Cancer statistics: a survival guide

Jason Oke *and* ***Tom Fanshawe*** *expose four simple biases that can change our understanding of cancer survival rates and skew comparisons made between countries*

In the UK, survival from cancer has increased dramatically over the last 40 years. At the beginning of the 1970s, only a quarter of men and one third of women survived beyond five years after their diagnosis. The latest figures show that net five-year survival is 49% in men and 59% in women (Figure 1). Survival from cancer in men has nearly doubled and in women it has increased by more than two-thirds.

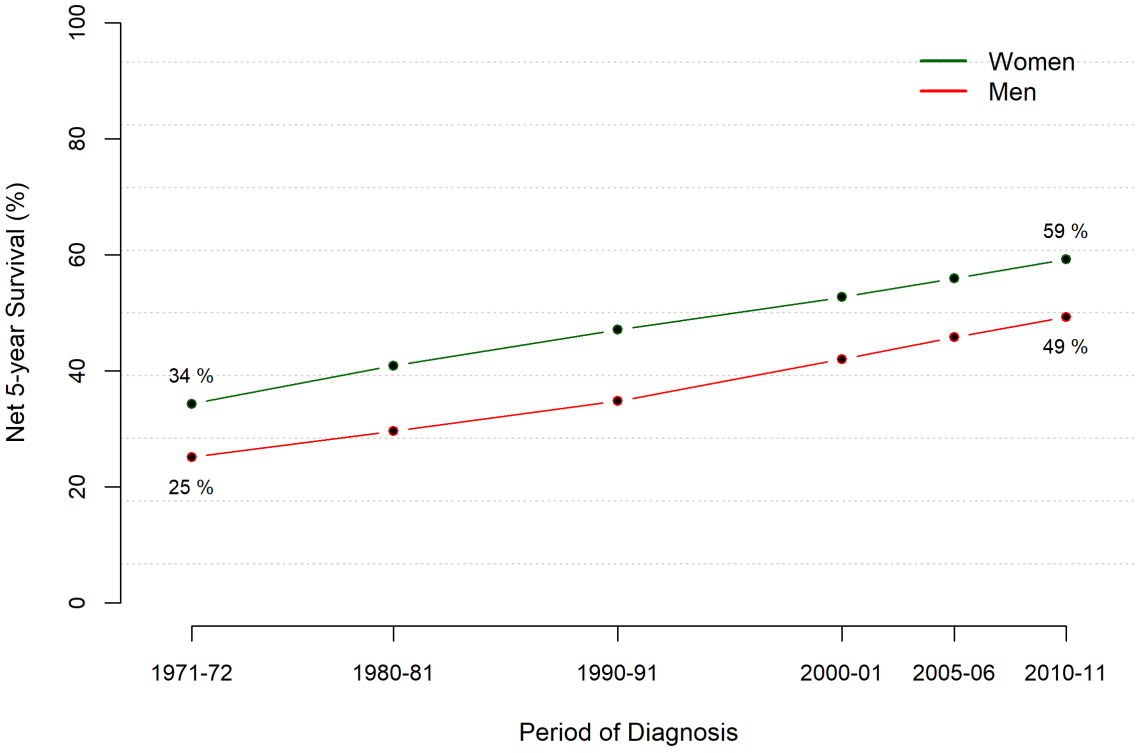

FIGURE 1 Plot of age-standardised five-year cancer survival (the percentage of people diagnosed with cancer who are still alive five years later, allowing for age) in England and Wales. Figure drawn using data from bit.ly/35Yypps. Please note that the interval of measurement varies over time. Confidence intervals are not presented, but typically these would be no more than ± 0.4%.

