Comparison of statistical methods for calculating life expectancy after spinal cord injury

Michael J DeVivo1 · Gordana Savic2 · Hans L Frankel2 · Mohamed Ali Jamous2 · Bakulesh M Soni3 · Susan Charlifue4 · James W Middleton5 · John Walsh6

Received: 14 December 2017 / Revised: 15 January 2018 / Accepted: 16 January 2018 © International Spinal Cord Society 2018

Abstract

Study design   Retrospective observational.

Objectives   To compare results of several different methods for calculating life expectancy in the same sample of people with spinal cord injury (SCI), and critically assess their advantages and disadvantages.

Setting   Two spinal centres in Great Britain.

Methods   Survival status of persons with traumatic SCI injured between 1943 and 2010 with follow-up to 2015 was determined. Standardised mortality ratios (SMRs) were calculated using age at injury and current (attained) age, and compared. Life expectancy was then estimated using the SMR methods and compared with the results of a method based on multivariate logistic regression of a person-year dataset. Life expectancy estimates calculated by applying SMRs based on current age to general population period (current) and cohort (projected) life tables were also compared.

Results   The estimated life expectancies were significantly higher when the SMRs were based on age at injury. They were also higher when a general population cohort life table was used, particularly for younger ages. With the exception of the ventilator-dependent group, the life expectancy estimates derived from logistic regression were slightly lower than those derived from SMRs based on current age and a general population period life table.

Conclusions   The multivariate logistic regression of person-years method offers several advantages compared to the SMR method for calculating life expectancy after SCI, the main ones being: greater statistical power and precision with smaller sample sizes, the ability to include more predictive factors and to distinguish the otherwise confounded effects of current age, time post-injury, and calendar time.

Introduction

Long-term survival and life expectancy can be estimated in a variety of ways [1]. Several methods have been used to study long-term survival and life expectancy among persons with spinal cord injury (SCI) and each approach has its strengths and limitations. Life expectancy estimates vary based on the chosen method of statistical analysis [2]. One of the aims of this paper was to highlight the most commonly used methods for estimating life expectancy among persons with SCI, and how the choice of methods affects reported findings. Another aim was to demonstrate the significant differences between the results derived from different methods of analysis, by recalculating life expectancy in five different ways in the same sample of individuals with traumatic SCI.
Standard life tables

Initially, investigators relied on the use of standard life tables or the Kaplan–Meier approach to estimate the cumulative survival function over time following SCI, and these approaches are still occasionally used [3, 4]. Cohort methods like these are best utilized when the length of follow-up is relatively short [5–8]. However, as treatment practices changed and life expectancy increased, limitations of these methods became magnified. For a typical SCI study population, median (50%) survival is now more than 30 years after injury. Therefore, by the time a cohort study concludes and results are available, Kaplan-Meier or standard life table estimates of survival are substantially out of date, because the results reflect what life expectancy was when persons were enrolled in the study (accounting for subsequent changes in health care) rather than what they would be at the end of the study.

Standardized mortality ratios (SMR)

Given these limitations of standard life tables, investigators began to seek ways to adjust period (current) general population life tables to reflect the effect of having an SCI. The typical approach that has been used is to calculate standardized mortality ratios (SMR). This involves determining an expected number of deaths based on mortality rates for the general population of comparable age, sex, race, and length of follow-up. The general population mortality rates used in these calculations should be concomitant with the period of exposure of the study population rather than the most recent general population mortality rates. The SMR is then calculated as the ratio of the actual number of deaths occurring in the SCI study population to the expected number of deaths that would have occurred in the general population over the same period.

Given known differences in long-term survival based on injury severity, investigators typically calculated a separate SMR for each category of injury [9–13]. Originally this involved creating four neurologic groups (incomplete paraplegia, complete paraplegia, incomplete tetraplegia, and complete tetraplegia); however, to improve homogeneity, the current recommendation is to construct five groups in accordance with the International Standards for Neurological Classification of SCI (C1–4 with American Spinal Injury Association (ASIA) Impairment Scale (AIS) grade A, B, or C; C5-8 ABC; T1–S4/5 ABC; AIS D regardless of injury level; and ventilator-dependent regardless of injury level) [14, 15]. Life expectancy was then estimated by multiplying the mortality rates in the general population life table by the SMR for that neurologic category and recalculating the life table accordingly. Often, the most current general population period life table was used for SCI life expectancy calculations under the assumption that the SMR was constant over time. However, this assumption is inconsistent with the findings of several studies now indicating a trend toward a higher SMR or lower percentage of normal life expectancy for SCI than in the past [16–19]. Therefore, current SCI life expectancy has been slightly overestimated by studies based on this SMR approach.

