Analysis

Nathan Green

17/03/2020

The raw annual risk of death is calculated by the proportion of people in each age, sex, ethnicity group who die each year (with those who emigrate being censored, of course). Separate survivorship curve for each year of birth, as life expectancy changes over time. There is sex differences; ethnicity doesn't make much of a difference.

ETHPOP is based on ONS data in what year?

```
library(readr)
library(purrr)
library(dplyr)
library(ggplot2)
library(scales)
library(reshape2)
library(survivorETHPOP)
```

1 ETHPOP

We want to compare between the ONS mortality statistics and ETHPOP. From here it details their method. They calculate a *central rate of mortality* as the average across 3 years.

$$m_x = \sum_{y1,y2,y3} deaths_i / \sum_{y1,y2,y3} pop_i$$

Finally, they calculate the *mortality rate* which is what we will be using to compare and is equivalent to the hazard.

$$q_x = 2m_x/(2 + m_x)$$

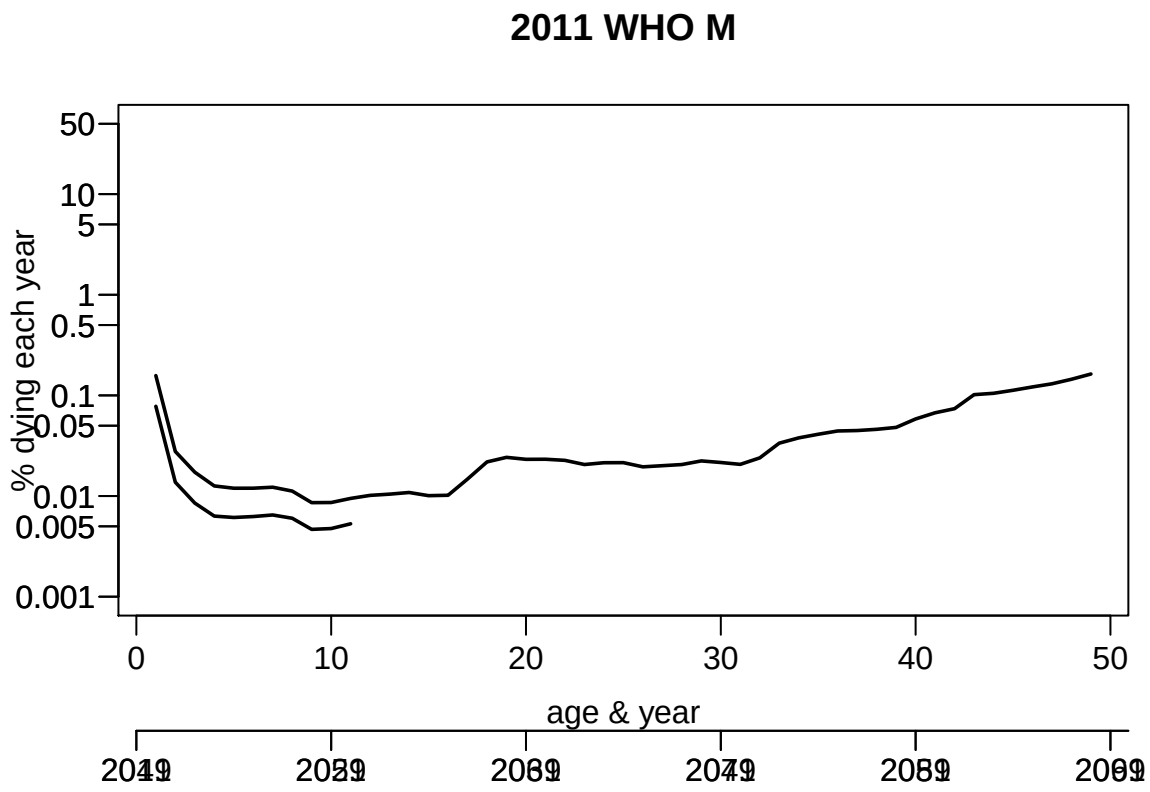
1.1 Individual categories hazards and survival

```
ETHPOP_lifetable <- make_ETHPOP_lifetable()
# save(ETHPOP_lifetable, file = here::here("data", "ETHPOP_lifetable.RData"))


head(ETHPOP_lifetable)
```

```
FALSE # A tibble: 6 x 16
FALSE   ETH.group   age sex   deaths  year    pop    id yr_age death_rate      mx      qx     Lx      T
FALSE   <chr>     <dbl> <chr>  <dbl> <dbl>  <dbl> <int> <chr>       <dbl>   <dbl>   <dbl>  <dbl>   <dbl
FALSE 1 BAN           0 F       9.37  2011  4844.   101 2011_0    0.00193 0.00193 0.00193  4826. 243722
FALSE 2 BAN           0 M      10.7   2011  4884.   101 2011_0    0.00219 0.00219 0.00219  4849. 250641
```

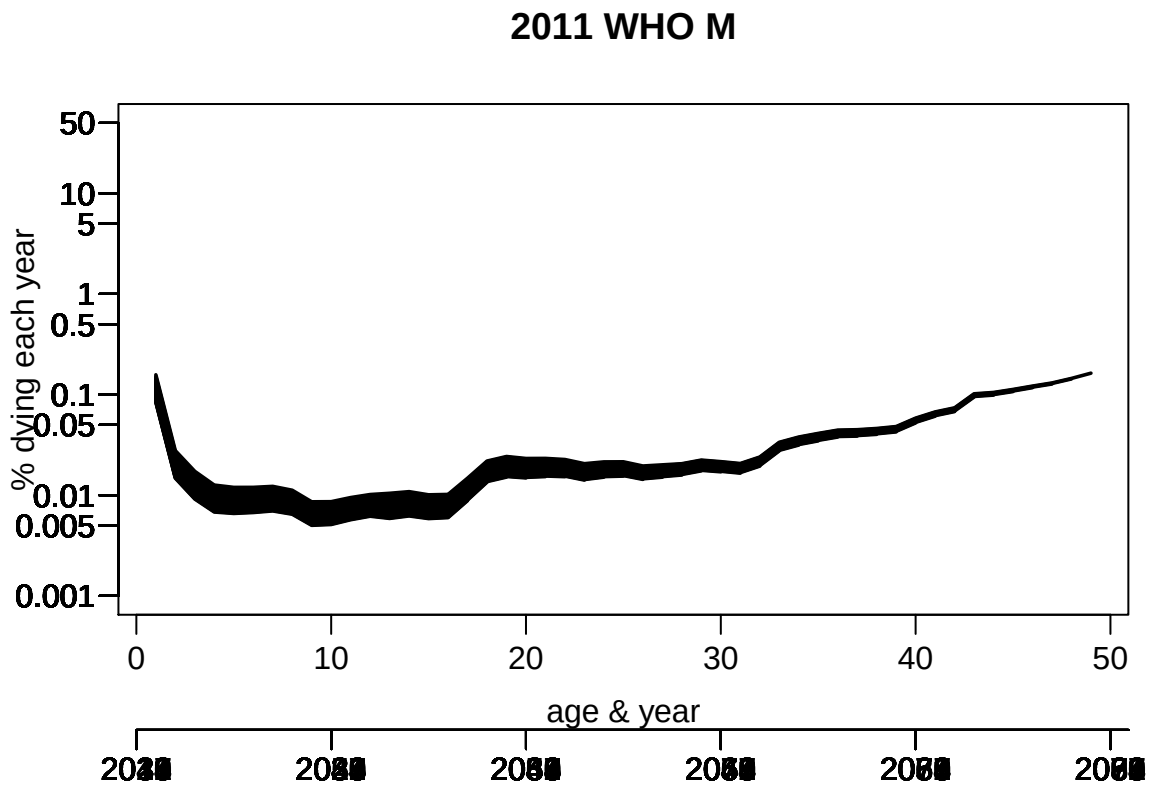
```
FALSE 3 BLA           0 F      16.4   2011 10529.   101 2011_0    0.00156 0.00156 0.00156 10686. 711828
FALSE 4 BLA           0 M      19.2   2011 10806.   101 2011_0    0.00178 0.00178 0.00178 10889. 777737
FALSE 5 BLC           0 F       4.91  2011  2728.   101 2011_0    0.00180 0.00180 0.00180  2681. 147650
FALSE 6 BLC           0 M       5.25  2011  2736.   101 2011_0    0.00192 0.00192 0.00192  2673. 149569
```

```
ETHPOP_lifetable %>%
  survivor_curve(group = list(sex = "M",
                              ETH.group = "WHO",
                              year = 2011)) %>%
  haz_plot()

ETHPOP_lifetable %>%
  survivor_curve(group = list(sex = "M",
                              ETH.group = "WHO",
                              year = 2049)) %>%
  haz_plot(add = TRUE)
```

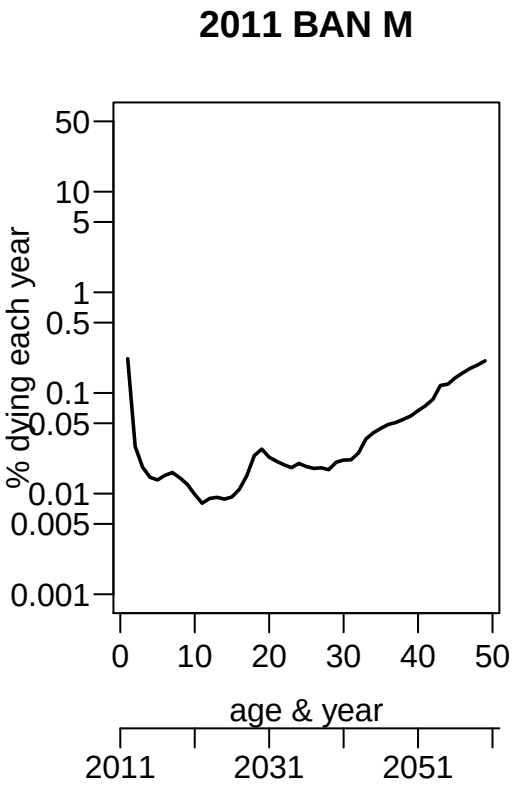


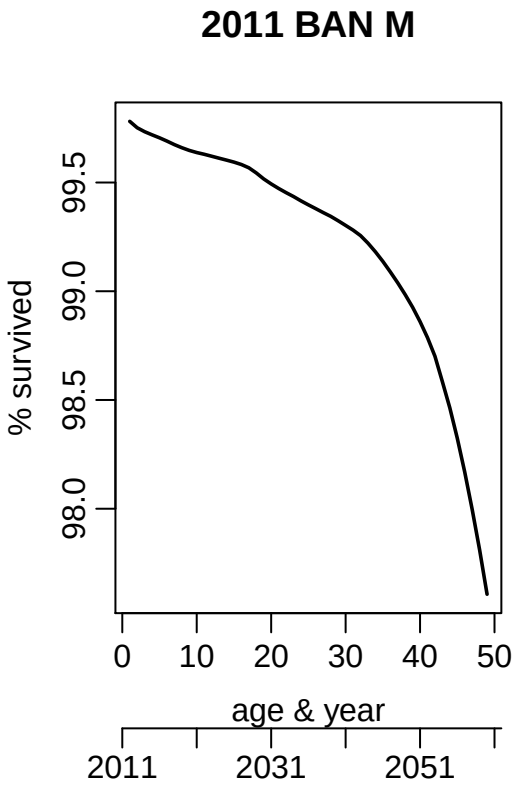

```
ETHPOP_lifetable %>%
  survivor_curve(group = list(sex = "M",
                              ETH.group = "WHO",
                              year = 2011)) %>%
  haz_plot()

for (i in 2012:2043) {
  ETHPOP_lifetable %>%
    survivor_curve(group = list(sex = "M",
                                ETH.group = "WHO",
```

```
                                year = i)) %>%
    haz_plot(add = TRUE)
}
```