The continuing improvement in survival appears extremely encouraging and strongly suggests that we can now treat and manage cancer better than ever before. Yet, despite this improvement, the UK's cancer survival rates are considered poor when compared to those in other European countries. For example, Baili *et al.* compared survival from cancer in European adults and found that the UK and Denmark had lower survival than other Northern European countries such as Sweden and France (Table 1).[1]

TABLE 1 Age standardised relative survival for all cancers combined at five years after diagnosis. Adapted from Table 2 of Baili *et al*.[1] Standard errors are derived using asymptotic formula for proportion and denominators from De Angelis *et al.*[2] Figures for US survival and population coverage are from the National Cancer Institute.[3, 4]

Country	**Five-year survival rate (%) (approx. standard error)**	**Proportion of population covered by cancer registration (%)**
European average	54.2 (0.02)	50
Sweden (highest)	64.7 (0.08)	100
France	58.6 (0.11)	25
Denmark	50.9 (0.10)	100
Ireland and UK	*50.1 (0.03)*	*100*
Bulgaria (lowest)	38.7 (0.10)	100
United States		
White ethnicity	68.2 (0.03)	24.9
Black ethnicity	63.7 (0.10)	25.6

Responding to the UK's long-standing deficit in cancer survival, in 2011 the Department of Health made a commitment to halve the gap in five-year cancer survival rates between England and the best performing countries in Europe, stating that: "If England was to achieve cancer survival rates at the European average, then 5,000 lives would be saved every year. If England was to achieve cancer survival rates at the European best, then 10,000 lives would be saved every year."[5]

Nevertheless, the use of survival rates as a measure of performance improvement in health care has been questioned. International comparisons can be affected by differences in how tumours are defined and how well registries record and collect follow-up data. Estimating survival also takes time to calculate, which is why, in 2019, we refer to cancers diagnosed in 2011 as "recent data". Similarly, within-country comparisons over time are influenced by changes in how the cancer is defined and diagnosed. The purpose of this article is to examine four pervasive biases when comparing survival between countries and over time, and to explain how they can make observed changes in survival misleading.

Lead time bias

Imagine two people who both develop cancer on their 60th birthday, initially without showing any symptoms. Patient A eschews modern medicine. He doesn't attend health screening or even bother the doctor when the symptoms appear later on. One day, at age 65, he is admitted as an emergency case and diagnosed, but by this time the cancer has spread, and he survives only a year longer.

Patient B is careful, heeds the advice of medical professionals and attends an annual health screen. The cancer is found early and treatment begins, but unfortunately in this case it is ineffective and Patient B also succumbs.

Figure 2 illustrates this scenario. Patient A's survival after diagnosis is just one year whereas Patient B's is five years. We might conclude, incorrectly, that Patient B's outcome was more favourable than

that of Patient A. This phenomenon is known as *lead time bias*. It shows that simply diagnosing cancer earlier can appear to improve survival rates.

