Assessing fitness, predicting outcome, and the missing axis

J. B. Carlisle*

Department of Anaesthetics, Torbay Hospital, South Devon Healthcare NHS Foundation Trust, Lowes Bridge, Torquay, Devon TQ2 7AA, UK.
* E-mail: john.carlisle@nhs.net

Summary. This article discusses how to estimate the risk of postoperative death, an outcome that affects an important minority of patients in the month(s) after scheduled surgery. In addition, it reflects on our inability to characterize the effects of surgery on quality of life. This outcome is of primary concern to the much more numerous survivors but is absent from all graphs of postoperative survival: it is the missing axis. The calculations discussed in the article are available online at https://sites.google.com/site/informrisk/.

Keywords: care, postoperative; statistics, survival analysis

Editor’s key points

- There is little information on quality of life after major surgery or conservative treatment.
- Mortality is partly predicted by age, sex, fitness and five comorbidities.
- https://sites.google.com/site/informrisk/ estimates perioperative mortality to support shared decision making.

We will now discuss in a little more detail the Struggle for Existence.†

The consequences of surgical intervention are difficult to interpret: postoperative follow-up is usually brief and incomplete, with little information on quality of life, and with little information on how long and good life would be without surgery. Quality of life is the axis missing from all graphs of postoperative survival. The current attempts by the UK government to measure societal happiness, which would provide a quality-of-life index, are unlikely to bear useful information for years.‡ However, age and death are linked in a predictable way with ill health and a worsening quality of life. Death tends to be preceded by 10–15 yr of increasing dependence, with proximity to death associated with a worse quality of life. The calculation of survival can therefore serve as a temporary proxy until we have better information on the quality of life.

Published data leave us predominantly with mortality 30 days after surgery. Scheduled surgery kills some patients who would otherwise live: the population of the UK would be more numerous if there was a moratorium on scheduled surgery for the next month. For instance, 10 people die each month in the UK after scheduled abdominal aortic aneurysm (AAA) repair, assuming a 30 day mortality of 0.024 in 420 patients having either endovascular aneurysm repair (EVAR) or open repair, while one might expect two deaths to occur per month in this AAA population, based upon common prognostic indicators combined with the risk of AAA rupture (see below).§ Similarly, most of the 95 of 1580 patients who die each month after scheduled resection of colorectal cancer would have lived—for a time—without surgery.‖ Secondly, some patients with lethal pathologies—such as AAAs or colorectal cancer—die from other causes, without the ‘lethal’ pathology being symptomatic, diagnosed or treated. Patients most likely to die from other causes are also most likely to be killed by ‘curative’ surgery on the primary pathology. Consideration of both statements leads to a broad conclusion: most scheduled surgeries might not be justified on the basis of published outcomes, while the most vulnerable populations would opt to avoid postoperative death and suffering by a more complete assessment of risk. Our drive to ‘do something’ leads to surgery and drugs when the ‘something’ should more often be a better assessment of benefit and harm, shared more adeptly with patients and their families. The precipitous drive to operate is so strong—in patients, families, and clinicians—that even an unbalanced emphasis on iatrogenic harm would inadequately redress the balance.

Death

The processes that determine survival act before, during, and after hospital admission. The perioperative selective pressures, of starvation, dehydration, immobility, and iatrogenic harm, including surgical injury and the inflammatory response to it, act within the context of the background survival. The aphorism ‘survival of the fittest’ remains constant throughout. The temporary rise in the absolute risk of mortality that follows scheduled surgery is greater in patients closer to death. In this article, I have supposed that the relative increase in risk is the same for everyone exposed to a similar selective pressure (operation), so that an operation that doubles the risk in the hardy will also double the risk in the frail, but the absolute change in risk will be different. I expect this assumption to be proven false, at least as a universal ‘truism’, but it is only by making the hypothesis in the first place that one can test it.

I present data provided by the UK government.⁵ Population births and deaths are almost completely recorded, with trends common to developed countries (I recommend that you access through the Internet one of the videos by the animated Hans Rosling).

