately, the Literature field cannot be counted directly here, as it also refers to individual catalogue numbers. It must therefore be counted on a short citation, which then excludes multiple entries in a single publication, e.g. cranial and pelvic age estimation in Sorsum by Moser et al. 2024.

```
ld %>%
  # individual citations truncated to the root element
  mutate(., short_citation = sub(",.*", "", literature)) %>%
  group_by(., site) %>%
  summarise(., recordings = n_distinct(short_citation)) %>%
  filter(., recordings > 1) %>%
  knitr::kable(caption = "Sites with multiple recordings.")
```

Table 2: Sites with multiple recordings.

site	recordings
Schwetzingen	2
Sorsum	2

This reference highlights a further characteristic of this data set: the survey of different bones inevitably leads to different numbers of individuals that do not necessarliy correspond to the minimum number of individuals (MNI) found at the site (cf Sorsum counted individuals). To understand this data set correctly: **It is a summary of opinions on age and gender characteristics, not necessarily the identification of individuals.**

Neolithic periods are abbreviated: **E**arly, **M**iddle, **Y**ounger, **L**ate and **F**inal.

```
ld$period <- factor(ld$period, levels = c("EN","MN", "YN", "LN","FN"))
ld %>%
  group_by(., period) %>%
  summarise(count = sum(count)) %>%
  knitr::kable(caption = "Number of age estimations in each period.",
               col.names = c("Period", "Count"))
```

Table 3: Number of age estimations in each period.

Period	Count
EN	1129.80
MN	258.00
YN	233.98
LN	1269.68
FN	357.00

Archaeological cultures per period.

```
ld %>%
  group_by(., period, culture) %>%
  summarise(count = sum(count), .groups = 'drop') %>%
```

```
  knitr::kable(caption = "Number of age estimations in each archaeological culture.",
               col.names = c("Period", "Culture", "Count"))
```

Table 4: Number of age estimations in each archaeological culture.

Period	Culture	Count
EN	LBK	1129.80
MN	Großgartach	49.00
MN	Hinkelstein	112.00
MN	Rössen	2.00
MN	Rössen, Großgartach	95.00
YN	Baalberge	233.98
LN	Bernburg	415.70
LN	TRB	563.98
LN	Wartberg	290.00
FN	Corded Ware	357.00

Number of age estimations per simplified sex determination. The high proportion of undetermined sex results from the tabular data and the subadult individuals.

```
ld %>%
  group_by(., sex) %>%
  summarise(count = sum(count)) %>%
  knitr::kable(caption = "Number of age estimations per sex.",
               col.names = c("Sex", "Count"))
```

Table 5: Number of age estimations per sex.

Sex	Count
female	708.10
indet	1852.36
male	688.00

Sites, related citations and count of age estimations.

```
ld %>%
  mutate(., short_citation = sub(",.*", "", literature)) %>%
  group_by(., site, short_citation) %>%
  summarise(count = sum(count), .groups = 'drop') %>%
  knitr::kable(caption = "Number of age estimations per site",
               col.names = c("Site", "Literature", "Count"))
```

Table 6: Number of age estimations per site

Site	Literature	Count
Aiterhofen-Ötmühlen	Baum 1990	98.000
Baalberge-cemeteries	Funke 2006	58.990
Baalberge-grave-sites	Funke 2006	174.990
Benzigerode	Berthold et al. 2008	43.000
Bredelem	mortAAR::gallery_graves	13.000
Bruchstedt	Bach 1978	49.000
Calden I	mortAAR::gallery_graves	24.000
Calden II	Pasda 2000	110.000
Dedeleben	Gall et al 1983	23.000
Derenburg	Bach 1981	90.000