As sample sizes, numbers of deaths, and lengths of follow-up increased, it became possible for investigators to calculate SMRs specific for age at injury or current (attained) age within each neurologic category [6–8, 20–23]. In general, these investigators found that the SMR declines substantially with increasing age. Moreover, the decrease in SMR by age is greater among persons with more severe injuries. As a result, we now know that use of a constant SMR for all ages will result in an overestimation of life expectancy at younger ages and an underestimation of life expectancy at older ages, with the bias being greater among more severely injured persons.

While calculation of SMRs grouped by age at injury has revealed an important trend toward decreasing SMR with increasing age, there are significant limitations to consider when interpreting these results. Most importantly, results across studies are not comparable unless the average length of follow-up is the same, which is rare. For example, assume there are two studies and, in both studies, the lowest age at injury group includes persons injured before age 20. Given the declining SMR with advancing age, if the first study is a 50 year follow-up study, the current experience of 65-year-old persons who are 50 years post-injury will be included in the youngest age group because they were injured before age 20, causing the calculated SMR for this young age at injury group to be significantly reduced when compared to the results of a second study limited to 10 years of follow-up.

Therefore, more recent studies have reported SMR values grouped more appropriately by current (attained) age rather than age at injury [22, 23]. This is accomplished by first creating a person-year data set in which each person followed for each year is treated as a separate observation. Thus, a person who dies during the fifth year after injury contributes five observations to the data file. For each year, a binary outcome measure is created indicating whether the person survived or died during the year. Therefore, for this example, the person would be considered alive for the first four observations and dead for the fifth observation. The expected and actual number of deaths then can be calculated from this person-year data set based on the current age during each follow-up year, and SMR values grouped by current age can be determined. These two studies [22, 23] went one step further by using actuarial methods [24] of graduation and data-smoothing techniques to adjust for excess variability and imprecision of SMR values caused by
small sample sizes. While this method lacks the power of statistical multivariate analyses described below, it allows more direct access to the data and development of assumptions regarding the determination of SMRs, particularly at older ages where exposure is often limited.

**Multivariable models**

**Cox regression**

Historically, the most common statistical method for evaluating the effects of multiple risk factors on long-term survival has been Cox proportional hazards regression analysis [25]. This approach, used in many studies of long-term survival following SCI, allows the simultaneous consideration of main and interactive effects of other known mortality risk factors such as gender, race, cause of injury, education, income, smoking, etc., in addition to age and neurologic status, and can accommodate changes in risk factors over time through the use of time-dependent covariates [11, 16, 19–21]. However, there are two significant limitations to the use of Cox models. First, the population is analyzed as a cohort. As such, the independent effects of current age, time post-injury, and calendar time are completely confounded in that they all increase uniformly with the passage of time. Thus, their independent effects on mortality risk cannot be assessed [26]. Moreover, the calculation of life expectancy based on a Cox model is mathematically cumbersome and requires several assumptions [27]. As a result, to our knowledge, no attempts have been made to estimate life expectancy following SCI directly from Cox models.

**Person–year logistic regression**

In 2000, Strauss et al. proposed the use of a pooled repeated observations method to calculate life expectancy following SCI [26]. Using this approach, the unit of analysis is the person-year rather than the person. The person-year data set is constructed as described previously. Risk factors are assessed separately each year, so each year has a unique combination of current age, year post-injury, and calendar year thereby allowing an assessment of their independent effects on mortality. Multiple logistic regression analysis is then used to determine the mortality odds ratio for each risk factor and the probability of dying each year given its unique set of risk factors. A life table can then be directly calculated from these age-specific probabilities of death in a more straightforward way than would be obtained from a Cox model. Many recent studies of life expectancy after SCI have adopted this logistic regression method of analysis [17, 18, 28–34].

**Methods**

To demonstrate the significant differences between SMRs grouped by age at injury vs. current age, we used the data set from the recently published 70 year follow-up study of traumatic SCI in Great Britain [18]. Details of the study sample, eligibility, data collection procedures, follow-up and death ascertainment can be found in that publication. The sample consisted of patients with traumatic SCI admitted to Stoke Mandeville and Southport spinal centres who were injured between 1943 and 2010, survived the first year post-injury, had residual neurological deficit on discharge and were British residents. Demographic and injury information was collected retrospectively from medical records. Survival status up to and including 31 December 2014, and death certificates for the deceased, were supplied by the Medical Research Information Service, Health and Social Care Information Centre (HSCIC), on behalf of the United Kingdom (UK) Office for National Statistics (ONS). The UK ONS also produced national life tables for the general population of England and Wales by age and gender.

