Chapter 5
5. The effect of risk factors on individual disability trajectories

Abstract
We studied compression or expansion of disability by estimating the average duration of life with disability of a cohort exposed to certain risk factors. However, such averages are composed of various individual disability trajectories. This chapter studies the role of risk factors on individual disability trajectories that underlie compression or expansion of disability.

We use microsimulation to show the diversity of individual trajectories underlying the disability experience at the population level. Multistate cohort models describe the dynamics of a synthetic cohort, estimated by empirically observed transition rates. Microsimulation generates a virtual population of individuals with disability trajectories starting from the same transition rates and which therefore are consistent with the cohort results from the cohort models. We verify the microsimulated lifepaths against the empirical data from the U.S. Health and Retirement Study (HRS).

The results show that individual disability trajectories vary widely between men and women and for different risk factor groups. Women and obese individuals have the highest risk of ever experiencing disability. We conclude that longitudinal microsimulation is a powerful instrument to study the distribution of individual disability trajectories and to assess the effect of risk factors on the process of compression and expansion of disability at higher ages.

Introduction
The impressive gains in life expectancy in recent decades have raised the question whether the decline in mortality causes an expansion or a compression of the duration of disability. Most studies of disability estimate the expected number of years for a synthetic cohort to live with a certain condition. However, the resulting life expectancies represent averages of the experience of a cohort. There is a large variability in disability trajectories in the last year of life (Gill et al. 2010). Very few individuals, if any, will actually experience the exact duration of health or disability estimated by the cohort approach. Microsimulation offers a
tool to show the variety of trajectories and the distribution of lifespans with or without disability (Cai et al. 2010). A particularly interesting application is to demonstrate the effect of risk factors on disability trajectories at higher ages. Multistate life table analyses have shown that smoking compresses life expectancy, both in good health and with disability, while obesity expands years with disability. The individual disability trajectories from the microsimulation are consistent with multistate cohort models, but they add the distribution of individual life paths to the mean effects of risk factor exposure.

According to Wolf (1986; Wolf 2001), microsimulation is essentially an exercise in sampling. It draws a sample for which the characteristics and dynamics are determined by a probability model. If the models are realistic, the virtual population closely resembles the real population. Transition rates are estimated based on observed, often censored or truncated survey data. A multistate life table summarizes the age-specific rates into a cohort biography, giving the average experience of the sampled individuals. Microsimulation starts from the same estimated transition rates, but simulates synthetic individual trajectories using Monte Carlo technique. The resulting biographies at the individual level can be aggregated in numerous ways, giving the distributions rather than only the mean (Imhoff and Post 1998). Microsimulation models can treat time as a discrete variable or use continuous time. Newly developed models and software like the package MSM in R (Jackson 2008), MicMac (Gampe et al. 2009) and SPACE (Cai et al. 2010) make use of microsimulation.

This paper presents an individual biography approach to study the distribution of trajectories of health and disability before dying, using the U.S. Health and Retirement Study. We will demonstrate the consistency between the cohort multistate approach and the individual microsimulation approach. Additional information from microsimulation entails the frequency distribution of disability trajectories and the distribution of age at death and duration of disability. The application demonstrates how microsimulation can offer another perspective on analyzing the effect of risk factors on compression or expansion of disability. It explores the distribution of individual disability trajectories, generated by gender, obesity, and smoking.
Data

Health and Retirement Study (HRS)

We used data from the Health and Retirement Study (HRS) data file containing the HRS and the Asset and Health Dynamics Among the Oldest Old (AHEAD) which began in 1992 and 1993, respectively, and were merged in 1998. More information is available at http://hrsonline.isr.umich.edu/. The HRS and AHEAD surveys include a nationally representative sample of initially non-institutionalized persons born between 1931-1941 (HRS, aged 51-61 in 1992) and in 1923 or earlier (AHEAD, aged 70 and older in 1993). Sampled persons were re-interviewed biannually. We used data from 7 waves from 1992 to 2004. Response was on average 86% (HRS) and 90% (AHEAD). We selected white non-Hispanic men and women. Data on vital status and month and year of death are obtained through the mortality register (the National Death Index) and exit interviews.