Site	Literature	Count
Derenburg-Meerenstieg	Fritsch et al. 2011	48.000
Dillingen-Steinheim	Nieszery 1995	27.000
Erfurt-Nordhäuser Str.	Bücke et al. 1989	13.000
Flomborn	Richter 1968/69	32.000
Greußen-LPG-Kiesgrube	Feustel et al 1966	7.000
Halberstadt-Sonntagsfeld	Fritsch et al. 2011	42.000
Jechtingen-Humbergäcker	Alt et al. 2014	95.000
Kleinhadersdorf	Neugebauer-Maresch/Lenneis 2015	51.000
Kruckow 69 37	Grimm 1983	1.000
Kruckow 69,37	Grimm 1983	1.000
Kruckow 69/37	Grimm 1983	12.000
Liepen 65/28	Grimm 1983	31.000
Liepen 65/29	Grimm 1983	2.000
Liepen 65/30	Grimm 1983	16.000
Mangolding	Nieszery 1995	13.000
Niederbösa	Feustel/Ullrich 1964/65	99.700
Niedertiefenbach	mortAAR::gallery_graves	50.000
Nordhausen	Feustel/Ullrich 1964/65	50.000
Odagsen	Rinne 2003	102.980
Orlishausen/Otterzunge	Bach et al 1975	12.000
Remstädt	Fuchs 2013	11.000
Schafstädt	Hummel 2000	70.000
Schwetzingen	Gerling 2012	211.000
Schwetzingen	Gerling/Francken 2007	210.800
Schönstedt	Bach / Bach 1972	64.000
Sengkofen	Nieszery 1995	31.000
Serrahn 64/24	Grimm 1983	1.000
Serrahn 64/25	Grimm 1983	24.000
Sondershausen	Bach 1978	59.000
Sorsum	Moser et. al. 2024	318.000
Sorsum	mortAAR::gallery_graves	42.000
Tauberbischofsheim-Dittigheim	Dresely 2004	58.004
Tauberbischofsheim-Impfingen	Dresely 2004	32.992
Taubertal	Dresely 2004	58.004
Trebur	Spatz 1999	124.000
Wandersleben	Gall et al 1983	46.000
Wandersleben-Cobstädter Str.	Bach 1986	246.000
Warburg I	Löwen 1997	32.000
Warburg III	Löwen 1997	44.000
Warburg IV	Löwen 1997	30.000
Wittmar	Rötting 1983	15.000
Worms-Rheindürkheim	Meier-Arendt 1975	17.000
Worms-Rheingewann	Meier-Arendt 1975	20.000
Wöllerspfad	Trautmann 2012	94.000

The age ranges depend mainly on the traditional age groups and the usual 5-year-intervals in life tables. Two different illustrations can show the dependence between the determined age at death and the accuracy of the age range.

```
ld %>%
  select(., from, to) %>%
  filter(., !is.na(from)) %>%
  filter(., !is.na(to)) %>%
  mutate(., range = to - from) %>%
  mutate(., agegroup = ifelse(to < 20, 'subadult', 'adult')) %>%
  arrange(., from, to) -> ages
```

```
id <- seq(nrow(ages))
plot(c(0,100), c(0,nrow(ages)), xlab = "Age from to",
     ylab = "age estimations", type = "n") +
segments(ages$from[id], id, ages$to[id], id)
```

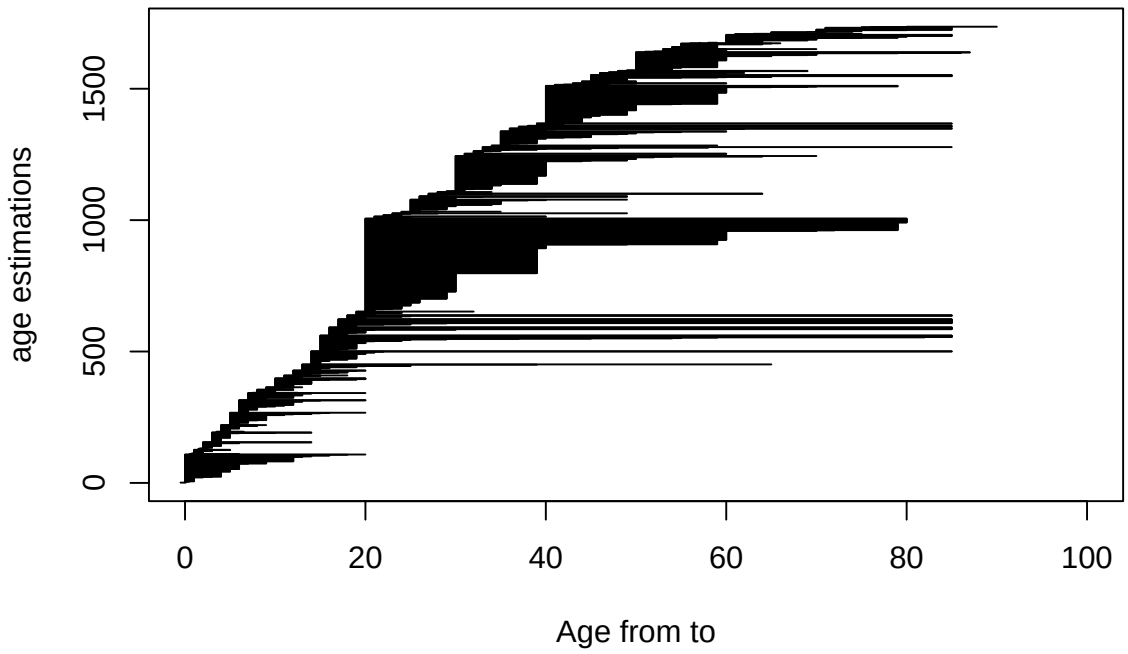


Figure 2: Age range per estimation.

```
## integer(0)
```

```
boxplot(range~agegroup, data = ages, xlab = "Age group", ylab = "Age range")
```