Data analysis began by creating a person-year data set as described above. The expected fraction of a death for each person-year was then obtained from the age-specific mortality rate for the general population of England and Wales for the mid-year of follow-up of the study population. SMRs were then calculated as the ratio of observed to expected deaths for each neurologic category grouped either by age at injury or current (attained) age. These SMRs subsequently were used to adjust the most recent England and Wales general population period life table to reflect the presence of SCI. Additionally, actuarial methods of graduation and smoothing of the derived SMR values based on current age were used as in previous studies of life expectancy following SCI conducted in Australia [22, 23]. Next, a logistic regression model was developed using the same person-year data set. This was used to calculate current age-specific mortality rates for each neurologic category from which standard life tables were then constructed to determine life expectancies. The complete logistic regression model that was used for calculating the estimated life expectancy appears in the previously mentioned paper [18]. The results of these four methods for calculating life expectancy were then compared. Finally, in order to demonstrate the difference in life expectancy estimates when adjustments are made based on projected future general population improvements in age-specific mortality rates, the current age SMR method was also applied to the 2014 cohort (projected) general population life table for England and Wales.
Statement of ethics

The British mortality study was approved by Berkshire Research Ethics Committee (REC), REC reference number 11/H0505/1.

We certify that all applicable institutional and governmental regulations concerning the ethical use of patient identifiable data were followed during the course of this research.

Results

Sample

The sample included 5483 persons with traumatic SCI injured between 1943 and 2010 and treated at one of the two participating British spinal centres (Stoke Mandeville and Southport). Their survival status was followed up to 2015. The mean age at injury for the entire sample was 35.12 years (range 0.5–90.99, median 30.68), 79.7% were male, 1.4% had a ventilator dependent tetraplegia, 6.3% C1-C4 tetraplegia AIS/Frankel grade ABC, 23.4% C5-C8 tetraplegia ABC, 41.2% paraplegia ABC, and 27.7% incomplete lesion (all Ds). On 31 December 2014, there were 2958 persons (54% of the sample) still alive, 2322 (42.3%) had died and 203 (3.7%) were lost to follow-up (could not be traced by the UK Office for National Statistics).

Standardized mortality ratios (SMRs)

The differences between SMRs grouped by age at injury vs. current age are presented in Table 1. The overall SMR for each neurologic category is the same (bottom row), but the age-specific SMR values are significantly higher, particularly for younger ages with greater injury severity, when the data are grouped by current age rather than age at injury.

Logistic regression

The multivariate logistic regression analysis identified gender, current age, time since injury, injury level and AIS/Frankel grade, ventilator dependency and study period as strong predictors of mortality, and they were used for calculating the estimated life expectancy, as reported in the previous life expectancy paper [18].

Life expectancy

Life expectancy estimates for males with SCI from the British study by neurologic category based on SMR values calculated by age at injury applied to the 2012–2014 general population period life table for England and Wales appear in Table 2a. The corresponding life expectancy estimates based on SMR values calculated by current (attained) age appear in Table 2b. The calculated life expectancies are significantly higher when the SMR used is based on age at injury, but the difference decreases as age increases.

Life expectancy estimates based on actuarial graduation and smoothing of current age SMR values appear in Table 2c. As expected, these results are very similar to those found in Table 2b, but differ slightly at older ages where there is less precision regardless of chosen method of calculation due to much smaller sample sizes.

Life expectancy estimates for males with SCI from the same British study by neurologic category based on a logistic regression analysis of person-year data appear in Table 2d. This analysis uses the same underlying data and the same risk factors grouped in the same way as were used to produce the estimates in Tables 2a, b, and c. With the exception of the ventilator-dependent group, where small sample sizes result in greater variability and less precision of estimation for all methods, the life expectancy estimates derived from logistic regression (Table 2d) are slightly lower than those derived from the use of SMRs based on current age (Tables 2b and c).

Table 3 shows much longer life expectancy estimates when cohort (projected) general population life tables were used with SMRs based on current age. The percentage of normal life expectancy also differs slightly from the percentage derived from period general population life tables.

Discussion

The results demonstrated the differences in estimated life expectancy calculations depending on the statistical analysis method used.