For Permissions, please email: journals.permissions@oup.com
has been constant or has changed. Mortality trends have varied for people of different ages, although not necessarily in patterns that one might expect. When one estimates survival into the future, one might assume that this average trend will continue or that it changes in various ways. The exploration of the various assumptions one has to make forms the basis for estimating the uncertainty of predictions: an expanding penumbra of uncertainty haunts the most probable survival trajectory as it is projected into the future. Whatever assumption one makes about future survival, it is likely that some things will remain constant:

- Your mortality rate increases 10% each year, doubling every 7 yr;
- If you are a man your mortality rate is almost double (1.7 times) that of a woman.

People of the same age rarely have exactly the same mortality risk. However, survival is associated with age more strongly than with any other variable: if permitted, only two factors with which to estimate mortality risk, age, and gender would be most predictive. The multivariable analysis of postoperative survival should probably force the inclusion of factors known to independently associate with general survival, which would be age, gender, aerobic fitness, and five co-morbidities, and then determine the extent to which these variables fail to explain the observed survival and determine what might explain the discrepancy.

**Fitness and death**

Aerobic fitness is independently associated with survival. However, before discussing the value of knowing someone’s fitness, I will use this opportunity to warn against reliance upon any single factor. A prognostic approach based upon clinical pathology combined with physiology cannot currently explain age-related increases in mortality. In time one might be able to measure the features of the telomere, RNA transcriptase, Golgi apparatus, apoptotic activity, or whatever, which when combined with clinical variables perfectly predicts the observed mortality. However, even then, such an approach would be a rather expensive way of asking someone their date of birth. From a prognostic point-of-view, the understanding of what makes people die is only useful if it improves the precision of one’s predictions.

The most widely used prognostic measure from cardiopulmonary exercise testing (CPET) is the peak oxygen consumption \( V_O_2 \) in \( \text{ml kg}^{-1} \text{min}^{-1} \). Other prognostic measures derived from CPET include the anaerobic threshold, ventilatory equivalents for oxygen absorption \( V_E/V_O_2 \), and carbon dioxide excretion \( V_CO_2/V_O_2 \). Expected peak \( V_O_2 \) average:

- \( 3.5 \times [18.4 - (0.16 \times \text{age})] \) for men;
- \( 3.5 \times [14.7 - (0.13 \times \text{age})] \) for women.

Mortality risk multiples by 1.15 for a shortfall in peak \( V_O_2 \) of 3.5 ml kg\(^{-1}\) min\(^{-1}\) and by 0.87 for the achievement of a peak \( V_O_2 \) of 3.5 ml kg\(^{-1}\) min\(^{-1}\) in excess of predicted.\(^{10}\) These equations can be adjusted for BMI as they overestimate mortality risk in the obese and underestimate mortality in the underweight.

Unique diagnostic information can be extracted from the absolute values of CPET variables and the patterns that they display during exercise. However, their ability to provide unique prognostic information is generally limited: the prognostic information extracted from one variable makes redundant most of the information contained in other CPET variables. However, different variables contain some unique prognostic information, demonstrated by closer associations of combinations of CPET variables with survival than individual measurements. There is no combination that is generally accepted as superior to other combinations. The peak \( V_O_2 \) must be one component of any combination that is used to predict general survival: it is the only CPET variable with quantified changes in mortality risk associated with deviation from detailed age- and gender-specific normal values. The ventilatory equivalent for \( CO_2 \) excretion \( V_E/V_CO_2 \) adds a prognostic value, particularly in studies of patients with heart failure and pulmonary disease.\(^{11,12}\)

Aerobic fitness alone cannot account for the increase in mortality risk between the ages of 20 and 90 yr. The decrease in fitness with age, characterized by falling peak \( V_O_2 \), matches the increasing mortality risk up to the age of \( \sim 45 \) yr, at which point fitness ‘accounts’ for \( \sim 66\% \) of the observed increase in mortality. By the age of 60 yr, fitness ‘accounts’ for 30% of the observed increase in mortality, at 70 yr 18%, at 80 yr 9%, and at 90 yr only 5%.

**Co-morbidities and death**

The diagnoses of myocardial infarction, heart failure, stroke, and peripheral arterial disease each chronically increase mortality risk \( \sim 1.5 \) times, compared with people who do not have these diagnoses.\(^{13-23}\) In the absence of myocardial infarction (MI) or stroke, the diagnoses of angina and transient ischaemic attack multiply risk by \( \sim 1.2 \) times. Recent strokes and MIs are associated with more substantial increases in risk. My attempts to calculate trajectories of risk following these events are incomplete, but suggest that it might take up to 2 yr for relative risk to settle at 1.5 times the population average, with risk as high as 15 times average at 3 months after an MI, six times average at 6 months and three times average at 1 yr. The absolute increase in risk associated with each diagnosis will be proportionately higher in people who had higher pre-morbid risks, for instance elderly men. A diagnosis will incur a variable risk among a homogenous group of people: the magnitude of risk associated with an MI or heart failure is probably best elicited through the extent of cardiorespiratory compromise measured during a CPET.