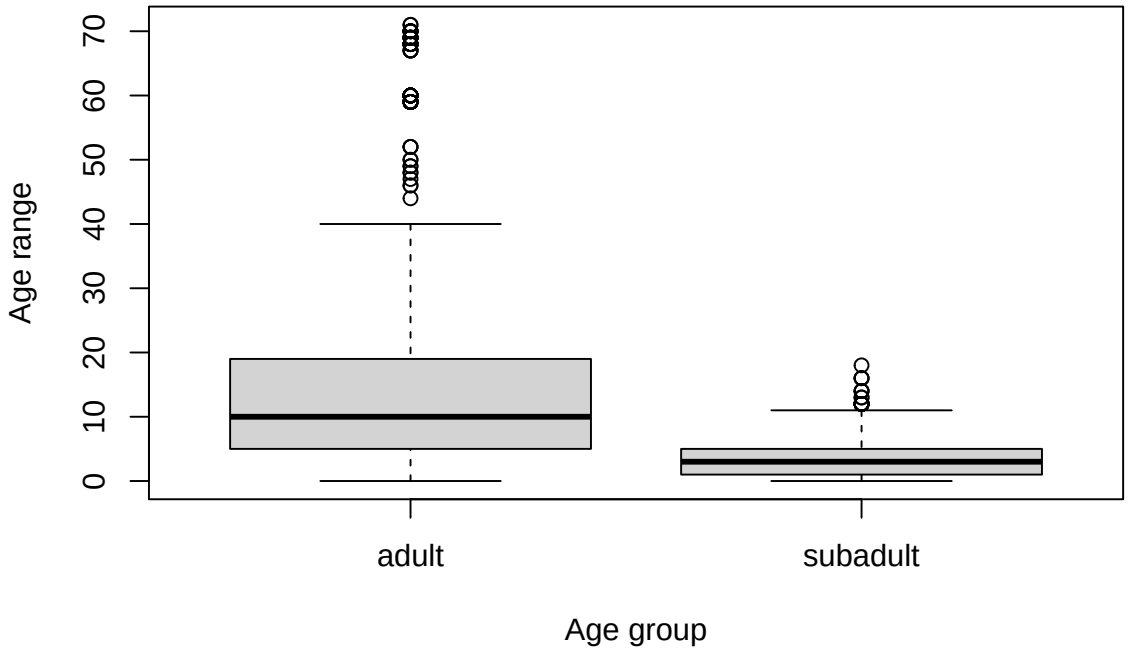


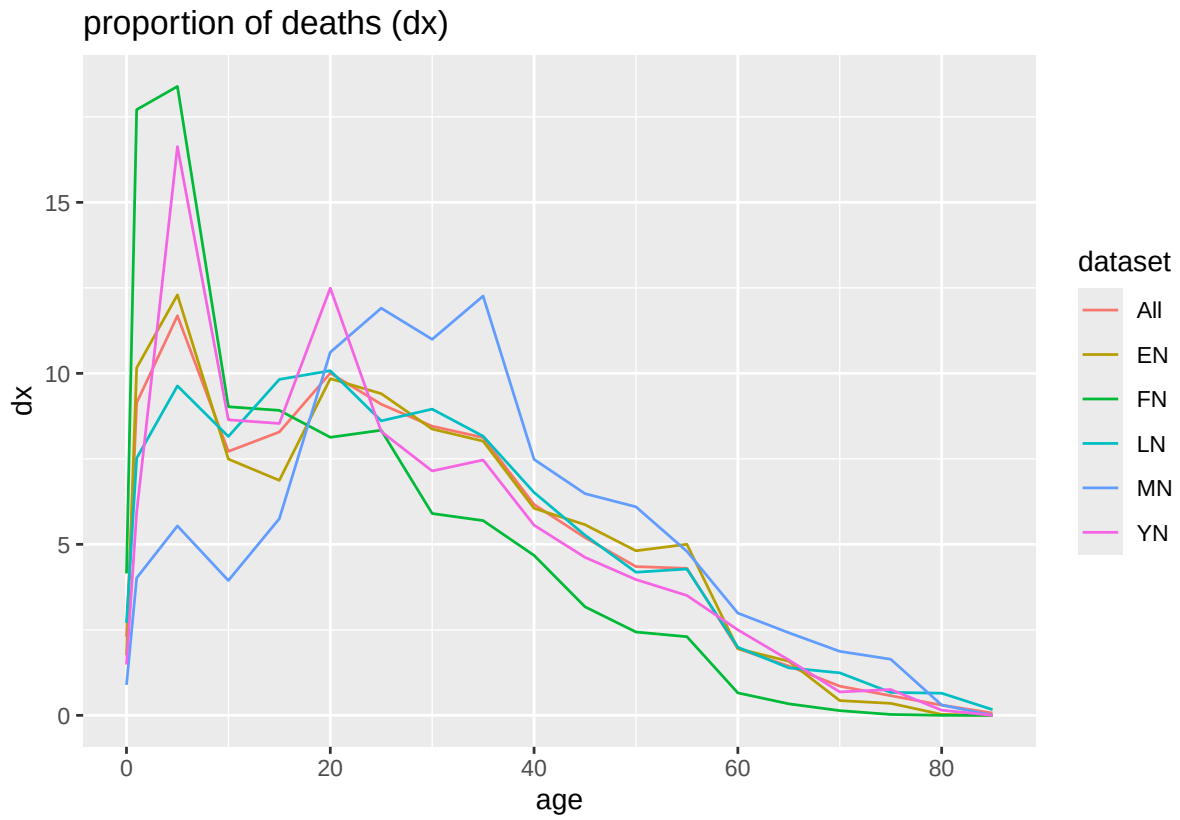
Figure 3: Age range for subadult (< 20) and adult individuals.