In the SMR method, the divergence of results at old ages results from having a single SMR for age 61 and older (Table 2b) vs. the graduated and smoothed estimates of SMR by individual ages (Table 2c). Given the declining SMR with advancing age, when a single SMR value is calculated for age 61 and older, the estimated SMR at ages closer to 61 will be underestimated while the SMR at significantly older ages will be overestimated, with the degree of underestimation or overestimation depending on the distribution of exposure within the over 60 age group. This will result in an overestimation of life expectancy for persons in their 60s but an underestimation of life expectancy for persons in their 80s when a single SMR is used for all ages over 60, whereas either the graduated and smoothed SMR values or the logistic regression can distinguish mortality risks at older ages. In fact, the logistic regression estimates for all ages are more precise because the effect of
age can be modeled more specifically than by calculation of SMRs in a few broad age categories. Since the age factor in the person–year data set reflects the current age in that person–year, it is not surprising that life expectancy estimates in Table 2d more closely approximate those in Tables 2b and c than Table 2a. The differences between SMRs in a few broad age categories. Since the age factor in

<table>
<thead>
<tr>
<th>AIS D</th>
<th>SMR</th>
<th>95% CI</th>
<th>Para ABC</th>
<th>SMR</th>
<th>95% CI</th>
<th>C5-8 ABC</th>
<th>SMR</th>
<th>95% CI</th>
<th>C1-4 ABC</th>
<th>SMR</th>
<th>95% CI</th>
<th>Ventilated</th>
<th>SMR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Age at injury (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46–60</td>
<td>1.42</td>
<td>1.21–1.67</td>
<td>2.76</td>
<td>2.36–3.21</td>
<td>5.54</td>
<td>4.66–6.54</td>
<td>7.28</td>
<td>5.64–9.24</td>
<td>9.86</td>
<td>5.39–16.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61+</td>
<td>1.08</td>
<td>0.90–1.28</td>
<td>1.79</td>
<td>1.40–2.25</td>
<td>3.33</td>
<td>2.75–4.00</td>
<td>4.40</td>
<td>3.32–5.73</td>
<td>6.58</td>
<td>3.76–10.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All ages</td>
<td>1.49</td>
<td>1.37–1.62</td>
<td>2.66</td>
<td>2.49–2.83</td>
<td>4.63</td>
<td>4.27–5.01</td>
<td>6.12</td>
<td>5.27–7.07</td>
<td>10.79</td>
<td>7.95–14.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Current age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–30</td>
<td>1.71</td>
<td>0.69–3.53</td>
<td>5.00</td>
<td>3.57–6.81</td>
<td>8.76</td>
<td>6.29–11.88</td>
<td>11.69</td>
<td>5.34–22.19</td>
<td>53.85</td>
<td>21.65–110.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61+</td>
<td>1.27</td>
<td>1.15–1.41</td>
<td>2.00</td>
<td>1.82–2.20</td>
<td>3.57</td>
<td>3.16–4.02</td>
<td>4.50</td>
<td>3.62–5.53</td>
<td>6.67</td>
<td>4.18–10.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All ages</td>
<td>1.49</td>
<td>1.37–1.62</td>
<td>2.66</td>
<td>2.49–2.83</td>
<td>4.63</td>
<td>4.27–5.01</td>
<td>6.12</td>
<td>5.27–7.07</td>
<td>10.79</td>
<td>7.95–14.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Standardized mortality ratios (SMRs) for the British SCI samplea, by neurologic grouping and either age at injury (Table 1a) or current age (Table 1b)

AIS American Spinal Injury Association (ASIA) Impairment Scale, CI confidence interval, SCI spinal cord injury, SMR standardized mortality ratio

When estimating life expectancy for any individual with an SCI, factors beyond those contained in the data sets typically used in SCI life expectancy research may need to be considered. Several publications from an American study reported on the effect of numerous psychosocial factors and self-reported health status on life expectancy after SCI, but did not include information on comorbid conditions, such as heart disease, cancer, diabetes, etc. [30, 34, 35]. In the absence of SCI-specific data on the effects of these factors, general population data are often used. The usual approach is to obtain a measure of the general population age-specific excess death rate (EDR) for the additional factors, add those EDRs to the effect of the SCI, and recalculate the life table accordingly [2]. This should include both positive and negative factors affecting life expectancy whenever possible. However, when including additional factors, one must recognize that this requires an assumption that the effects of each additional risk factor are completely independent of all other factors and that the effect in the general population is the same as it is for persons with SCI. In many cases, these assumptions will not be entirely valid, so careful consideration is warranted when there are important additional risk factors to be included in an individual projection of life expectancy.