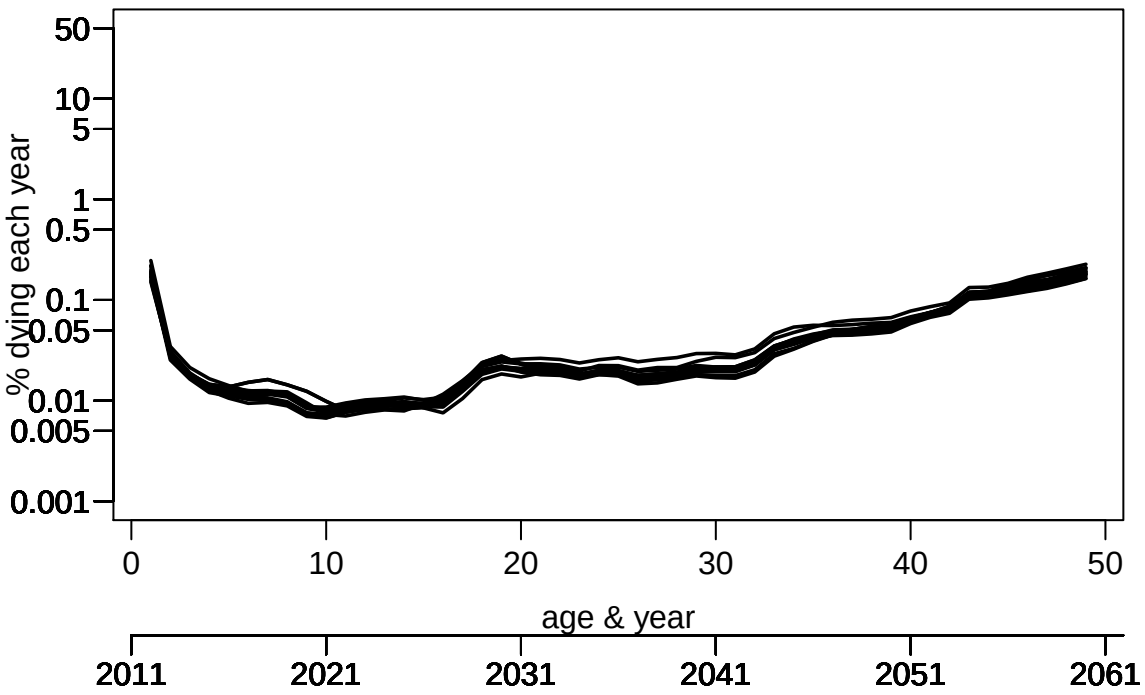

```
par(mfrow=c(1,2))
dat <- ETHPOP_lifetable %>%
  survivor_curve(group = list(sex = "M",
                              ETH.group = "BAN",
                              year = 2011)) %>%
  haz_plot() %>%
  surv_plot()
```





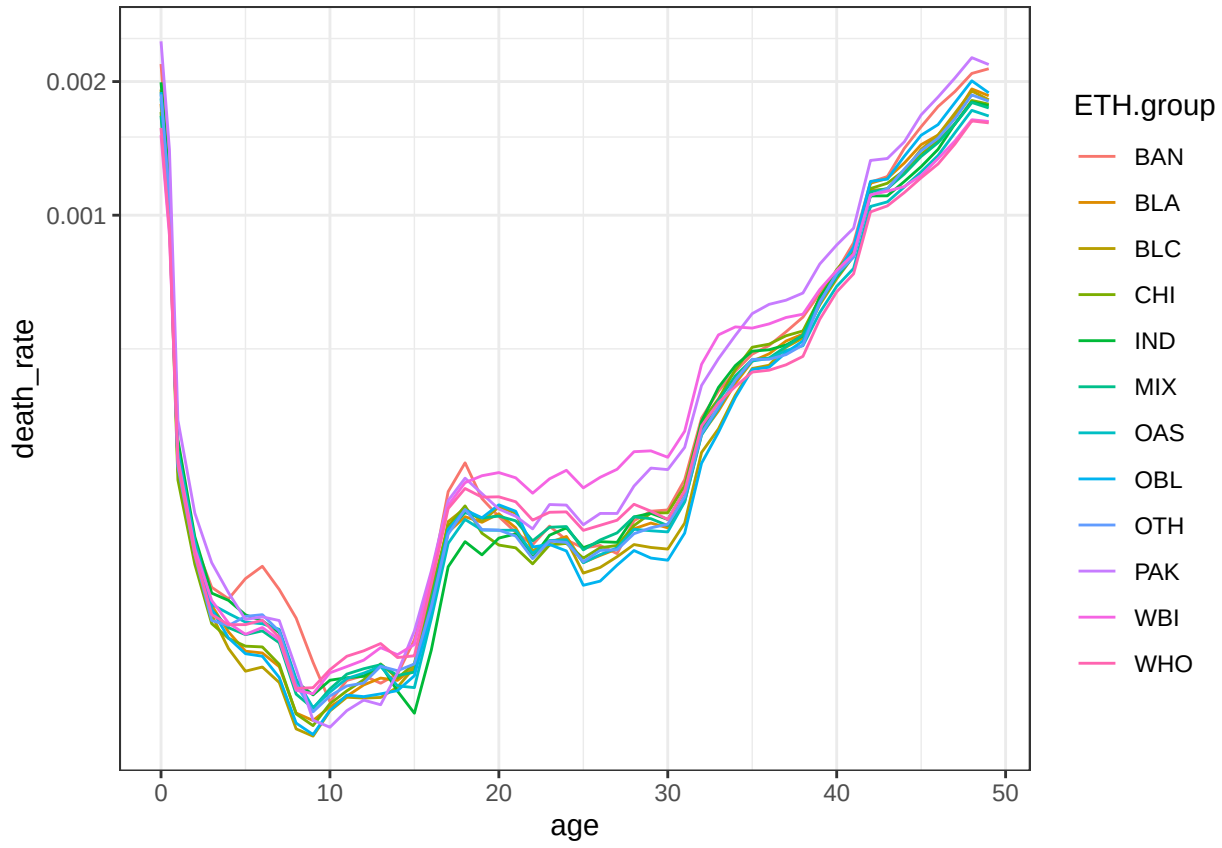
1.1.1 Ethnic groups

```
ethnic_grps <- c("BAN", "BLA", "BLC", "CHI", "IND", "MIX",
                 "OAS", "OBL", "OTH", "PAK", "WBI", "WHO")

ETHPOP_lifetable %>%
  survivor_curve(group = list(sex = "M",
                              ETH.group = "BAN",
                              year = 2011)) %>%
  haz_plot()

for (i in ethnic_grps) {
  ETHPOP_lifetable %>%
    survivor_curve(group = list(sex = "M",
                                ETH.group = i,
                                year = 2011)) %>%
    haz_plot(add = TRUE)
}
```

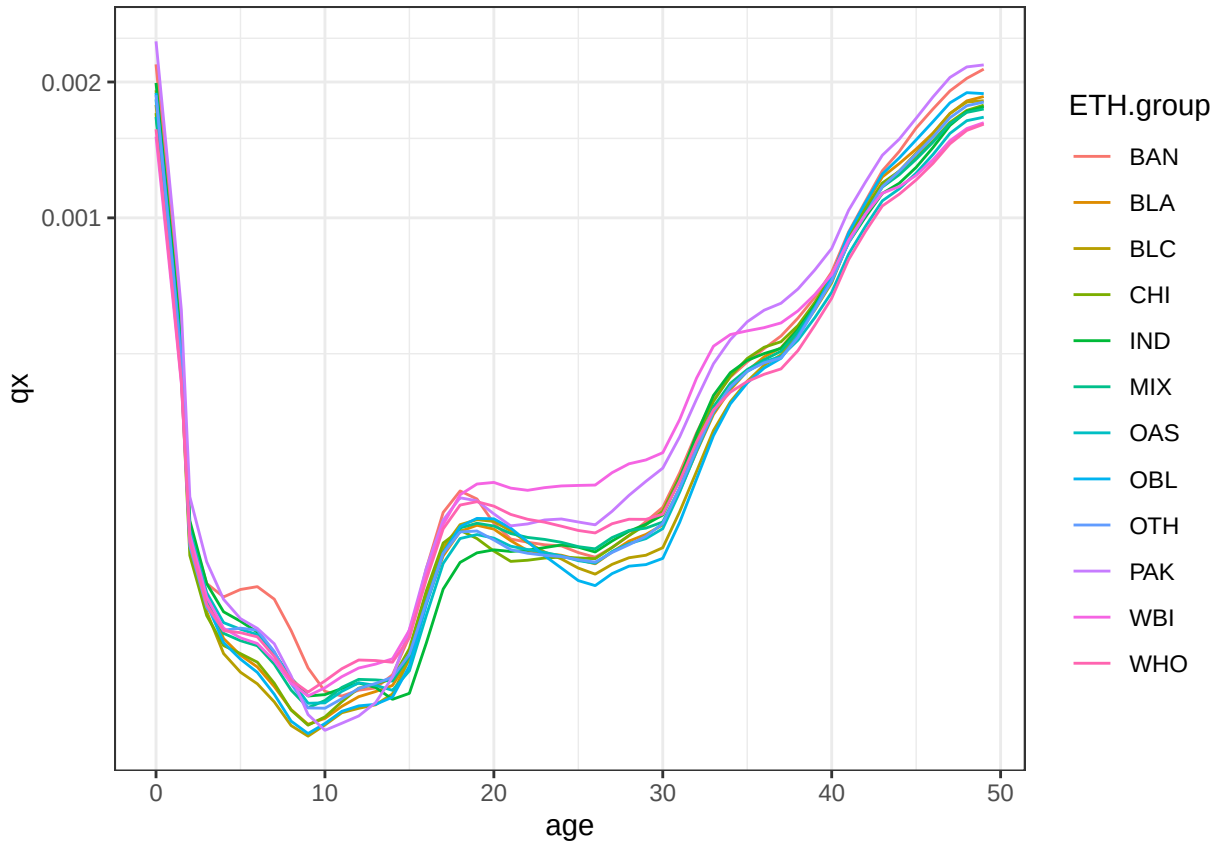



```
baseyr_2011 <- ETHPOP_lifetable$id[ETHPOP_lifetable$yr_age == "2011_0"][1]

ETHPOP_lifetable_2011M <- ETHPOP_lifetable %>%
  filter(id == baseyr_2011,
         sex == "M")

ETHPOP_lifetable_2011F <- ETHPOP_lifetable %>%
  filter(id == baseyr_2011,
         sex == "F")
```