2.4 Overview of age at death (dx).

The following plots are created with the R-package mortAAR[2] without an interpolation (method = standard) for various groupings.

2.4.1 Periods

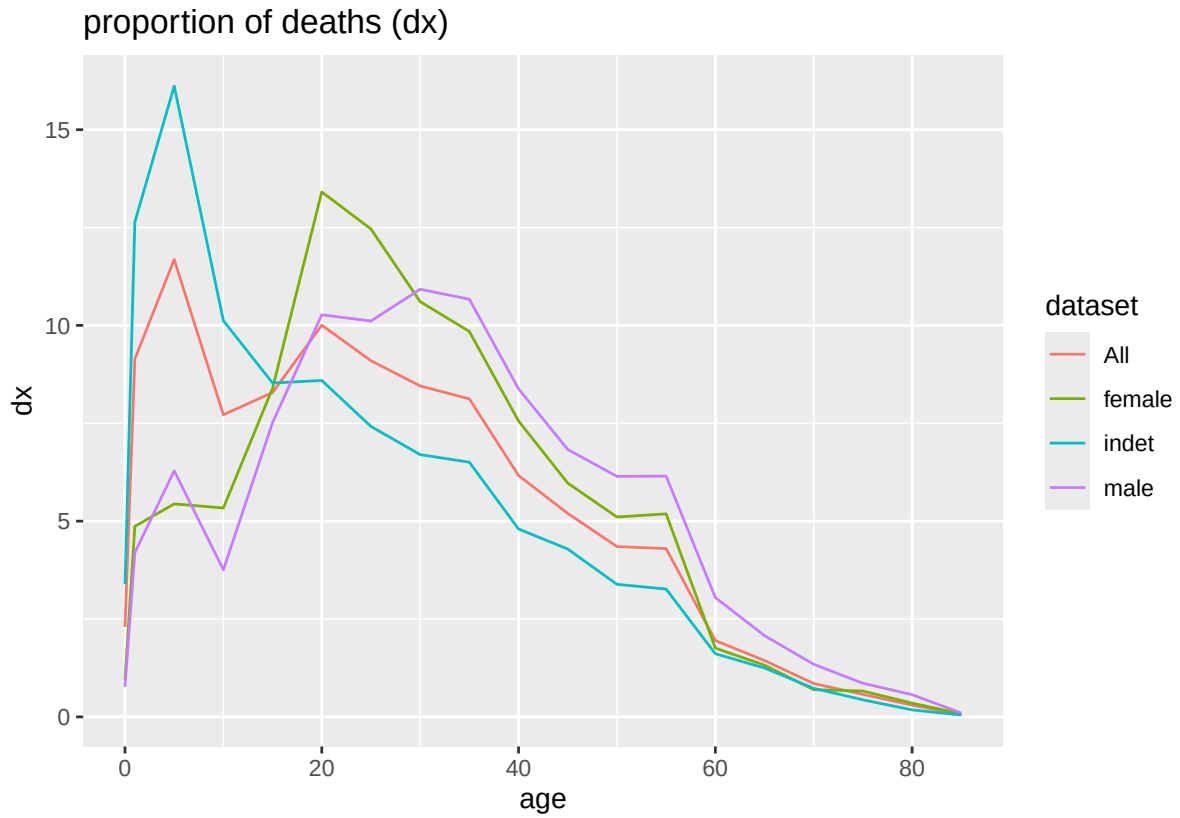
```
prep.life.table(ld, dec = "count", agebeg = "from", ageend = "to",
  group = "period", method = "standard", agerange = "included") %>%
  life.table(.) %>%
  plot(., display = c("dx"), line_vis="color")
```



2.4.2 Sex

```
prep.life.table(ld, dec = "count", agebeg = "from", ageend = "to",
  group = "sex", method = "standard", agerange = "included") %>%
  life.table(.) %>%
  plot(., display = c("dx"), line_vis="color")
```

[2]If you observe strange plot lines in the ggplots of mortAAR, please consider the installation of the latest version of mortAAR from github.