Excess death rates

When estimating life expectancy for any individual with an SCI, factors beyond those contained in the data sets typically used in SCI life expectancy research may need to be considered. Several publications from an American study reported on the effect of numerous psychosocial factors and self-reported health status on life expectancy after SCI, but did not include information on comorbid conditions, such as heart disease, cancer, diabetes, etc. [30, 34, 35]. In the absence of SCI-specific data on the effects of these factors, general population data are often used. The usual approach is to obtain a measure of the general population age-specific excess death rate (EDR) for the additional factors, add those EDRs to the effect of the SCI, and recalculate the life table accordingly [2]. This should include both positive and negative factors affecting life expectancy whenever possible. However, when including additional factors, one must recognize that this requires an assumption that the effects of each additional risk factor are completely independent of all other factors and that the effect in the general population is the same as it is for persons with SCI. In many cases, these assumptions will not be entirely valid, so careful consideration is warranted when there are important additional risk factors to be included in an individual projection of life expectancy.

Use of life expectancy estimates in individual cases

Life expectancy is by definition the average length of remaining life. Some persons will live longer while others will not live as long as their projected life expectancy. When used to resolve matters of litigation, a settlement based on a single estimate of life expectancy will almost always result in considerable deviation from the actual expenses to be incurred in the future by the person with SCI. To avoid situations where persons with SCI have insufficient funds to cover their remaining lifetime expenses, structured settlements and annuities should be encouraged whenever possible.
Applying a single measure of life expectancy to a life care plan assumes with 100% certainty that an individual will live exactly as long as their life expectancy and no longer. This assumption is never correct and only rarely close to what actually happens. To avoid the necessity of this assumption, a better alternative is to use the distribution of estimated survival probabilities from the calculated SCI-specific life table, as has been done in all previous studies of lifetime costs of care following SCI [36–40]. This approach also better accommodates changes in the life care plan as individuals age as well as the effects of future inflation and the choice of discount rate for use in present value calculations.

**Future improvements in life expectancy**

One issue of concern is the adjustment of current SCI life expectancy results to reflect possible future improvements.
as have occurred in the general population. In the United States, no such improvement in SCI life expectancy has occurred in the past 35 years, but this does not preclude future improvement [17]. In Great Britain, there have been small improvements in SCI life expectancy over the past three decades, but these have not kept pace with improvements in the general population [18]. For litigation in both Great Britain and Australia, courts have typically used cohort (projected) life expectancy tables that have built in assumptions regarding future improvements in life expectancy. In the United States, period (current) life tables are typically used without adjustment for future increases. As seen by comparing the results of Table 2b and Table 3, there are substantial increases in life expectancy when applying the same SMRs to reduced future general population mortality rates that incorporate anticipated long-term improvements in life expectancy. This is particularly true at younger ages when the anticipated improvements have more time to occur. If SMRs had been applied to projected improved future general population mortality rates 30 years ago (the approach used to produce estimates in Table 3), life expectancies for individuals with SCI would have been substantially overestimated. Moreover, as seen from the differences in Tables 2b and 3, while it may be tempting to apply the percentage of normal life expectancy derived from a period life table to a cohort life table to obtain a life expectancy estimate accounting for potential future improvement, such an approach is not exact, even under the assumption of a constant SMR over time. To facilitate comparison of studies of life expectancy after SCI, investigators should describe in detail any adjustments that have been made to accommodate potential future increases in life expectancy into current estimates and report both unadjusted and adjusted findings.

Conclusions

To facilitate comparison across studies and minimize bias in calculated life expectancy estimates, future studies of long-term survival after SCI that use the SMR method of analysis should be based on current age rather than age at injury. Cox models and logistic regression of person-year data are more powerful methods of analysis, requiring smaller sample sizes, and are preferable to methods based on SMRs. These multivariable models also have the advantages of being able to include more predictive factors as well as being capable of producing current estimates of life expectancy without the necessity of assuming a constant SMR over time since calendar time can be included as a risk factor in either type of model. Additional important advantages of the logistic regression method are its ability to distinguish the effects of current age, time post-injury, and calendar time and the more straightforward estimation of life expectancy when compared to Cox models.

Acknowledgements The study was financially supported by Buckinghamshire Healthcare NHS Trust Charitable Spinal Fund and Ann Mason Legacy for Spinal Research Fund, UK. The original part of the study [21] was supported in part by the National Institute on Disability and Rehabilitation Research Grant #H133.G90010, USA. The authors wish to thank Medical Records staff at Stoke Mandeville and Southport Spinal Centres, and especially Mrs Pauline Cato, for their help with medical notes retrieval.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References