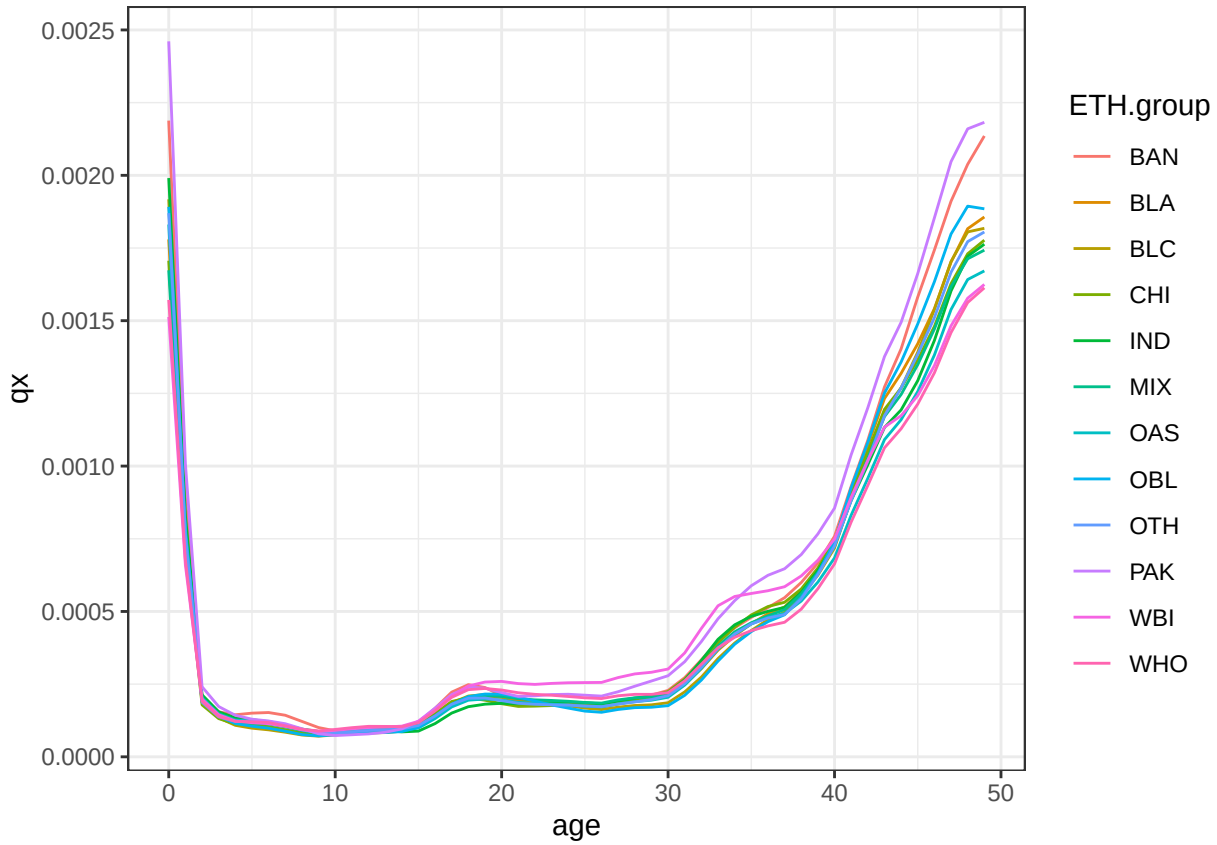
```
## ggplot

ggplot(ETHPOP_lifetable_2011M, aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  # scale_y_continuous(trans='log2') +
  coord_trans(y = "log10") +
  theme_bw()
```



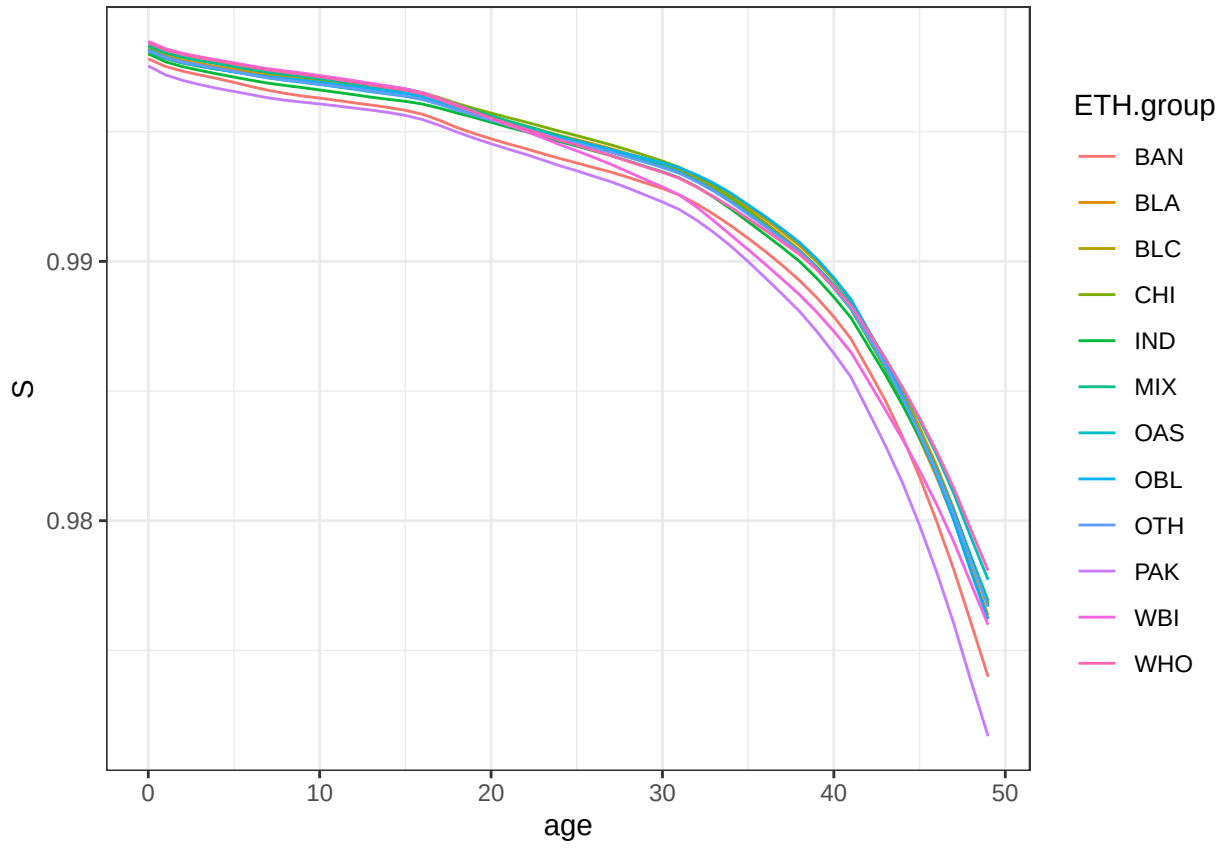

```
ggplot(ETHPOP_lifetable_2011M, aes(x = age, y = qx, colour = ETH.group)) +
  geom_line() +
  # scale_y_continuous(trans='log2') +
  coord_trans(y = "log10") +
  theme_bw()
```




```
ggplot(ETHPOP_lifetable_2011M, aes(x = age, y = qx, colour = ETH.group)) +
  geom_line() +
  # scale_y_continuous(trans='log2') +
  theme_bw()
```




```
ggplot(ETHPOP_lifetable_2011M, aes(x = age, y = S, colour = ETH.group)) +
  geom_line() +
  theme_bw()
```