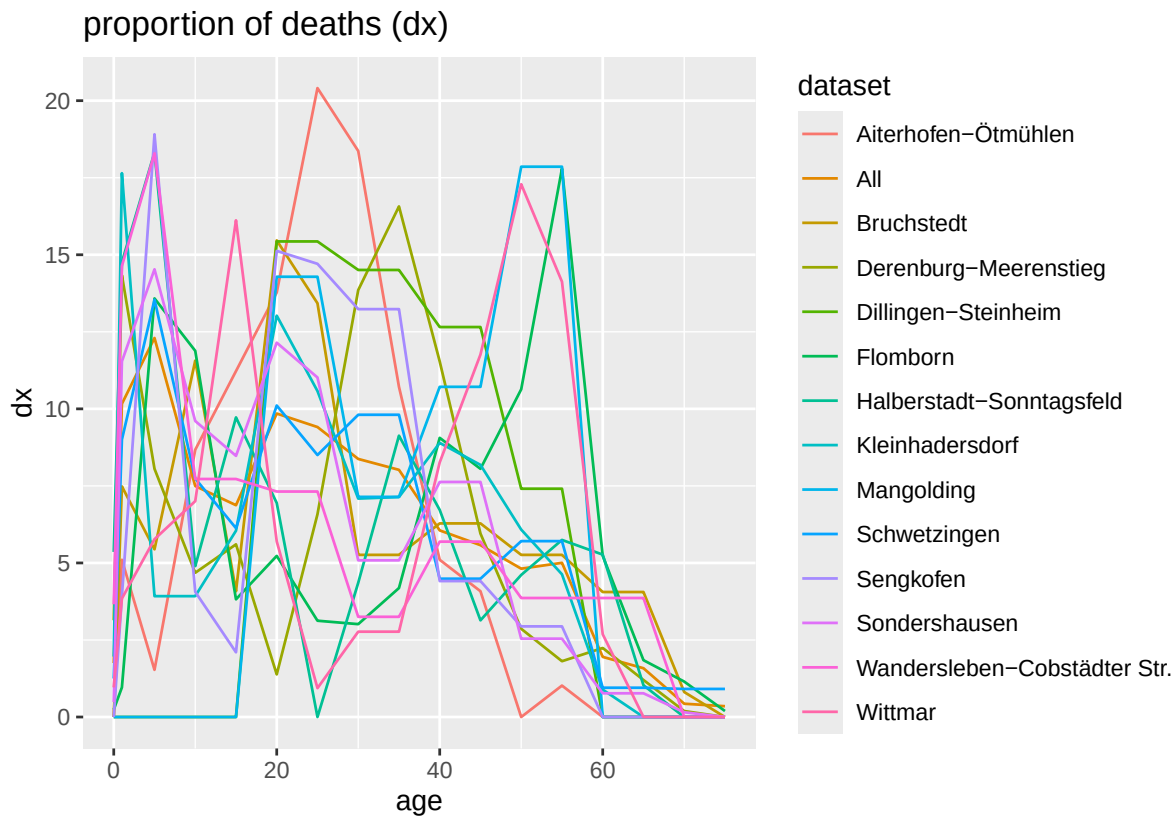


2.4.3 Sites per period

The following diagrams show very heterogeneous data. It is obvious that each data set has its own weaknesses and requires careful and detailed analysis. Furthermore, this does not rule out significant differences in their relation to previous living conditions.