Renal failure is the fifth co-morbidity that consistently increases mortality risk, by between 1% and 2% for each 1 ml min\(^{-1}\) decrease in estimated glomerular filtration rate (eGFR), particularly below 75 ml min\(^{-1}\) 1.73 m\(^{-2}\) with the relative risk increasing by 1.8–3.3 times with a reduction in...
Renal function declines with age, so inclusion of renal impairment in the risk model requires comparison with the average function, adjusted for age and sex. One popular formula to calculate eGFR is the ‘modification of diet in renal disease’:

- \[ \text{eGFR} = 32788 \times \left( \frac{\text{creatinine} (\mu \text{mol litre}^{-1})}{1.73 \text{ m}^2} \right)^{-1.154} \times \text{age}^{-0.203} \]
  - for men;
- \[ \text{eGFR} = 32788 \times \left( \frac{\text{creatinine} (\mu \text{mol litre}^{-1})}{1.73 \text{ m}^2} \right)^{-1.154} \times \text{age}^{-0.203} \times 0.742 \]
  - for women.\(^26\)

Patients with one of these diagnoses may reach a prevalence of more than 20% in some historical elderly cohorts; the assumption that the ‘average’ age-related mortality risk represents people without this diagnosis would be incorrect in such a cohort. For simplicity, I have not attempted to account for disease prevalence in different cohorts, which would lead to a relative 12% overestimation of risk in a cohort with a 20% prevalence of these diagnoses and a relative 20% overestimation of risk if the prevalence was 40%.

### Death online

I have published a calculator online through a website that is being developed to support shared decision-making: [https://sites.google.com/site/informrisk/](https://sites.google.com/site/informrisk/).\(^27\) If you subscribe to that website’s group, you will gain editing access for the calculator, allowing you to enter your own data. You can access videos describing how to use the calculator, with descriptions of the calculations and their strengths and weaknesses. After you have entered the data, the calculator will provide both tables and graphs of survival over subsequent years, and once you have entered the data, the calculator will provide both tables and graphs of survival over subsequent years, and expected mortality 1 month after surgery. You can use it prospectively for your patients, and for retrospective audits of outcome following decisions to have (or avoid) surgery, as far back as 1981.

There are two stages that are prone to the most uncertainty. First,

- how much does an operation temporarily increase baseline risk?

Observational studies that report mortality in the month after scheduled surgery, such as that by Noordzij and colleagues,\(^28\) can indicate the postoperative increase in risk. Noordzij and colleagues reported deaths after 3.7 million elective operations in the Netherlands between 1991 and 2005. For each operation, they reported the median age and sex distribution of patients, and the proportion with co-morbidities. I therefore estimated the expected monthly mortality rate for each population and compared this with the observed. Table 1 shows the results for five operations. My local data suggest that open AAA repair temporarily increases mortality risk 12 times, half the relative increase that I calculated from Noordzij and colleagues’ paper. There are numerous possible reasons for this discrepancy, which might or might not be specific to AAA surgery. The lowest expected mortality rate (0.10%) was in the fairly young, female and healthy cholecystectomy population, so the observed mortality rate of 1.54% was over 15 times that expected, or over 7 times after my adjustment.

Irrespective of the inaccuracies of these figures, it is clear that scheduled surgery is a potent selective pressure, equivalent to temporarily ageing patients by 7 yr after total hip replacement (2 is 1.1)\(^7\) or 26 yr after open AAA repair (12 is 1.1).\(^26\) Mortality is elevated for prolonged periods after surgery, which will be accompanied and surpassed by the duration of ill health and increased dependency experienced by some survivors.

Secondly,

- What is the survival trajectory—and quality of life—without surgery?

The answer to this question is of less immediate professional concern to the anaesthetist or intensivist, but it is fundamental for the patient and critical to the validity of choice, consent, and dissent. The information void I alluded to at the beginning of this article is most profound when considering the natural history of surgical pathology treated without surgery.