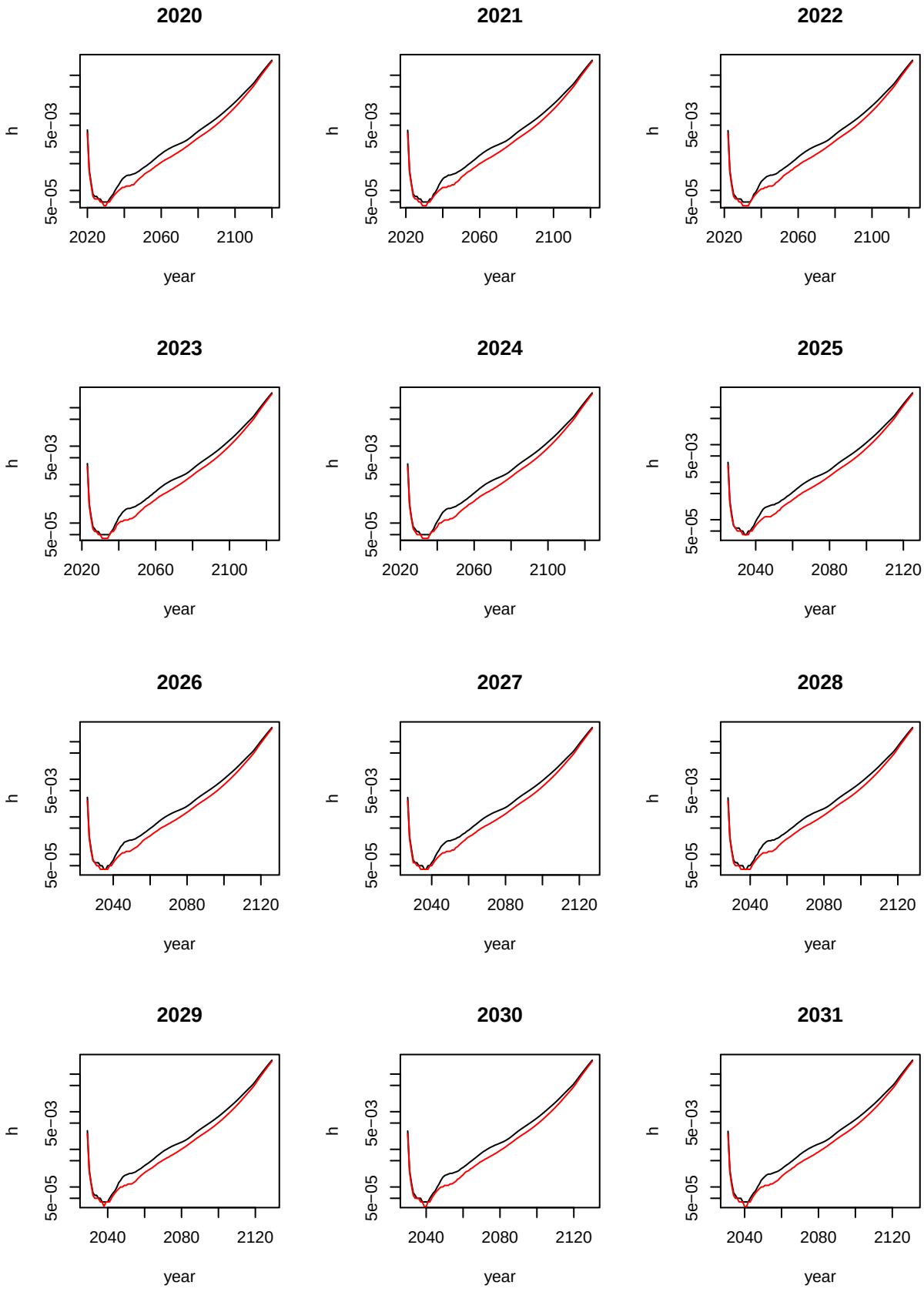
2 ONS lifetables

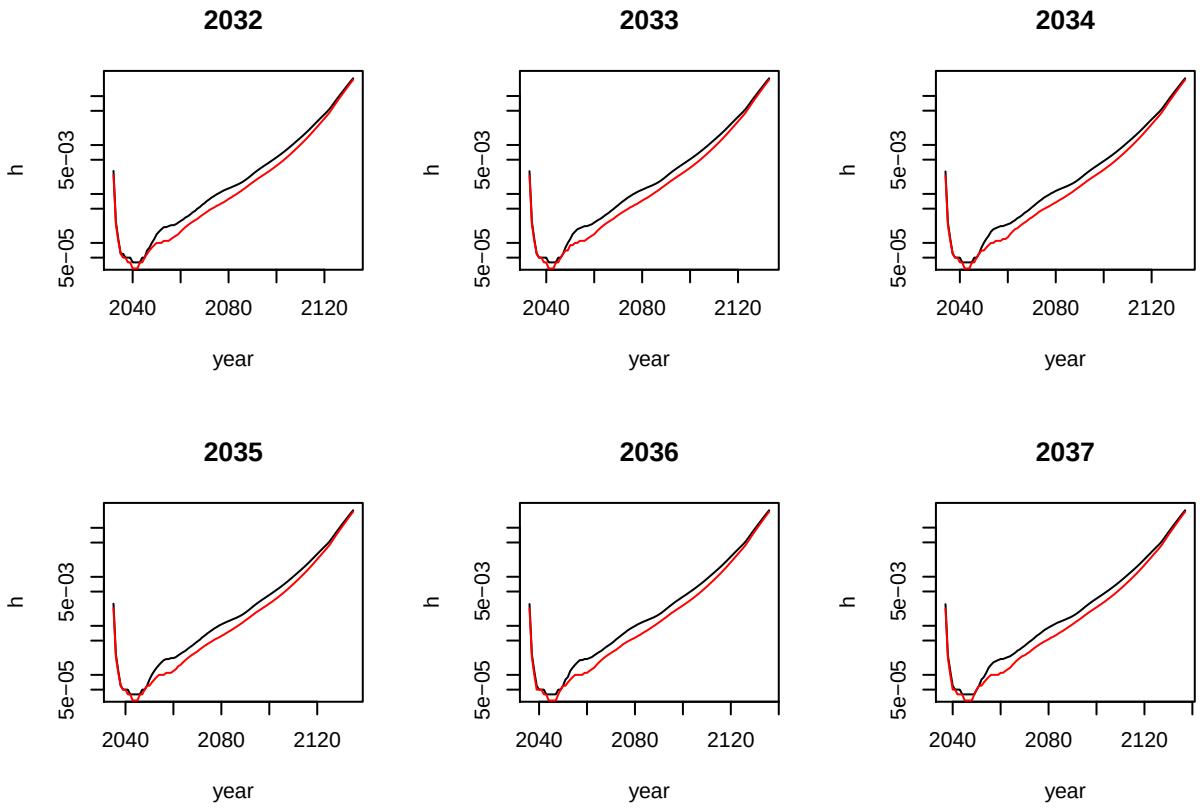
Read in and check.

```
lifetables <- read_lifetables()
# save(lifetables, file = here::here("data", "lifetables.RData"))


ONS_lifetables <-
  do.call(rbind, lifetables) %>%
  mutate(new_yr = year < dplyr::lag(year, default = Inf),
         id = cumsum(new_yr)) %>%
  group_by(id) %>%
  mutate(baseyr = min(year)) %>%
  ungroup() %>%
  select(-id, -new_yr) %>%
  mutate(baseyr = as.factor(baseyr))
```

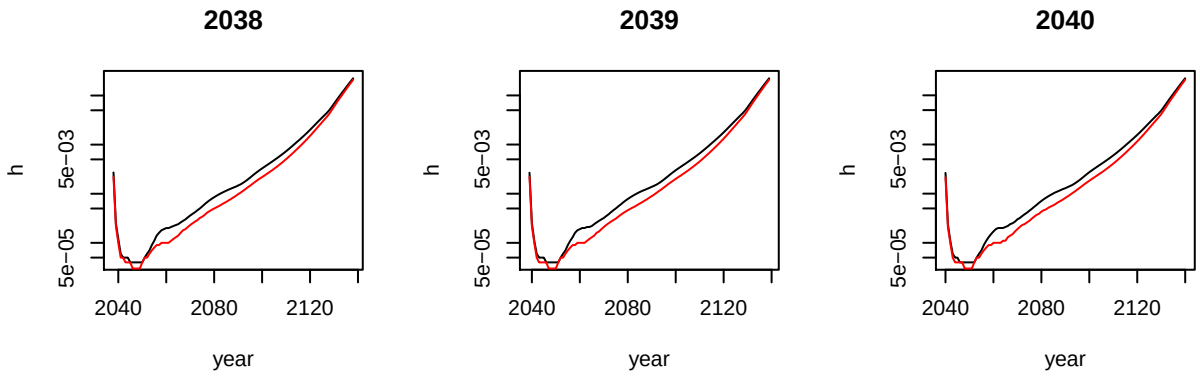
```
par(mfrow= c(2,3))
for (i in seq_along(lifetables)) {
  plot(x = lifetables[[i]]$year[lifetables[[i]]$sex == "M"],
       y = lifetables[[i]]$qx[lifetables[[i]]$sex == "M"], log = "y", type = "l",
       main = names(lifetables)[i], ylab = "h", xlab = "year")
```

```
  lines(x = lifetables[[i]]$year[lifetables[[i]]$sex == "F"],
        y = lifetables[[i]]$qx[lifetables[[i]]$sex == "F"], log = "y", type = "l", col = "red")
}
```

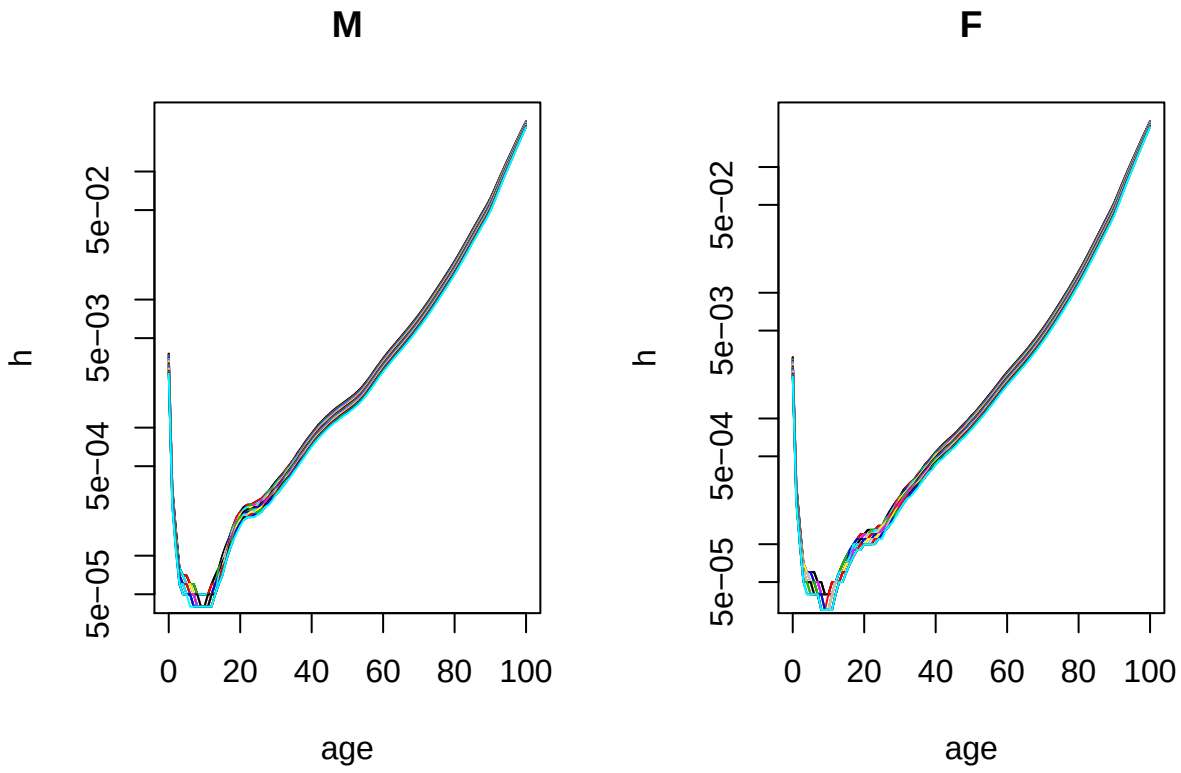

2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
h
5e−05
5e−03
2020
2040
2060
2080
2100
2120
year