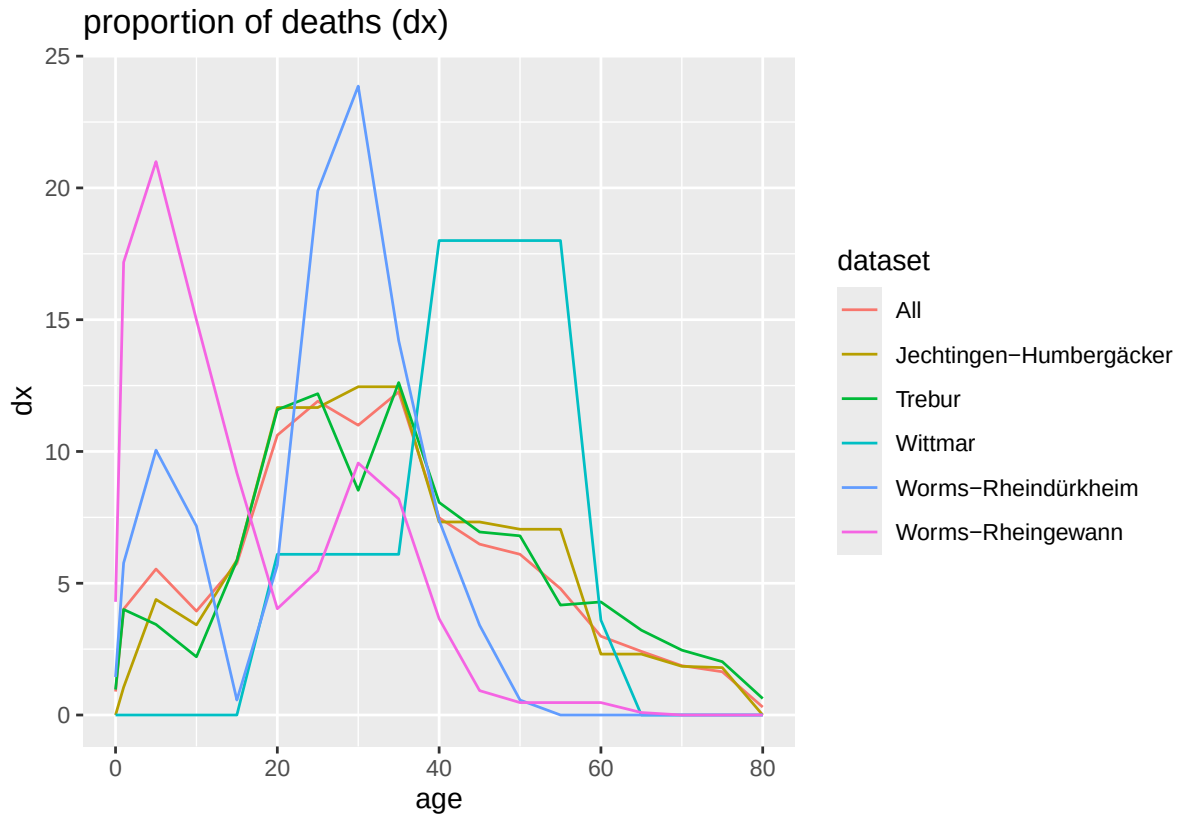
The sites from the **Early Neolithic**.

```
ld %>%
  filter(.,period == "EN") %>%
  prep.life.table(., dec = "count", agebeg = "from", ageend = "to",
  group = "site", method = "standard", agerange = "included") %>%
  life.table(.) %>%
  plot(., display = c("dx"), line_vis="color")
```



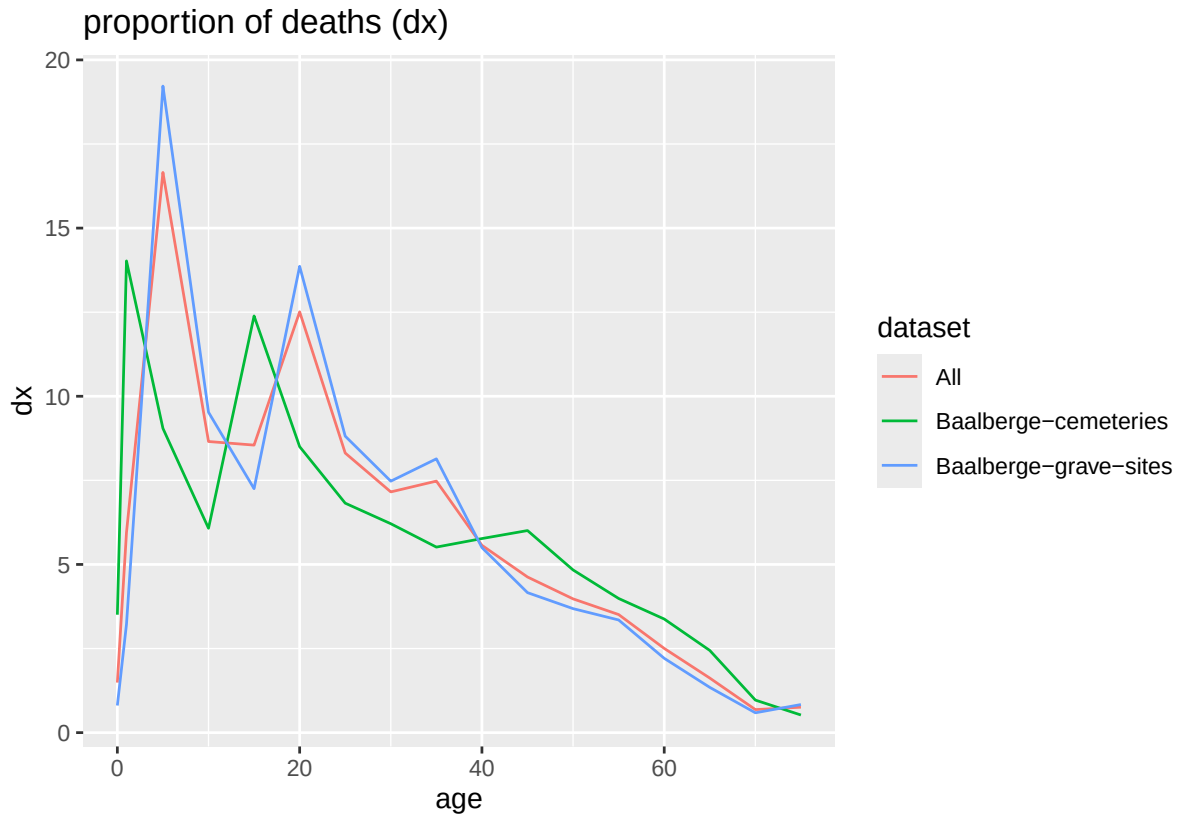
The sites from the **Middle Neolithic**.

```
ld %>%
  filter(.,period == "MN") %>%
  prep.life.table(., dec = "count", agebeg = "from", ageend = "to",
  group = "site", method = "standard", agerange = "included") %>%
  life.table(.) %>%
  plot(., display = c("dx"), line_vis="color")
```