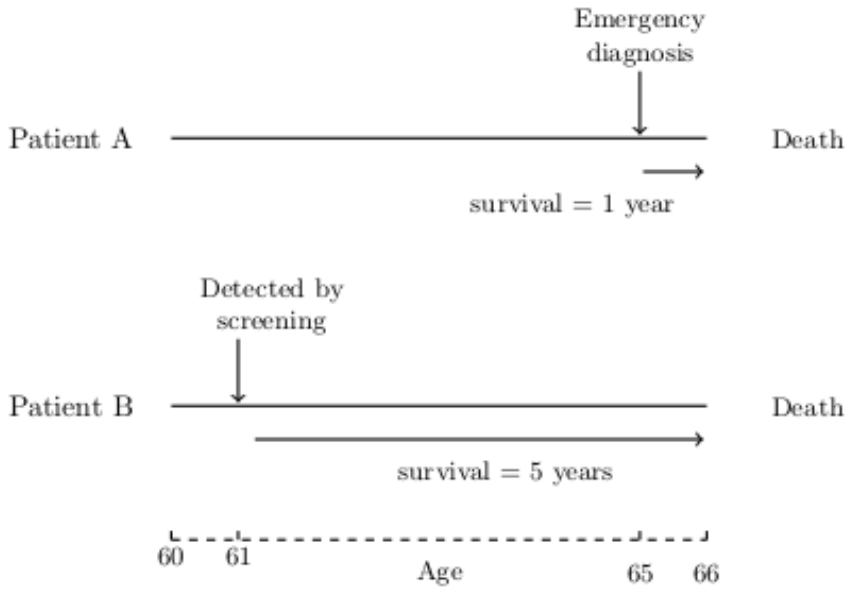

FIGURE 2 Hypothetical example of two patients with the same natural history of cancer but different diagnostic pathways. Patient A receives no treatment; Patient B receives treatment, but it does not extend life.

Early detection methods such as cancer screening will advance the time of diagnosis (lead time) but the fact that screen-detected cancers have better survival rates than cancers detected by symptoms does not on its own provide evidence that screening saves lives. Similarly, the fact that some countries have better five-year survival than others does not prove that their patients live longer and have better outcomes: they might just have more Patient Bs than Patient As.

Overdiagnosis bias

In 1981, Doll and Peto observed that "if someone ever invents a method of screening apparently healthy men for prostate cancer, the apparent incidence rates may be expected to rise quickly by several hundred percent".[6] Soon after, just such a test was introduced – the prostate specific antigen test. The incidence of diagnosed prostate cancer (the number of new cases diagnosed per year) more than doubled between 1980 and 2013. Prostate cancer survival rates also more than doubled. Doll and Peto were able to foresee this dramatic rise because they knew that prostate cancer had the potential to be overdiagnosed.

Cancer overdiagnosis is the detection and diagnosis of disease that bears all the hallmarks of a cancer but is of a type that does not progress to cause harm. As Van den Bruel wrote in 2015: "Every large dangerous cancer was small once, but not every small cancer will become large and

dangerous".[7] Overdiagnosis should not be confused with a false positive result, which may be overturned by more in-depth investigation.

Overdiagnosed cancers are most likely to be diagnosed incidentally (when looking for another problem) or through population-wide screening. The problem is that when the cancer is in the early stages, it is difficult to distinguish between lethal cancers and *indolent* types that are not a threat to the patient. Autopsy studies suggest that nearly 60% of men will have an undiagnosed prostate cancer by the age of 80, but the majority of these would be harmless and never cause symptoms.[8]

The overdiagnosis of cancer can lead to overtreatment, stretch healthcare resources and impact on survival measures. By diagnosing more cases of indolent prostate cancer, survival rates appear to increase (see "The impact of overdiagnosis").

The former mayor of New York and presidential candidate Rudy Giuliani, a prostate cancer survivor, is quoted (bit.ly/38foyBg) as saying: "My chance of surviving prostate cancer… in the United States? Eighty-two percent. My chance of surviving prostate cancer in England? Only 44 percent under socialised medicine." At the time, most prostate cancer in the US was detected via screening, whereas in Britain most diagnoses were made after the patient developed symptoms. It is possible that some of the difference can be attributed to overdiagnosis even if treatments are equally good in the two countries, so the figures quoted by Giuliani do not necessarily reflect the chances of surviving for patients with comparable disease severity and progression.

Length time bias

Suppose we were interested in estimating average alcohol consumption in a public house. One way to do this would be simply to turn up at a random time and survey all the people in the pub at that time about the amount they had drunk or were intending to drink. However, as some people stay in the pub all night and others just pop in for a swift half-pint, any snapshot of time like this would inevitably include all of the all-nighters but miss most of those who were only there a short time. If people who spend more time in the pub also drink more, our estimate of alcohol consumption per person would be too high. This is called *length time bias*.

Screening for cancer incurs length time bias in much the same way. If we screen (sample) at a single point in time, we will tend to find many people with slow-growing disease but fewer with faster-growing cancers (Figure 3). As people with slow-growing, less aggressive cancers tend to survive longer whether detected or not, this may cause us to overestimate the average survival time.