Disability and risk factors

We defined disability by the Katz basic activities of daily living (ADL): Walking, Bathing, Dressing, Toileting and Feeding (Katz et al. 1963). We classify as ADL disabled anyone answering ‘with difficulty’ to at least one of the ADL items. The definitions are in line with other literature using disability measures from the HRS (Reynolds et al. 2005). We will from now on refer to ADL or severe disability as disability, indicated in short by S, to avoid confusion with death (D). To illustrate the effect of risk factors on individual trajectories of disability, we distinguish three risk factor groups: 1) smokers, defined as those who smoke at entrance into the survey 2) obese, defined as those with a BMI equal to or over 30 at first observation in the survey 3) non-smoking normal weight, defined as those with normal weight (BMI between 18.5 and 25) and non-smoking at entry into observation. Self-reported weight and height at baseline are used to calculate BMI (kg/m²).

Methods

In order to calculate life expectancies in particular health states for an entire cohort, we construct multistate life tables, based on transition intensities to health and disability states and to death. Based on the same rates, individual health trajectories can be simulated. Health states are indicated as either 1) healthy, 2) disabled, or 3) dead. The overall outcomes are
consistent, but the longitudinal microsimulation gives information at a more detailed, individual level. Both methods rely heavily on the accuracy and fit of the transition rates.

Figure 1: Markov model with 3 states.

\[
\begin{align*}
1 & \text{- Healthy (H)} \\
2 & \text{- Disabled (S)} \\
3 & \text{- Death (D)} \\
\end{align*}
\]

\[\mu_{12} \quad \mu_{21} \quad \mu_{13} \quad \mu_{23}\]

**Estimation of transition rates**

We estimated the hazard rates of transitions to death and disability by age for males and females, including the possibility of recovering from disability, based on a Markov model as shown in Figure 1. Exact months of deaths are available and for transitions to disability and recovery we assume transitions halfway between two waves. Age-specific transition rates are estimated by exponential smoothing of occurrences and exposures using Poisson regression to reduce the variability of the rates (Mamun 2003). The assumption is that the hazards of death and disability increase exponentially with age, which fitted the empirical transition rates very well. This implies a Gompertz hazard, but we assume the rates to be constant within the 1-year age intervals for calculating probabilities. Hence, our model is piece-wise constant with rates increasing exponentially with age.

When estimating the rates, making life tables and simulating individual biographies, we stopped at age 105, using the remaining life expectancy at age 105 of the average American male and female population regardless of their state of health. This assumption was made because there is no empirical data in our HRS sample above that age. Besides, exponential
increase of death rates at very high ages is still disputed (Horiuchi and Wilmoth 1998). As the number of people surviving to ages over 105 is very small, this causes negligible changes in life expectancy outcomes.

Because of very low rates of relapse (becoming disabled again after recovery), these rates are not considered here. This implies that simulated individuals who have recovered from disability experience the same hazards as those who stayed healthy all along.

Multistate life table analysis

We built multistate cohort life tables for males and females separately based on the estimated age-specific transition rates. The synthetic cohort starts with 100,000 individuals in the healthy state. Because of the possibility of recovering from disability, we need matrix algebra to solve the set of linear equations (Hougaard 1999). The main outcomes of the multistate life tables are mean durations at age \( x \geq 55 \); total life expectancy, life expectancy with disability and healthy life expectancy. Confidence intervals for the multistate life table outcomes were calculated using bootstrapping with 250 replicates.

Longitudinal continuous-time microsimulation

Longitudinal microsimulation creates individual biographies that are synthetic and fully determined by (1) empirical transition rates and (2) a random mechanism, known as the Monte Carlo technique, which identifies which individuals experience a transition during a given interval (Imhoff and Post 1998). Given that the longitudinal microsimulation is based on the same set of transition rates as in the multistate analysis, the resulting average sojourn times per state are very close to results from the multistate life table. The small difference stems from Monte Carlo variation which becomes smaller as the number of simulated biographies is increased.