On the website I cited I have used the risk of a patient dying from the rupture of an AAA to illustrate how the comparison of survival curves, one with surgery and one without, could support informed decision-making, for both patients with AAAs considering surgery and men considering ultrasound screening. The increased risk of rupture with AAA diameter is poorly characterized, but AAAs probably behave more predictably than pathologies such as colorectal adenocarcinomas. At the moment I have estimated that the annual risk of rupture for an AAA with diameter measured in millimetres is \((0.0001 \text{ mm}^2 \times 0.0068 \text{ mm}+0.1215)\). This equation gives the following annual percentage risks of rupture: 45 mm, 1.8%; 50 mm, 3.2%; 55 mm, 5%; 60 mm, 7.4%; 65 mm, 10.2%; and 70 mm, 13.6%.

<table>
<thead>
<tr>
<th>Operation</th>
<th>Number of operations</th>
<th>Expected mortality (%)</th>
<th>Observed mortality (%)</th>
<th>Observed/expected (O/E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oesophagectomy</td>
<td>20 364</td>
<td>0.21</td>
<td>5.5</td>
<td>27</td>
</tr>
<tr>
<td>Open AAA repair</td>
<td>78 826</td>
<td>0.42</td>
<td>10.2</td>
<td>24</td>
</tr>
<tr>
<td>Cholecystectomy</td>
<td>104 071</td>
<td>0.10</td>
<td>1.54</td>
<td>15.2</td>
</tr>
<tr>
<td>Total hip replacement</td>
<td>387 232</td>
<td>0.39</td>
<td>1.6</td>
<td>4.1</td>
</tr>
<tr>
<td>Hernia repair</td>
<td>367 885</td>
<td>0.14</td>
<td>0.28</td>
<td>2.1</td>
</tr>
</tbody>
</table>

\(eGFR\) to 15 ml min\(^{-1}\) 1.73 m\(^{-2}\).\(^{24,26}\) Renal function declines with age, so inclusion of renal impairment in the risk model requires comparison with the average function, adjusted for age and sex. One popular formula to calculate eGFR is the ‘modification of diet in renal disease’:
I have used AAA repair as an example because the pathology is usually physically asymptomatic until rupture (or elective surgery), although some patients are burdened by fear after diagnosis: for example, one patient I saw with a screen-detected AAA had remained within a 30 mile radius of his home for nearly 3 yr, until its diameter exceeded 5.5 cm and he proceeded to surgery. Fear of death is usually calmed by openly discussing it, with patients becoming paradoxically more sanguine after they have been reminded of their certain mortality and all the other threats to their survival, which will either be unaffected or increased by AAA repair. The removal of one ‘ticking time bomb’ accelerates the clock speed on a number of other bombs!

Obituary

Charles Darwin died in 1882, aged 73 yr. He would not have appreciated my abuse of his words, ‘the greater the struggle for existence, the greater the bearing on surgical selection’. People avoid things that will kill or incapacitate them, so it is surprising how many people subject themselves to scheduled surgery. The frail often avoid threats meticulously, preferring to remain in the relative safety of home. Mrs Smith made a rare exception when she visited hospital to be told by a surgeon that the caecal tumour, detected on a computed tomographic scan and unaccompanied by hepatic or pulmonary metastases, was the likely cause of her anaemia. Her next door neighbour, Mr Brown, accepted the invitation for AAA screening, knowing that by doing so he would reduce his risk of dying from aeurysmal rupture. Meanwhile, Mr Davies—a keen runner in his youth—was admitted for a total knee replacement, in the same week that his niece had her gall bladder removed by keyhole surgery. Six months later, Mrs Smith remains asymptomatic and uncut: she does not get up enough head of steam to notice her anaemia, while her bowel habit is unchanged. Mr Brown is worried. His abdominal aorta is now 4.6 cm diameter, but was 4.4 cm diameter 6 months ago. Mr Davies has not taken up running again because he is dead, his niece having just attended his funeral. She is occasionally bothered by dyspeptic symptoms and has a numb area on her side with occasional shooting pains. Still, ‘mustn’t grumble’, she thinks, ‘at least I’m alive’.

Declaration of interest

None declared.

Funding

None.

References