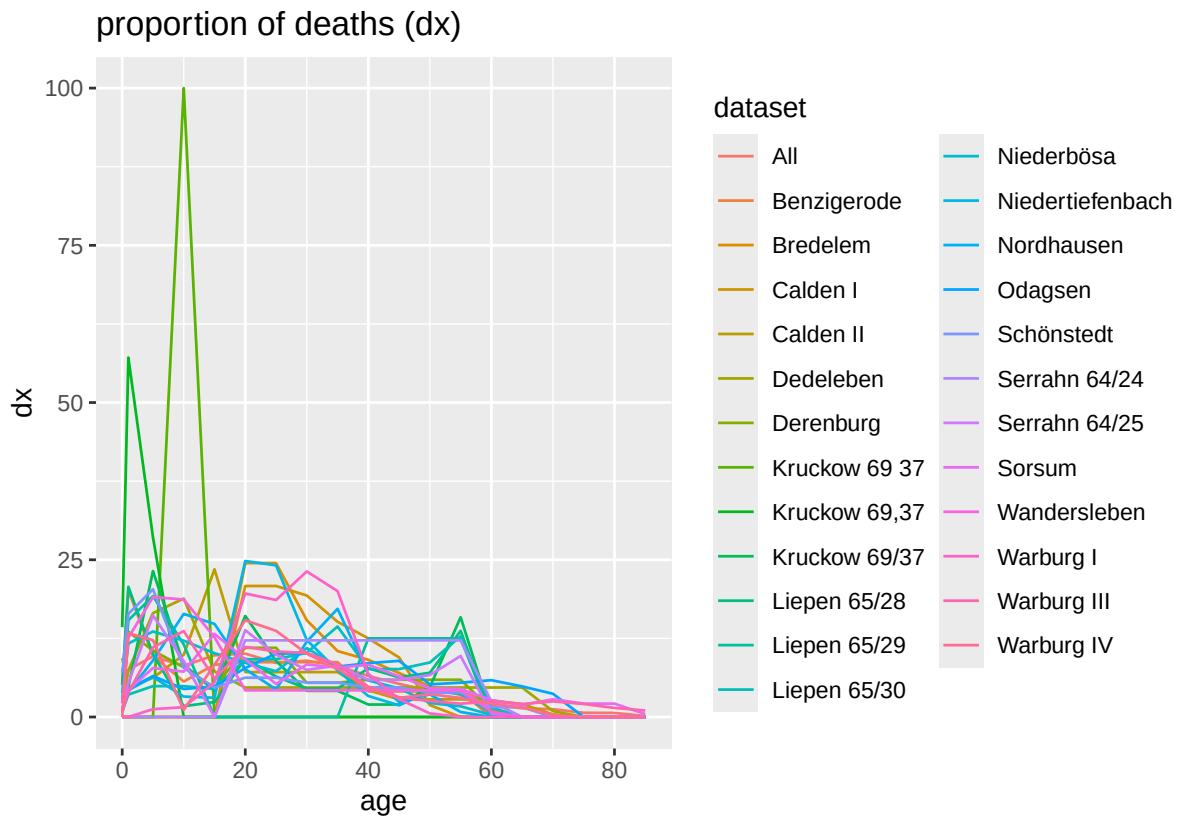
The sites from the **Younger Neolithic**.

```
ld %>%
  filter(.,period == "YN") %>%
  prep.life.table(., dec = "count", agebeg = "from", ageend = "to",
  group = "site", method = "standard", agerange = "included") %>%
  life.table(.) %>%
  plot(., display = c("dx"), line_vis="color")
```