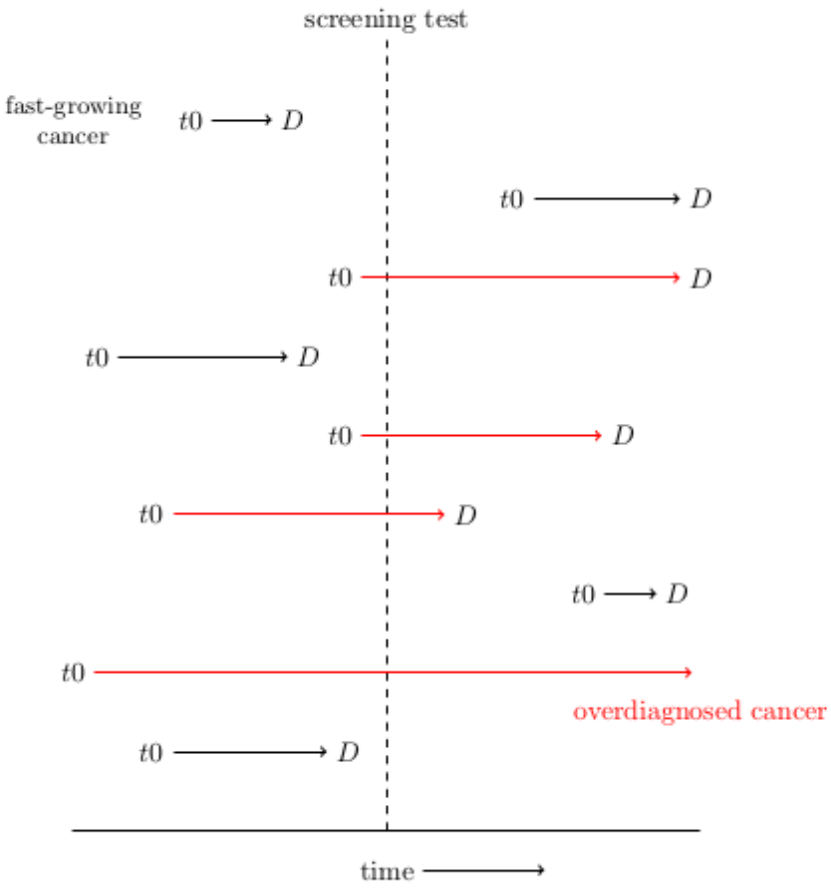

FIGURE 3 Example of length time bias. *t*0 = start of disease (cancer is detectable but not symptomatic) and *D* = Death. Length of arrows represent different survival times. Red arrows indicate which cancers are detected by screening.

Length time bias is related to overdiagnosis. As screening tends to detect slower-growing cancers, it will also tend to find those that will never progress at all. These are the overdiagnosed cancers discussed in the previous section. When comparisons of survival are made between countries with and without active screening programmes, or comparisons are made before and after the introduction of screening, the effects of length time bias needs to be considered along with lead time bias.

Stage migration

The American humourist Will Rogers is attributed with a remark – thought provoking if not entirely complimentary – about migration during the economic depression of the 1930s: "When the Okies left Oklahoma and moved to California, they raised the average intelligence level in both states."

Quoting Rogers, the epidemiologist Alvan Feinstein observed that changes to the way cancer is *staged* can artificially result in improvements in survival rates.[9] Cancer stage reflects the extent of a cancer in the body and is usually based on the size of the tumour, whether lymph nodes contain cancer, and whether the cancer has spread from the original site to other parts of the body.

The Will Rogers effect can happen when patients previously classified in a 'good' stage are migrated to a 'worse' stage, perhaps as a result of using new diagnostic technology. If the survival rate of the people migrating is worse than the prognosis of other people in their original group, but better than that of other people in their new group, survival can appear to improve in both groups.