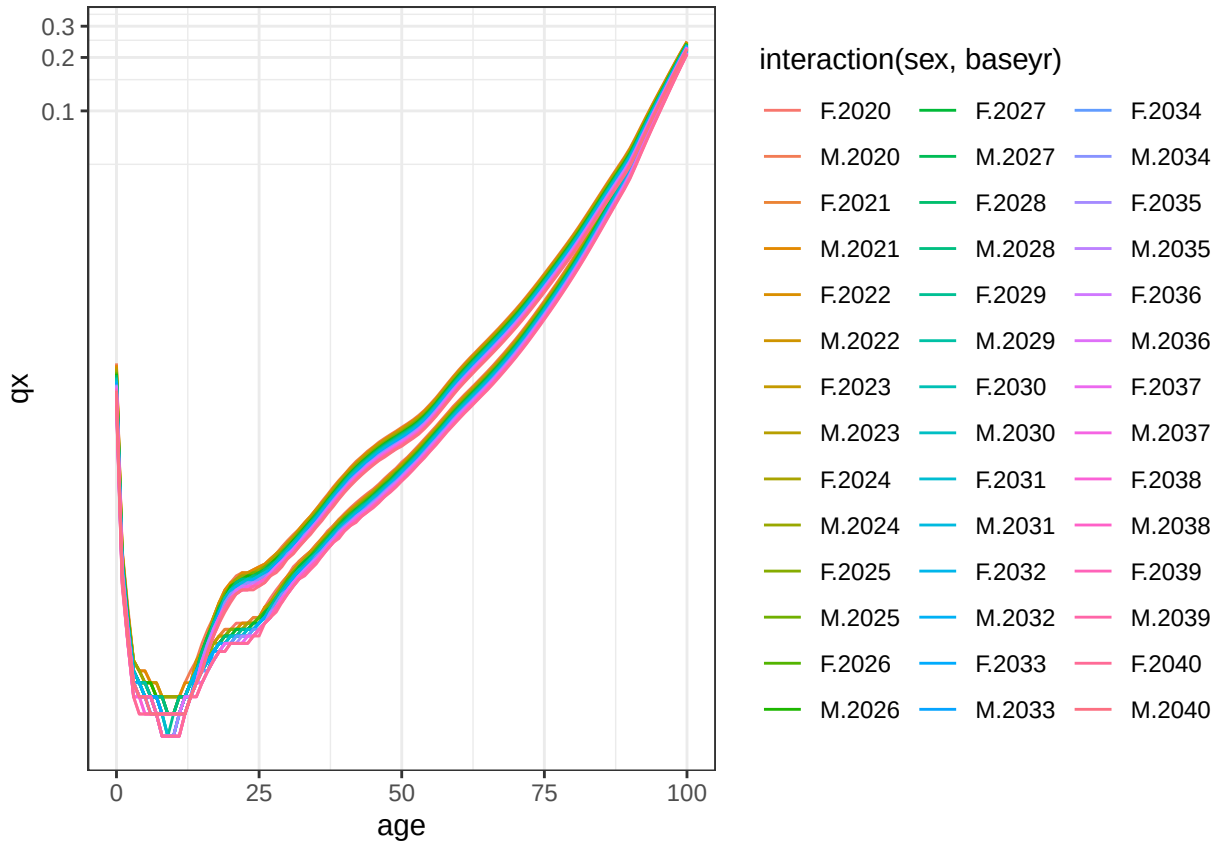

```
par(mfrow= c(1,2))
```




```
for (j in c("M","F")) {
  plot(x = lifetables[[1]]$age[lifetables[[1]]$sex == j],
       y = lifetables[[1]]$qx[lifetables[[1]]$sex == j], log = "y", type = "l",
       main = j, ylab = "h", xlab = "age")
  for (i in seq_along(lifetables)) {
    lines(x = lifetables[[i]]$age[lifetables[[i]]$sex == j],
          y = lifetables[[i]]$qx[lifetables[[i]]$sex == j], log = "y", type = "l", col = i)
  }
}
```



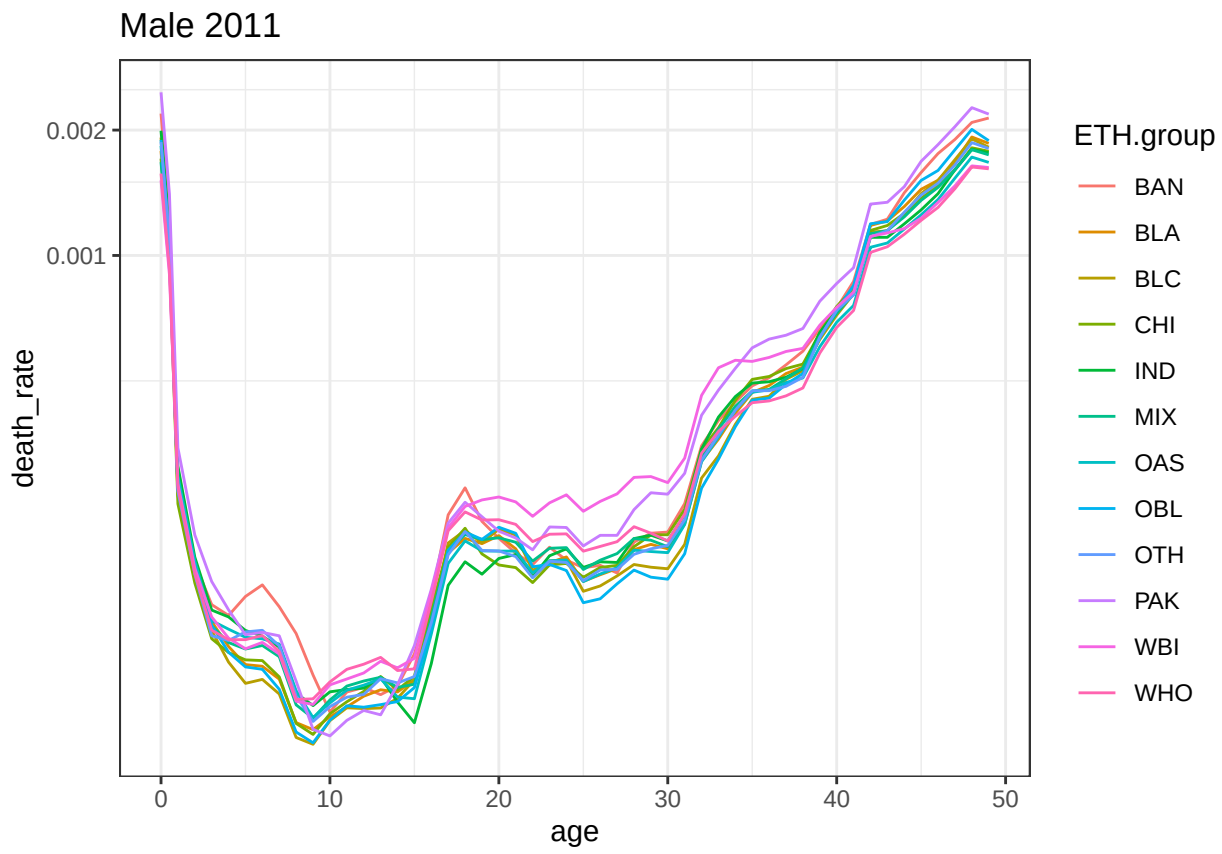

```
ggplot(ONS_lifetables, aes(x = age, y = qx, colour = interaction(sex, baseyr))) +
  geom_line() +
  # scale_y_continuous(trans='log2') +
  coord_trans(y = "log10") +
  theme_bw()
```