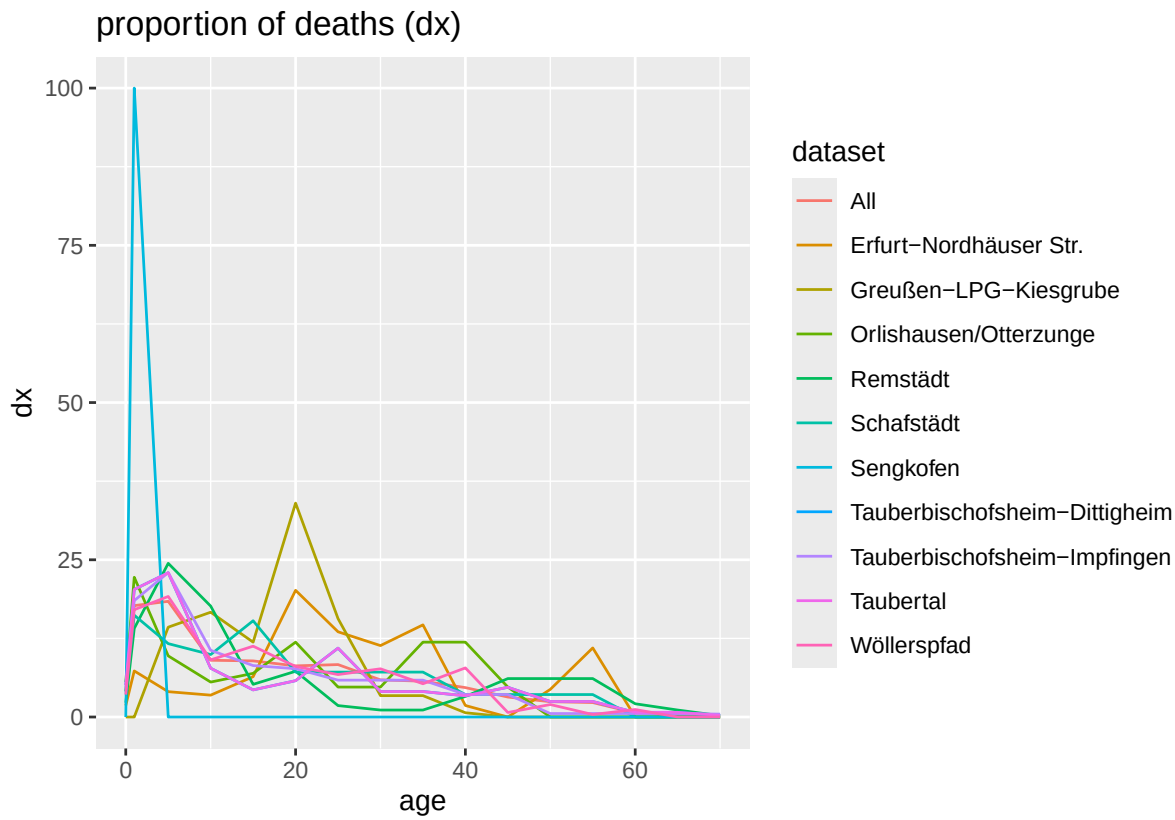
The sites from the **Late Neolithic**.

```
ld %>%
  filter(.,period == "LN") %>%
  prep.life.table(., dec = "count", agebeg = "from", ageend = "to",
  group = "site", method = "standard", agerange = "included") %>%
  life.table(.) %>%
  plot(., display = c("dx"), line_vis="color")
```



The sites from the **Final Neolithic**.

```
ld %>%
  filter(.,period == "FN") %>%
  prep.life.table(., dec = "count", agebeg = "from", ageend = "to",
  group = "site", method = "standard", agerange = "included") %>%
  life.table(.) %>%
  plot(., display = c("dx"), line_vis="color")
```


proportion of deaths (dx)
dx
100
75
50
25
0
0
20
40
60
age
dataset
All
Erfurt−Nordhäuser Str.
Greußen−LPG−Kiesgrube
Orlishausen/Otterzunge
Remstädt
Schafstädt
Sengkofen
Tauberbischofsheim−Dittigheim
Tauberbischofsheim−Impfingen
Taubertal
Wöllerspfad