We use microsimulation in continuous time, allowing for multiple events to happen within one interval, here 1 year. Because the HRS gives information on health status only at interview times (about every two years), the outcome of the microsimulation will give more transitions than are observed in the empirical data. The waiting time distribution used here is the piece-wise constant exponential distribution, with intervals of 1 year.
The results of the microsimulations can be verified by comparing them to the empirical data. In order to test this we simulated the same number of individuals as in the empirical data (7,191 males) and for each individual simulated the segments of life observed in the data. As mentioned earlier, the continuous-time microsimulation can generate several events in the time-interval between two interviews. Hence, the number of transitions in the microsimulation is larger than in the empirical sample. However, when comparing prevalences at two specific ages, the transition probabilities between these ages are the same in the empirical sample data and the microsimulation. Figure 2 shows the survivorship functions in healthy and disabled state, comparing the simulated with the empirical data. Note that every time a microsimulation is run, the results will change slightly. This test assures that the estimation of the rates and the simulation of lifepaths are well replicating the actual empirical data.

Figure 2: Comparison of empirical and simulated survivors by disabled and healthy state.
We use the R-package MSM and the function sim.msm to create the synthetic biographies (Jackson 2008). Sim.msm allows for simulation individual trajectories from a continuous-time Markov model based on the transition intensity matrix estimated from the empirical data. We simulate 50,000 males and females starting at age 55 in a healthy state, running through the single age and sex-specific transition rates up to age 105.

Uncertainty

Two different types of uncertainty can be distinguished in our analyses. One is sampling variability and the other is Monte Carlo variation. The sampling variability is captured in the confidence intervals obtained by bootstrapping. At every resample, the bootstrap method draws from the empirical data, estimates the rates and recalculates the life expectancies. Hence the confidence intervals from bootstrapping include the uncertainty around the parameter estimates. This is different for the microsimulation. The distribution that results from the microsimulation only represents individual randomness. Bear in mind that the simulation of lifespans starts from the point estimates of the transition rates without error margins. The standard deviations of the simulated lifespans solely represents individual variability. Every microsimulation run is slightly different because of Monte Carlo variation, but when the number of simulation heads towards infinity, this variation moves towards zero.

Results

Descriptives

The selection of non-Hispanic white individuals aged 55 and over who participated at least 3 years and reported BMI, smoking and ADL resulted in a sample of 16,176 individuals. The selection is shown in Table 1. Average follow-up is 5.8 years. Males contributed 41,920 personyears, 1,090 events to disability and 1,603 deaths. The females in the sample accounted for 52,257 personyears, 1,516 transitions to disability and also 1,603 deaths. A total of 4.3% of men and 5.1% of women recovered from disability while under observation.

Multistate life table analysis: the cohort biography

The actual life expectancy of the total unselected white American population in 2003 was 24.6 years for men and 28.1 for women at age 55 (Centers for Disease Control and
Prevention and National Center for Health Statistics 2006). The comparable life expectancy of our study population was respectively 25.5 [25.0:26.0] and 30.5 [30.0:31.0] years (excluding underweight individuals). The results of the multistate analysis shows that life years with disability at age 55 are significantly higher for females (8.1 [7.7:8.5]) than for males (5.3 [5.0:5.7]). Healthy and disabled life spans by risk factor groups are depicted in Figure 3. As expected the males and females who were non-smoking and had normal weight at baseline have the longest healthy life expectancy. There is no significant difference between the healthy life span for obese and smoking females. Obese individuals live significantly longer with disability than smokers or non-smoking normal weight males and females. These results are consistent with earlier findings in the literature (Al Snih et al. 2007; Flegal et al. 2005; Reuser et al. 2009).

Figure 3: Life expectancy at age 55 by health state (HE) and disabled state (SE) for smoking, obese and healthy (non smoking normal weight) males and females, 95% confidence intervals obtained by bootstrapping.
Continuous-time microsimulation: individual biographies

Based on the sex and age-specific transition rates we simulated 50,000 synthetic biographies for males and 50,000 for females from age 55 to 105. Adding up the sojourn times per health state gives approximately the same results as the life expectancies that resulted from the multistate cohort analysis. Total average life expectancy of the simulated males and females resulted in 25.45 and 30.60 years compared to 25.54 and 30.51 when using multistate life table analysis. Years with disability were estimated at respectively 5.25 and 8.00 compared to 5.33 and 8.10 from the cohort approach. Note that every time 50,000 new individuals are simulated, the results from the microsimulation change slightly.

Table 1: Selection of the sample.