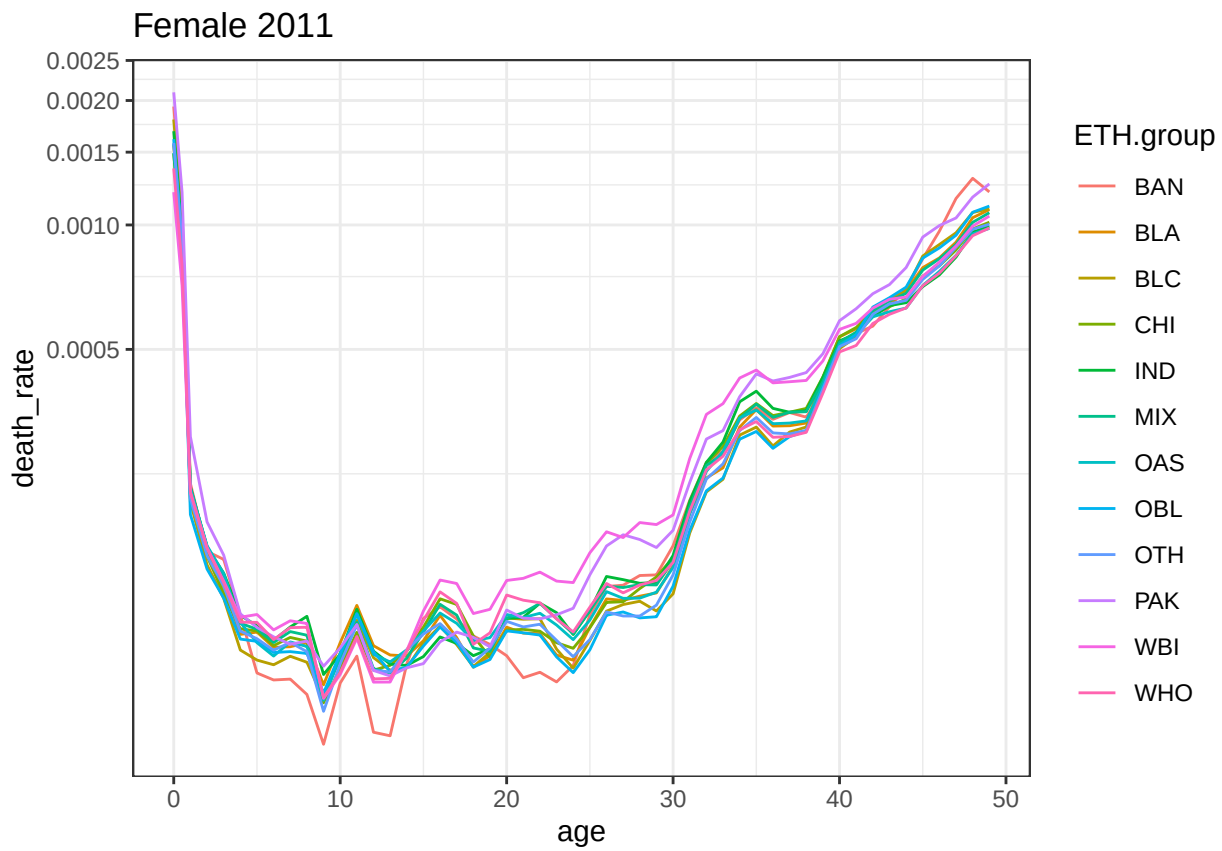
3 Comparison with ONS and ETHPOP

2011

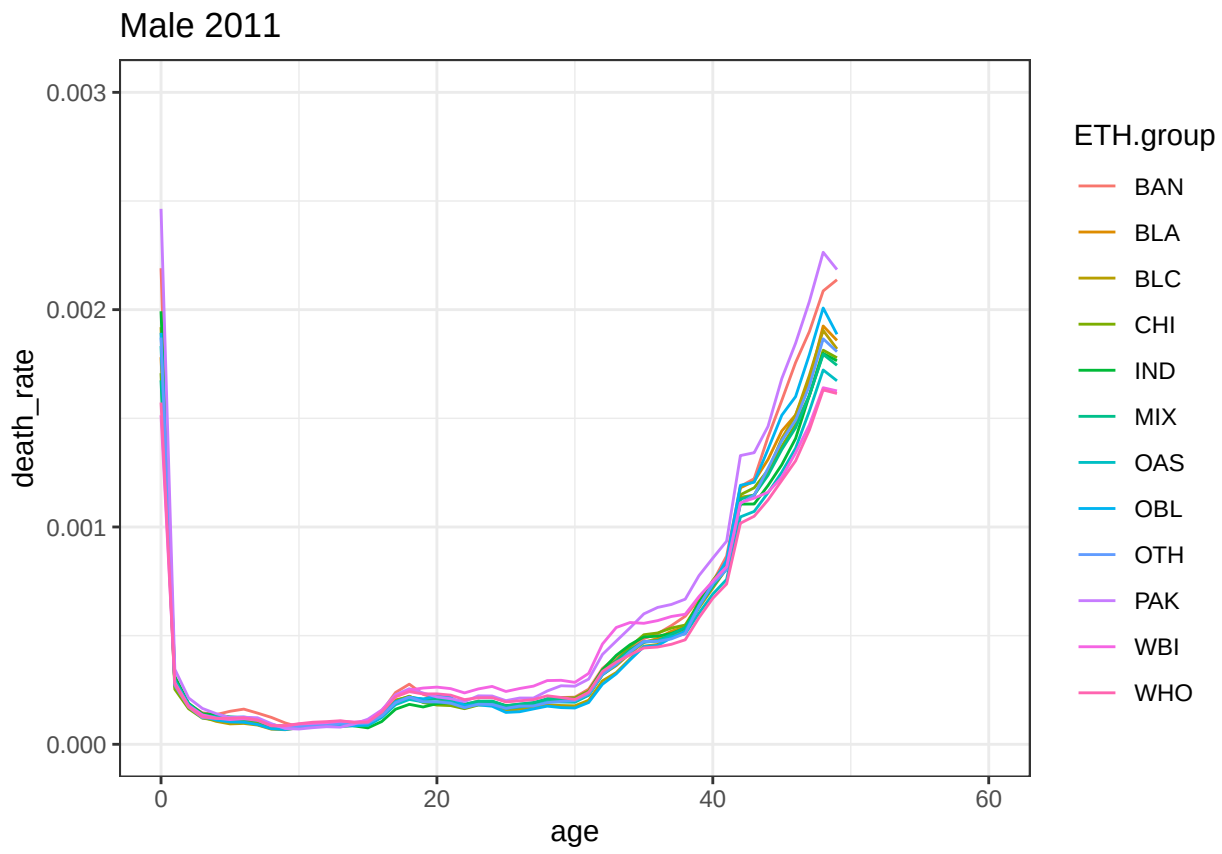
```
ggplot(ETHPOP_lifetable_2011M, aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  coord_trans(y = "log10") +
  ggtitle("Male 2011") +
  theme_bw() +
  geom_line(aes(age, qx, colour = "ONS"),
            data = ONS_lifetables[ONS_lifetables$baseyr == 2011 & ONS_lifetables$sex == "M", ],
            colour = "black")
```



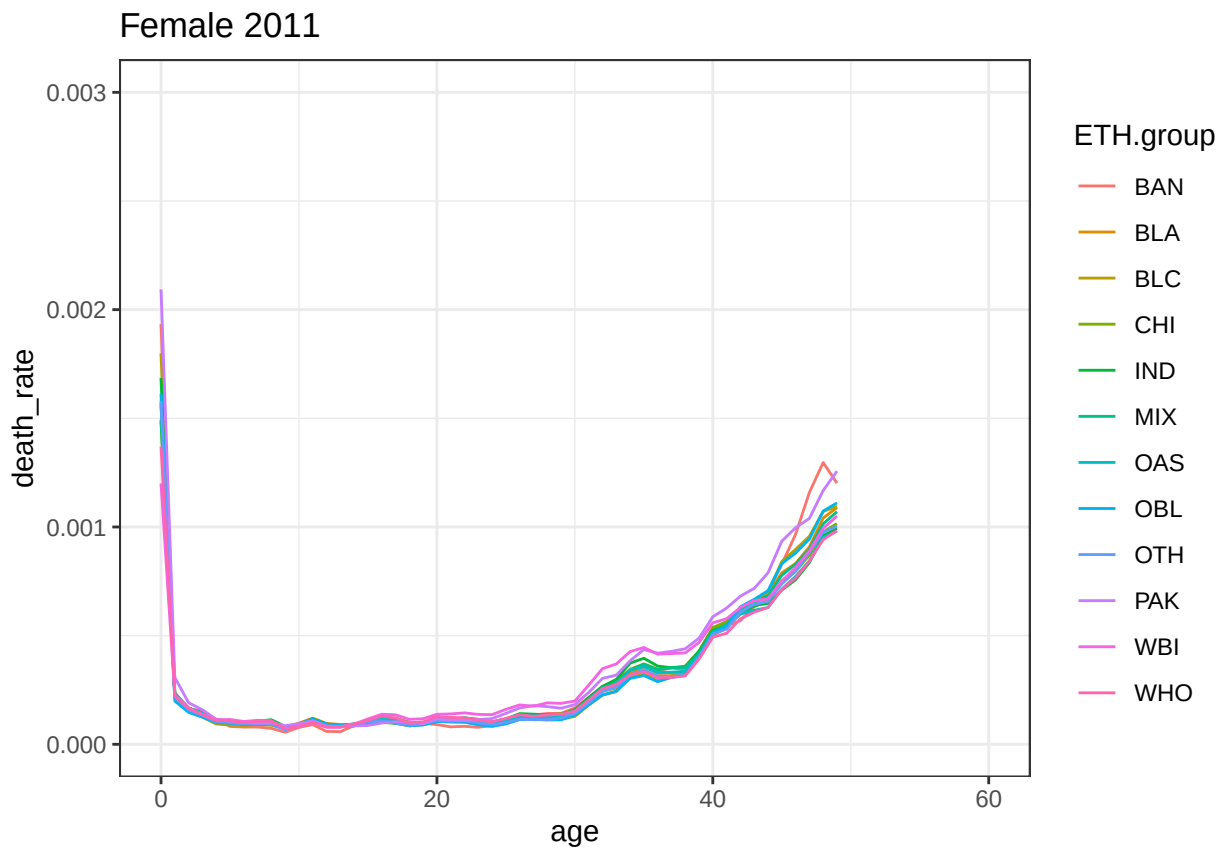

```
ggplot(ETHPOP_lifetable_2011F, aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  ggtitle("Female 2011") +
  coord_trans(y = "log10") +
  theme_bw() +
  geom_line(aes(age, qx, colour = "ONS"),
            data = ONS_lifetables[ONS_lifetables$baseyr == 2011 & ONS_lifetables$sex == "F", ],
            colour = "black")
```




```
ggplot(ETHPOP_lifetable_2011M, aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  ggtitle("Male 2011") +
  theme_bw() +
  geom_line(aes(age, qx, colour = "ONS"),
            data = ONS_lifetables[ONS_lifetables$baseyr == 2011 & ONS_lifetables$sex == "M", ],
            colour = "black") +
  ylim(0, 0.003) + xlim(0, 60)
```

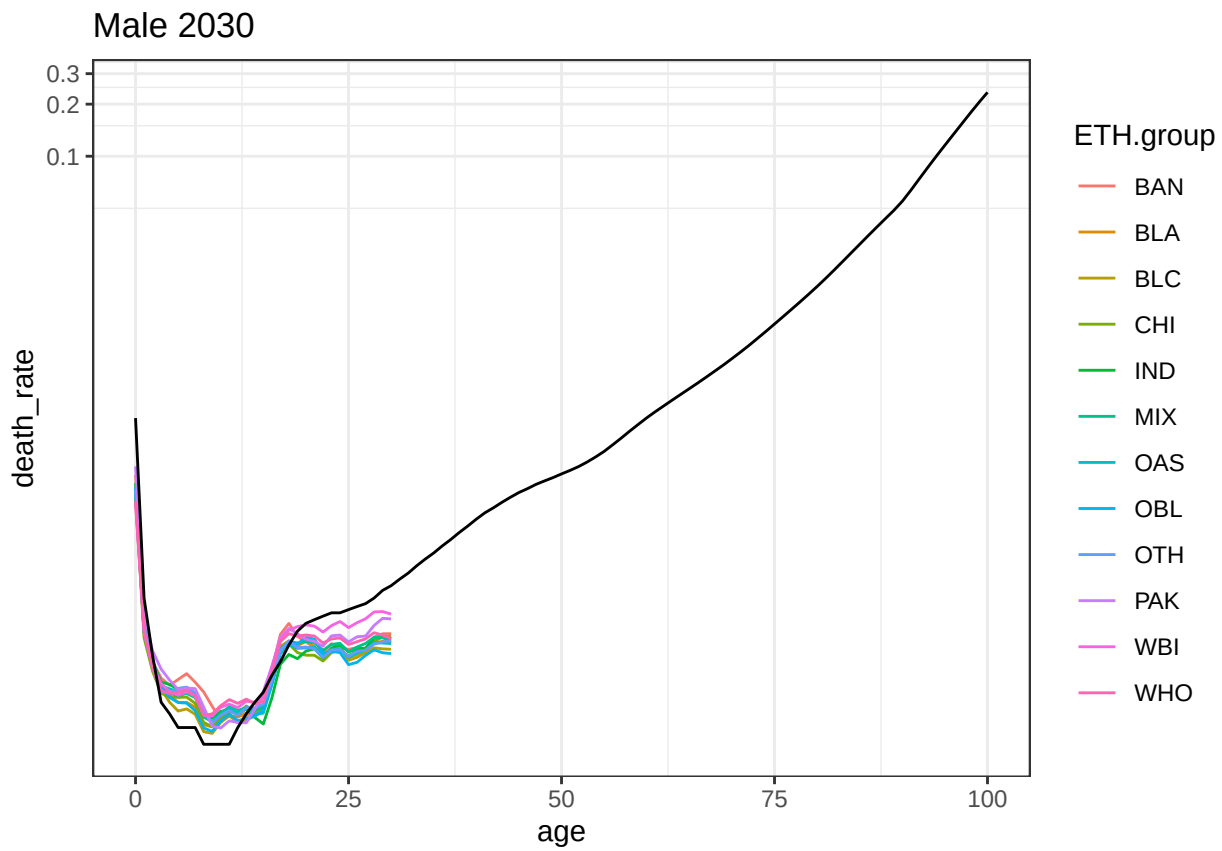



```
ggplot(ETHPOP_lifetable_2011F, aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  ggtitle("Female 2011") +
  theme_bw() +
  geom_line(aes(age, qx, colour = "ONS"),
            data = ONS_lifetables[ONS_lifetables$baseyr == 2011 & ONS_lifetables$sex == "F", ],
            colour = "black") +
  ylim(0, 0.003) + xlim(0, 60)
```