Feinstein and colleagues demonstrated this effect in 131 patients diagnosed with lung cancer. They staged patients using both the old system and more modern imaging procedures, and recorded six-month survival.

Table 2 shows the results: of the 42 people classified as stage I-II under the old system, 25 migrated to stage III (a worse stage) with the new technology. Because the survival of the 25 was lower (56%) than those who did not migrate (84%), but higher than the people remaining in the worse stage (36%), their migration improved survival in both the new stage III group (42% vs 36%) and the new stage I-II group (84% vs 76%). Feinstein *et al*. concluded that the apparent improvement in survival was likely to be nothing more than a statistical artefact caused by reclassifying patients to a higher stage.

TABLE 2 Illustration of stage migration. Data adapted from Feinstein *et al*.[9]

		Stage using new technology - survival % (*N*)		Six-month survival by stage (under old technology)
		I – II	**III**	
Stage under old system	**I – II**	84% (*N* = 42)	56% (*N* = 25)	*76% (N = 67)*
	III	-	36% (*N* = 64)	*36% (N = 64)*
Six-month survival by stage (using new technology)		*84% (N = 42)*	*42% (N = 89)*	-

The Will Rogers effect has the potential to affect any disease in which changes in diagnostic practices occur, often with the addition of new imaging. Different diagnostic protocols are often used in different countries, so this is another reason why comparing survival rates between countries can be difficult.

A possible solution?

Some have suggested that a way to mitigate the four biases discussed in this article – lead time, overdiagnosis, length time and stage migration – is to use cancer-related mortality (the number of people who die of cancer, as a proportion of the whole population) as an alternative measure to survival. Mortality rates are less affected than survival rates by these biases because they are based on the entire population at risk, rather than only those who are diagnosed, and so cancer-related mortality is arguably less affected by changes in diagnostic practices.

For example, Welch *et al*. compared the change in incidence over 45 years with change in survival and cancer-related mortality for the 20 most common cancers in the US.[10] They found that survival rates were closely related to incidence, primarily as a result of changes in diagnosis, but almost unrelated to mortality rates. At the time when Rudy Giuliani compared American and British survival statistics, mortality from prostate cancer was strikingly similar in both countries: 26 deaths per 100,000 men in the US versus 27 per 100,000 in Britain.

Mortality rates have their limitations, however. Deaths may arise from cases diagnosed many years earlier, and so may not reflect current management of cancer. They also need to be adjusted for genuine changes in cancer incidence over time, such as might be caused by increasing or decreasing exposure to risk factors like tobacco smoking.

Conclusion

As this article has shown, cancer survival statistics can be more difficult to interpret than they first appear. They can be affected by biases that make comparisons over time and between countries difficult. By advancing the time of diagnosis, and not necessarily by changing the outcome, we can improve apparent survival rates (lead time bias). Changing the definitions of cancer, to include a milder spectrum of disease (e.g. overdiagnosis), or by introducing a screening programme (length time bias) or new imaging technology (stage migration), can have the same effect.

Many cancer treatments have improved substantially in recent years, saving and prolonging lives. But when advances are made, as demonstrated in randomised control trials, it can be difficult, in registry data, to distinguish these from spurious differences caused by the issues described above. These issues have been known to epidemiologists and statisticians for some time, but still cause confusion, even to some health professionals.

Cancer statistics play a key role in shaping health policy and are more visible to the public than ever (bit.ly/2sybkv9), and so increasing awareness of these biases is important. The first question to ask when interpreting changes in survival is therefore whether they truly reflect differences in the way cancer is treated, or only changes in the way it is diagnosed.

About the authors

Jason Oke is a senior medical statistician in the Nuffield Department of Primary Care Health Sciences at the University of Oxford.

Thomas Fanshawe is a senior medical statistician in the Nuffield Department of Primary Care Health Sciences at the University of Oxford.