<table>
<thead>
<tr>
<th></th>
<th>men</th>
<th>women</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial sample</td>
<td>13086</td>
<td>17110</td>
<td>30196</td>
</tr>
<tr>
<td>Non-whites</td>
<td>2405</td>
<td>3535</td>
<td>5940</td>
</tr>
<tr>
<td>Hispanics</td>
<td>803</td>
<td>1099</td>
<td>1902</td>
</tr>
<tr>
<td>BMI &lt; 18.5</td>
<td>93</td>
<td>461</td>
<td>554</td>
</tr>
<tr>
<td>Aged &lt; 55</td>
<td>1067</td>
<td>1665</td>
<td>2732</td>
</tr>
<tr>
<td>Participated less than 3 year after first report of BMI</td>
<td>1467</td>
<td>1328</td>
<td>2795</td>
</tr>
<tr>
<td>Missing data on BMI, smoking or ADL disability</td>
<td>56</td>
<td>41</td>
<td>97</td>
</tr>
<tr>
<td>Final sample</td>
<td>7195</td>
<td>8981</td>
<td>16176</td>
</tr>
</tbody>
</table>

Additional information resulting from microsimulation is the distribution of disability trajectories. The most frequently followed health trajectory is trajectory 123, as shown in Table 2: 51% of males and 65% of females experience disability before dying. To die in a healthy state without ever experiencing disability is rarer for women (31%) than for men (45%). Table 2 also demonstrates how the trajectories followed are very different for the risk factor groups. Current smokers are more likely to die without ever having experience disability. Trajectories with recovery are rare. Obese individuals are most likely to go through disability.

The simulated lifepaths show that males and females who die between ages 55 and 70 are much more likely to die healthy (78% and 71% respectively). They live on average 7.9 and 7.8 years without disability, which is respectively 88% and 85% of their lives after age 55. Men and women who reach age 90 have little chance of escaping from disability (26% and
19% for men and women respectively) and can expect to live 31.0 and 30.6 years without disability after age 55, accounting for 78% and 74% of their lives after 55.

The collection of simulated lifepaths allows us to study the distribution of lifespans. For illustration, Figure 4 shows the distributions of total lifespan for smoking males and obese females. Distributions can be characterized by their mean, standard deviations, percentiles or quartiles of interest. Table 3 demonstrates the 25% quartile, the mean, the 75% quartile and the standard deviation of age at death and life years with disability for simulated males and females. Consistent with the frequencies shown in Table 2, the share of individuals remaining healthy throughout their lives after 55 (0 years of disability) is larger than 25% for all risk groups. The standard deviation of disabled lifeyears is larger for women than for males. The range of life years with disability is significantly larger for the obese men and women than for the smoking and non-smoking normal weight individuals. Extension of average duration of disability could be caused by 1) the same number of individuals spending more years disabled or 2) more individuals suffering from disability. As the distribution of disabled lifespan is wider for the obese, it means that the expansion of disability by obesity is due to more individuals suffering disability compared to those in other groups.

Table 2: Percentage of simulated males and females per health trajectories and risk factor group resulting from longitudinal microsimulation (frequency > 1%).

<table>
<thead>
<tr>
<th>Trajectory</th>
<th>Males</th>
<th>Females</th>
<th>Current smokers</th>
<th>Obese</th>
<th>Non-smoking normal weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>50.8%</td>
<td>64.6%</td>
<td>45.6% 64.6%</td>
<td>61.5%</td>
<td>68.1% 49.5% 64.2%</td>
</tr>
<tr>
<td>13</td>
<td>45.2%</td>
<td>30.7%</td>
<td>53.3% 33.9%</td>
<td>34.6%</td>
<td>27.0% 45.7% 30.9%</td>
</tr>
<tr>
<td>1213</td>
<td>2.0%</td>
<td>1.9%</td>
<td>0.7% 0.6%</td>
<td>1.6%</td>
<td>2.3% 2.4% 1.8%</td>
</tr>
<tr>
<td>12123</td>
<td>1.9%</td>
<td>2.8%</td>
<td>0.4% 0.9%</td>
<td>2.3%</td>
<td>2.6% 2.3% 3.0%</td>
</tr>
</tbody>
</table>

**Conclusion and discussion**

Microsimulation is an innovative and useful method to analyze multistate transition rates. Because both methods start out from empirically estimated transition rates, the aggregated results are consistent. More studies have started to use microsimulation to analyze individual

As opposed to the cohort approach offering average state specific life expectancy of an entire cohort, the individual biography or microsimulation approach provides the complete range and distribution of individual lifepaths that can be aggregated and analyzed in many different ways. Hence, the microsimulation analyses give a richer output than the multistate life table analysis.