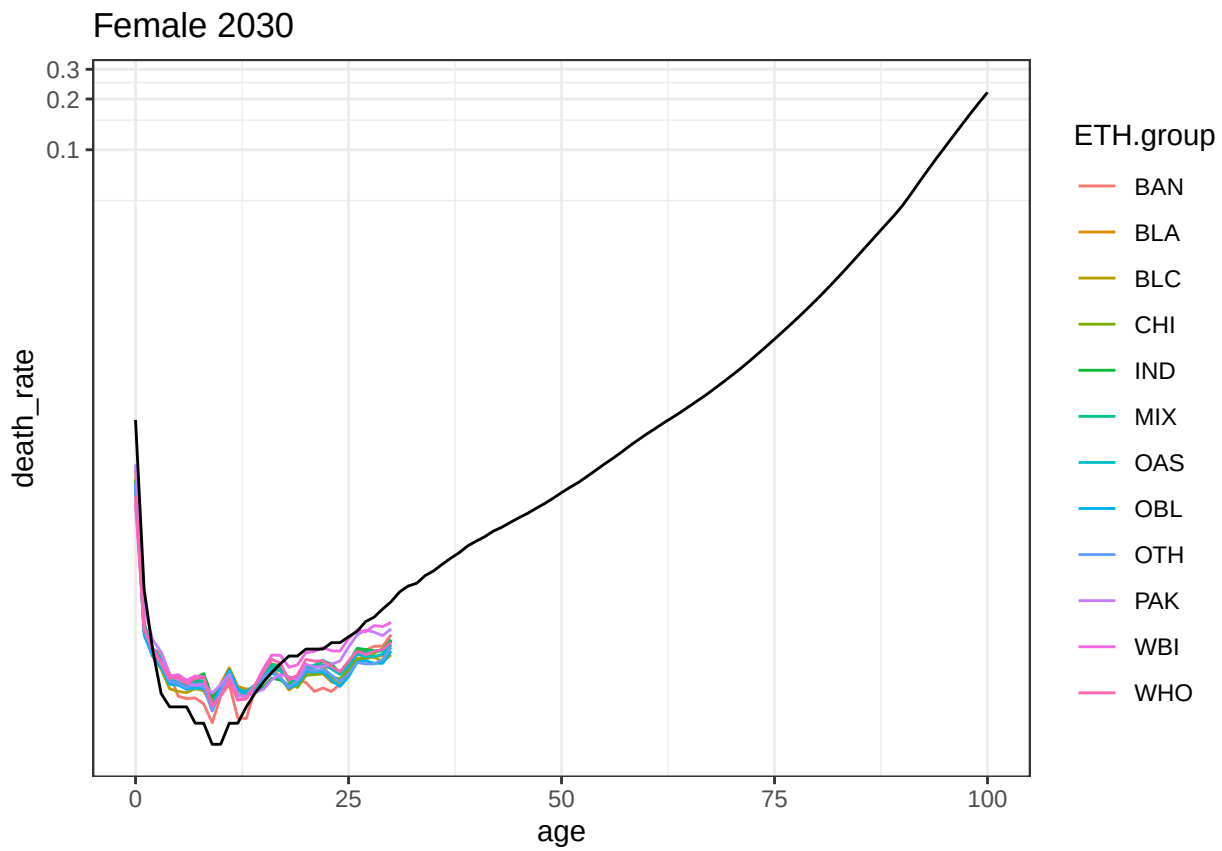


2030

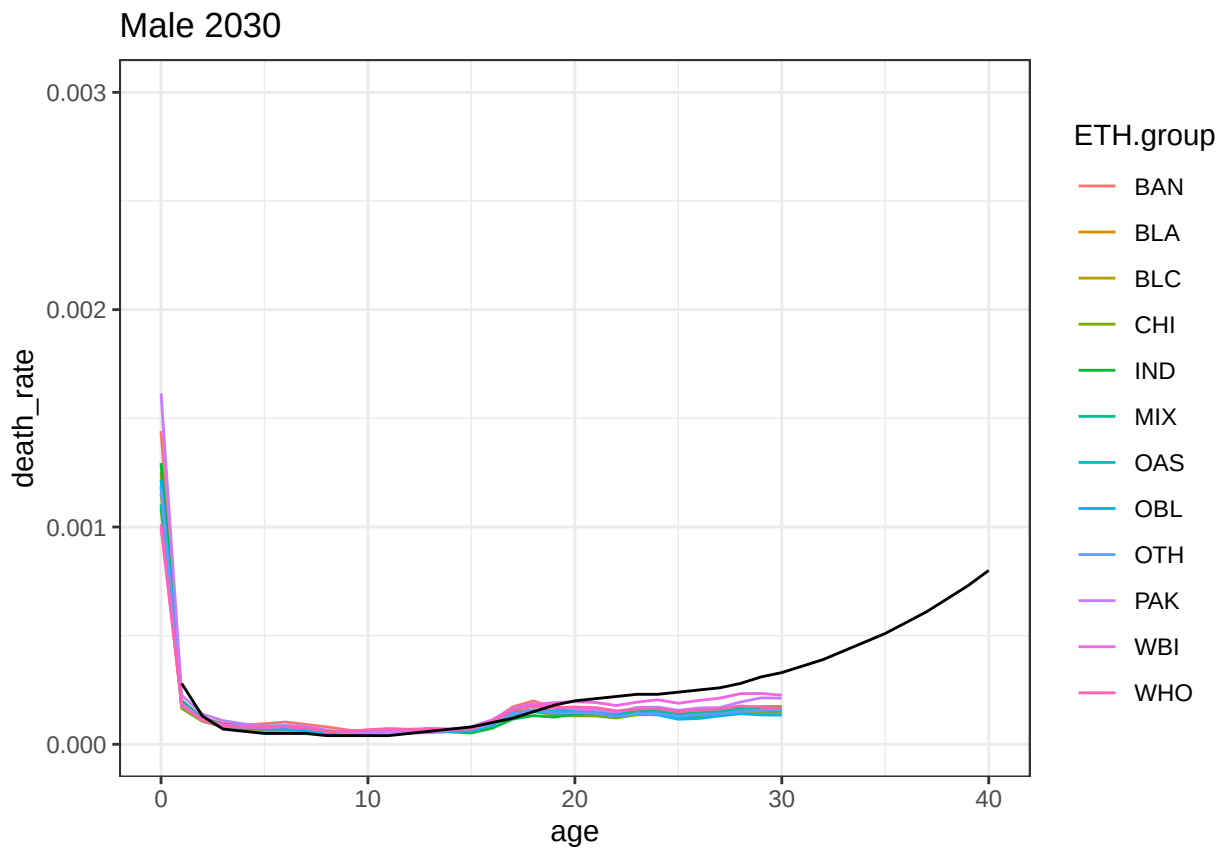
```
baseyr_2030 <- ETHPOP_lifetable$id[ETHPOP_lifetable$yr_age == "2030_0"][1]

ETHPOP_lifetable[ETHPOP_lifetable$id == baseyr_2030 & ETHPOP_lifetable$sex == "M", ] %>%
  ggplot(aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  coord_trans(y = "log10") +
  ggtitle("Male 2030") +
  theme_bw() +
  geom_line(aes(age, qx, colour = "ONS"),
            data = ONS_lifetables[ONS_lifetables$baseyr == 2030 & ONS_lifetables$sex == "M", ],
            colour = "black")
```