BOX

The potential impact of overdiagnosis

Suppose we simplify prostate cancer into two types: (i) a progressive type that causes symptoms and death if untreated, and (ii) a non-progressive indolent type that will never lead to symptoms and never cause death. Before the era of prostate specific antigen testing, the diagnosis would have occurred after symptomatic presentation and so most cancers would be type i. In 1971, the age-standardised incidence was 30 per 100,000 and the five-year survival was 34-38%. These figures provide us with an estimate of the rate of progressive prostate cancer in the population. If we assume that the rate of progressive cancer has been stable over time, then the difference between the current incidence and the historical incidence provides us with an estimate of the amount of the indolent form that Doll and Peto predicted.

In 2016, the age-standardised incidence was 171 per 100,000. Subtracting the historical rate from the current implies that 141 per 100,000 diagnosed cases are overdiagnosed. As they are the non-progressive indolent type, the five-year survival for these cancers will be close to 100%. Then, the overall five-year survival rate in 2016 would be

$$\left(\frac{30}{171}\right) \times 38\% + \left(\frac{141}{171}\right) \times 100\% = 89\%.$$

The most recent five-year survival figures for prostate cancer from Cancer Research UK are 85% (bit.ly/35X61np).

It may be unrealistic to assume that both the rate of progressive disease and the prognosis for these patients is stable over time. We could have assumed that the rate of progressive prostate cancer increased from 30 to 45 per 100,000 and five-year survival also improved from 38% to 57%, as well as some overdiagnosis of indolent disease (126 per 100,000). In this scenario, the five-year survival rate would also be

$$\left(\frac{45}{171}\right) \times 57\% + \left(\frac{126}{171}\right) \times 100\% = 89\%.$$

Clearly, both of these illustrative calculations are consistent with big improvements in survival, but the former is driven entirely by overdiagnosis whilst the latter is driven by both overdiagnosis and an increase in progressive disease. In practice, it is difficult to tell which one is the correct explanation.

References

1. Baili, P., Di Salvo, F., Marcos-Gragera, R., Siesling, S., Mallone, S., Santaquilani, M., *et al*. (2015) Age and case mix-standardised survival for all cancer patients in Europe 1999-2007: Results of EUROCARE-5, a population-based study. *Eur J Cancer*, **51**(15), 2120-9.
2. De Angelis, R., Sant, M., Coleman, M. P., Francisci, S., Baili, P., Pierannunzio, D., *et al*. (2014) Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE-5-a population-based study. *Lancet Oncol*, **15**(1), 23-34.
3. National Cancer Institute. Five-Year Cause-Specific Survival (%) for Malignant cancers, 2009-2015 Both sexes by Race/Ethnicity and Primary Site. 2019.
4. National Cancer Institute. Number of Persons by Race and Hispanic Ethnicity for SEER Participants (2010 Census Data).
5. Department of Health. Improving outcomes: A strategy for cancer aims to help the reformed NHS deliver cancer outcomes that are amongst the best in the world. 2011.
6. Doll, R., Peto, R. (1981) The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. Oxford; New York: Oxford University Press. 1197-312 pp.
7. Van den Bruel, A. (2015) The triumph of medicine: how overdiagnosis is turning healthy people into patients. *Fam Pract*, **32**(3), 364.
8. Bell, K. J. L., Del Mar, C., Wright, G., Dickinson, J. and Glasziou, P. (2015) Prevalence of incidental prostate cancer: A systematic review of autopsy studies. *Int J Cancer*, **137**(7), 1749-57.
9. Feinstein, A. R., Sosin, D. M. and Wells, C. K. (1985) The Rogers, Will Phenomenon - Stage Migration and New Diagnostic-Techniques as a Source of Misleading Statistics for Survival in Cancer. *New Engl J Med*, **312**(25), 1604-8.
10. Welch, H. G., Schwartz, L. M. and Woloshin, S. (2000) Are increasing 5-year survival rates evidence of success against cancer? *J Am Med Assoc*, **283**(22), 2975-8.