Figure 4: Distributions of total lifespan for male smokers and female obese as a result of 100,000 simulations.

We have shown an application on health and disability transitions illustrating large differences in pathways to death for males and females and smoking or obese individuals. When looking at the frequency distribution of health trajectories we concluded that females
have less chance to die without ever having experienced disability. Hence, females are more likely to pass through disability and also to recover from it. Current smokers have the highest chance to follow a lifepath without disability because of early death: obese individuals the lowest because of early onset of disability. The distributions of age at death and life expectancy with disability show the highest variability for obese individuals and the lowest for non-smoking normal weight individuals. This means that not only is the disability duration longer for the obese, but also, more obese individuals will experience disability.

Table 3: Distributions of age at death and life years with disability, indicated by the 25% and 75% quartiles, the mean and the standard deviation.

<table>
<thead>
<tr>
<th></th>
<th>Age at death</th>
<th>life expectancy with severe disability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25%  mean</td>
<td>75%  st.dev.</td>
</tr>
<tr>
<td>Males</td>
<td>72.94 80.51</td>
<td>88.53 10.86</td>
</tr>
<tr>
<td>Smoking males</td>
<td>67.31 75.38</td>
<td>83.19 10.50</td>
</tr>
<tr>
<td>Obese males</td>
<td>72.61 80.75</td>
<td>89.26 11.39</td>
</tr>
<tr>
<td>Healthy males</td>
<td>75.12 82.02</td>
<td>89.81 10.45</td>
</tr>
<tr>
<td>Females</td>
<td>78.13 85.46</td>
<td>93.85 11.24</td>
</tr>
<tr>
<td>Smoking females</td>
<td>72.96 80.25</td>
<td>88.01 10.51</td>
</tr>
<tr>
<td>Obese females</td>
<td>77.62 85.11</td>
<td>93.53 11.38</td>
</tr>
<tr>
<td>Healthy females</td>
<td>81.08 86.94</td>
<td>94.11 9.81</td>
</tr>
</tbody>
</table>

The transition rates are estimated based on empirical occurrence-exposure rates of individuals who were aged between 55 and 105 in the years 1992 to 2004. Because the interviews are held bi-annually, we miss information on events during the interval, but this holds for both short unobserved periods of disability and recovery. We assume response to be independent of health status. In reality, it is more likely that individuals with severe health problems are less likely to respond, which would slightly underestimate our disability incidence. When aggregating the simulated lifepaths, the state-specific life expectancies are close to the multistate life table analysis. Continuous-time microsimulation can show several transitions within a year, which is not observed in the bi-annual survey. Therefore, the number of transitions are higher in the simulated than in the observed lifepaths. An elegant way to verify the microsimulation results is to cut off the simulated lifepaths in the same way as the empirical lifepaths are observed and compare prevalences at different ages. The estimation of the transition rates is based on the Markov assumption that only the current
state affects the transition probability to another state. Transition probabilities are often
dependent: in medicine, those with a history of disease are at increased risk of relapse.
Further research should investigate the validity of this Markov assumption and the effect of
past disability or the duration of disability on incidence and mortality.

References

"The effect of obesity on disability vs mortality in older Americans." Archives of Internal
Medicine 167(Apr. 23):774-780.

Cai, L. and J. Lubitz. 2007. "Was there compression of disability for older Americans from

"Estimation of multi-state life table functions and their variability from complex survey data
using the SPACE Program." Demographic Research 22(6):129-158.

Centers for Disease Control and Prevention and National Center for Health Statistics. 2006.

associated with underweight, overweight, and obesity." Journal of the American Medical
Association 293(15):1861-1867.

Conference of the International Microsimulation Association, 8-10 June 2009, Ottawa,
Canada.