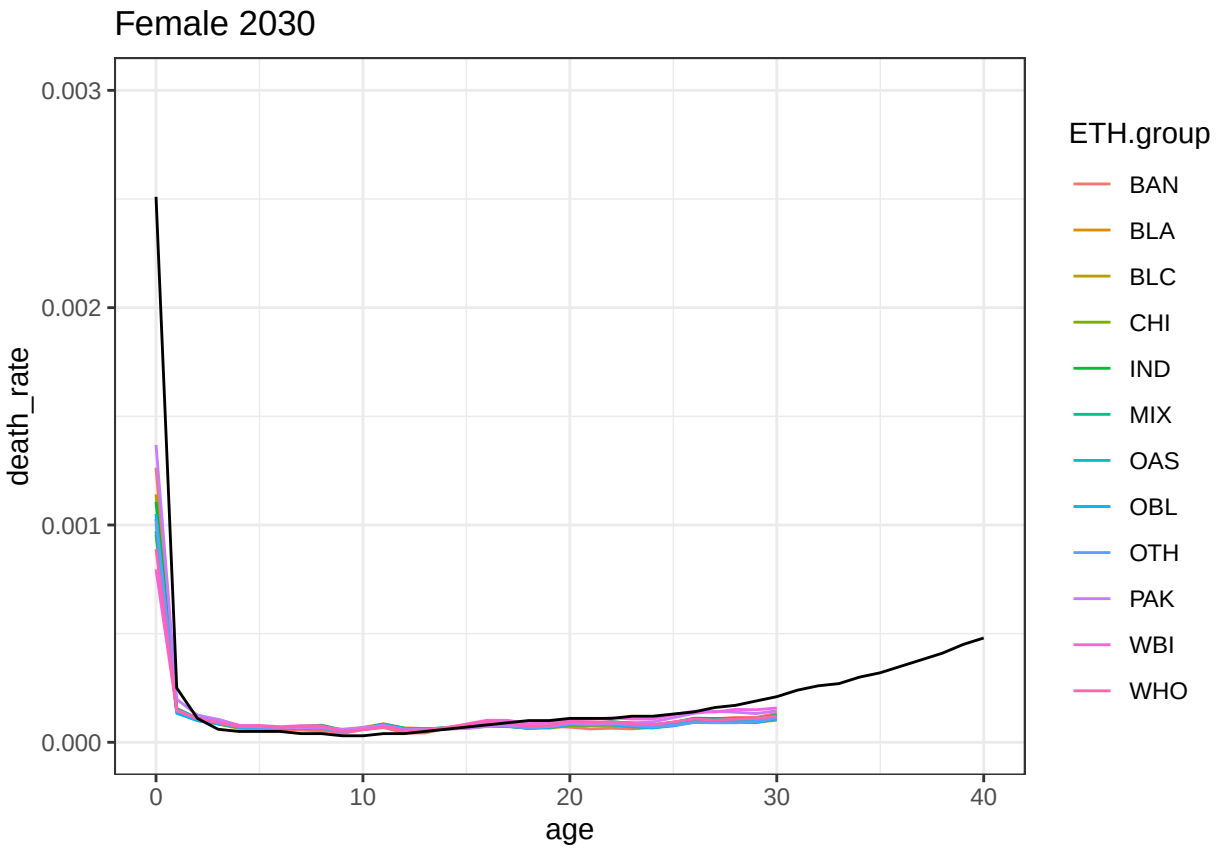

```
ETHPOP_lifetable[ETHPOP_lifetable$id == baseyr_2030 & ETHPOP_lifetable$sex == "F", ] %>%
  ggplot(aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  ggtitle("Female 2030") +
  coord_trans(y = "log10") +
  theme_bw() +
  geom_line(aes(age, qx, colour = "ONS"),
            data = ONS_lifetables[ONS_lifetables$baseyr == 2030 & ONS_lifetables$sex == "F", ],
            colour = "black")
```




```
baseyr_2030 <- ETHPOP_lifetable$id[ETHPOP_lifetable$yr_age == "2030_0"][1]

ETHPOP_lifetable[ETHPOP_lifetable$id == baseyr_2030 & ETHPOP_lifetable$sex == "M", ] %>%
  ggplot(aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  ggtitle("Male 2030") +
  theme_bw() +
  geom_line(aes(age, qx, colour = "ONS"),
            data = ONS_lifetables[ONS_lifetables$baseyr == 2030 & ONS_lifetables$sex == "M", ],
            colour = "black") +
  ylim(0, 0.003) + xlim(0, 40)
```



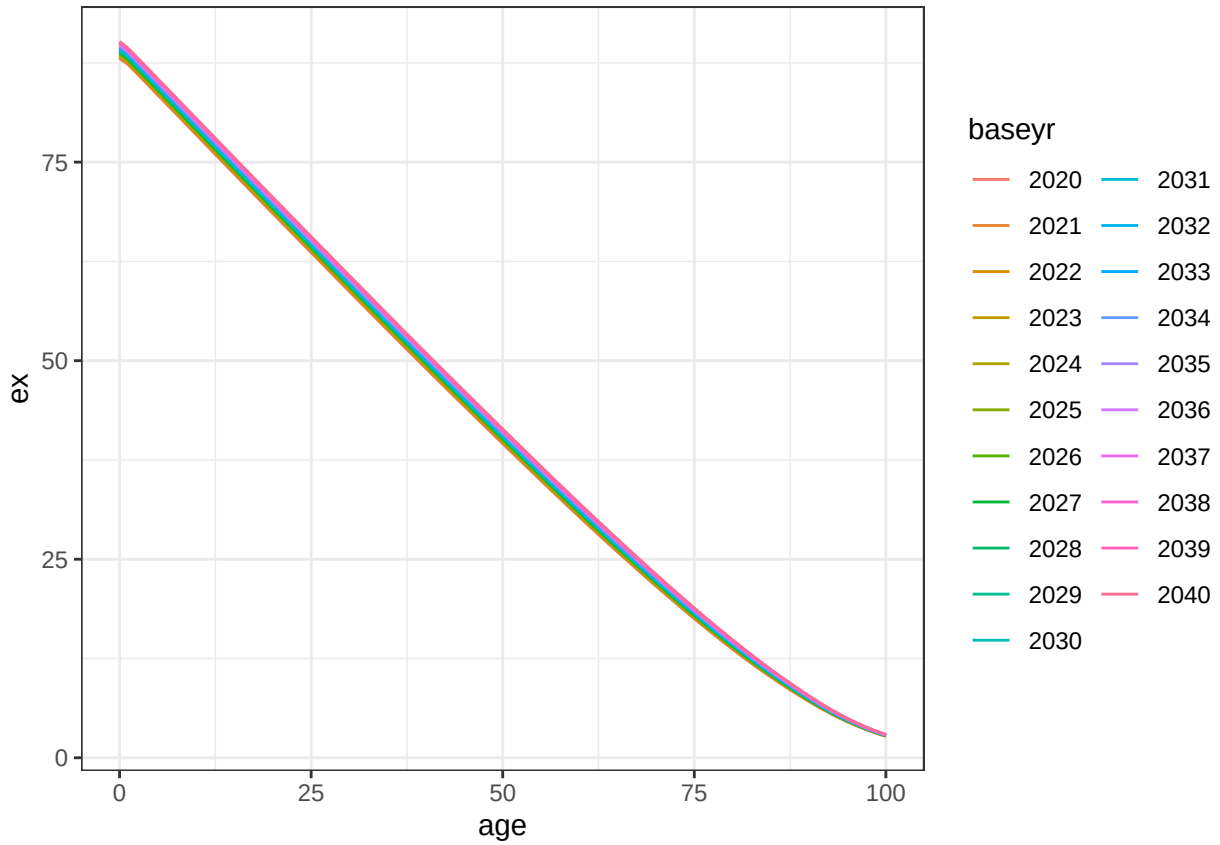

```
ETHPOP_lifetable[ETHPOP_lifetable$id == baseyr_2030 & ETHPOP_lifetable$sex == "F", ] %>%
  ggplot(aes(x = age, y = death_rate, colour = ETH.group)) +
  geom_line() +
  ggtitle("Female 2030") +
  theme_bw() +
  geom_line(aes(age, qx, colour = "ONS"),
            data = ONS_lifetables[ONS_lifetables$baseyr == 2030 & ONS_lifetables$sex == "F", ],
            colour = "black") +
  ylim(0, 0.003) + xlim(0, 40)
```



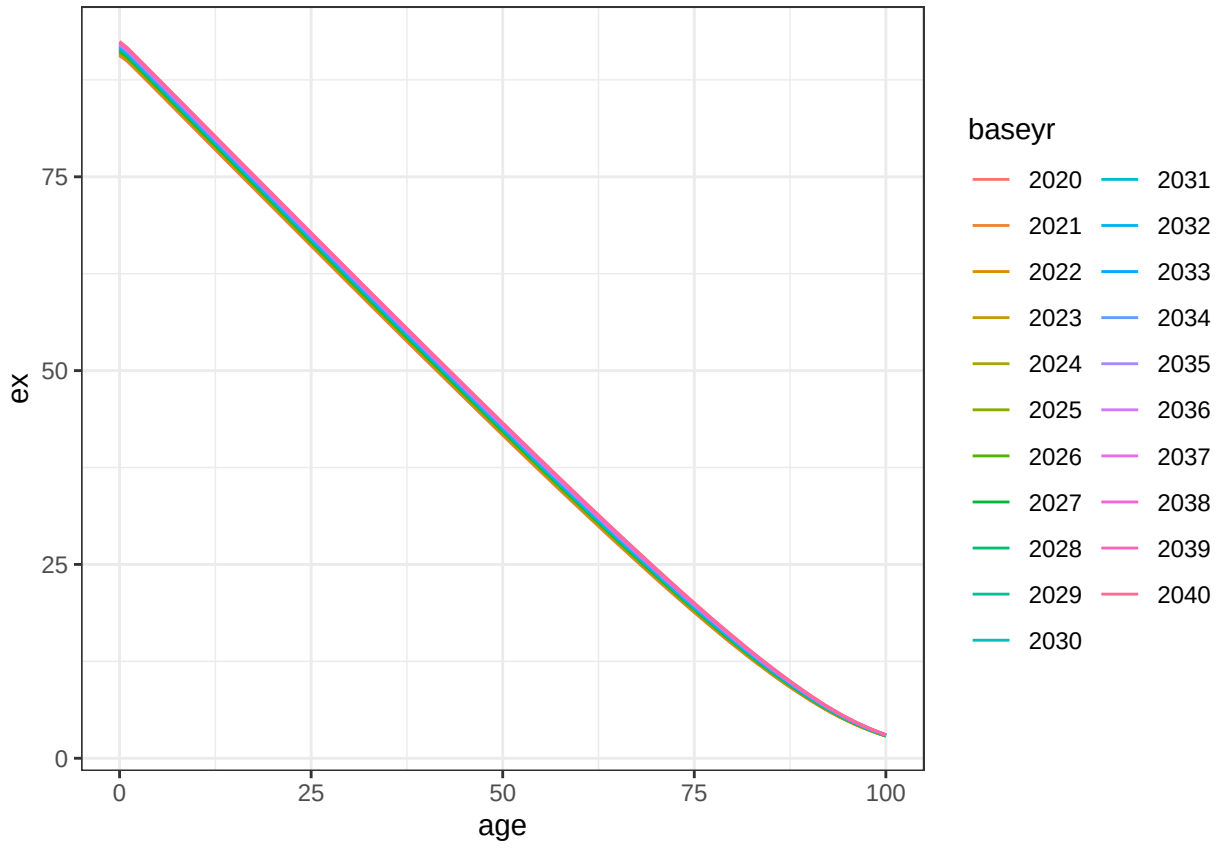
3.1 Life expectancy

ONS

```
ggplot(ONS_lifetables[ONS_lifetables$sex == "M", ], aes(age, ex, colour = baseyr)) +
  geom_line() +
  theme_bw()
```




```
ggplot(ONS_lifetables[ONS_lifetables$sex == "F", ], aes(age, ex, colour = baseyr)) +
  geom_line() +
  theme_bw()
```


ex
75
50
25
0
0
25
50
75
100
age
baseyr
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040