Editorial

The best available approximation to the truth

Stephen J. Richards[1]
Longevitas Ltd, Edinburgh, UK

[1]Correspondence to Stephen J. Richards, Longevitas Ltd, 4 Caledonian Place, Edinburgh, EH11 2AS. Email stephen@longevitas.co.uk

> [w]e are much better off if we know the best available approximation of the truth
>
> Sagan (1997, p.266)

The astronomer and astrophysicist Carl Sagan emphasised that science always strives for underlying truths. However, his quotation above highlights that, at any given point in time, science offers the best workable approximation to the truth, not necessarily the truth itself. There are parallels with actuarial work, where often what is needed is not the truth in great detail, but an approximation to the truth which reliably captures the broad essence of a phenomenon.

But what is today's "best approximation of the truth" for actuarial work, particularly in the field of mortality and longevity projections? The modern actuary could be forgiven for envying her historical counterparts. Historical mortality-improvement assumptions would have been set simply, usually with a single deterministic scenario supplied by the CMI. In contrast, today's actuary must consider a wide variety of stochastic projection models, and possibly also projections disaggregated by cause of death. Without a focused specialism in this particular area, how is an actuary to assess the suitability of a particular model? More importantly, how can we judge if a particular model is fundamentally flawed or not?

One approach is for a firm to have dedicated specialist resources of its own. For example, in the 20th century few UK life offices spent much time analysing mortality rates, and the task would often have been delegated to a junior actuary or student for a few weeks' work. In contrast, many large life offices in the UK today have a dedicated longevity actuary, or even a small team. Some companies have taken this a step further: Legal and General plc created a "Longevity Science Advisory Panel" staffed with external experts to rigourously assess claims made for models in the longevity sphere.

However, not every company has the resources to set up its own advisory body with such deep specialist expertise. Instead, actuaries in other companies will need their own version of Sagan's "baloney detection kit" (Sagan, 1997) when dealing with claims made for a model. Sagan advocated that "everyone should have the essential tools to effectively and constructively evaluate claims to knowledge". So what should these tools be for actuaries examining projection models? Perhaps the primary tool should be the willingness to ask questions. For example, if a projection model has a category for yet-to-be-invented technologies to "reverse ageing", how can such a model be parameterised in a meaningful way? Oliver & Chung's paper in this special section reminds us how uncertain we can be of the benefit of a particular drug or technology undergoing a medical trial, while Goldacre (2012) shows how alarmingly distorted some publications can be. What reliance — if any — should be placed on models claiming to allow for technologies not in such trials, or not yet even invented?

The second tool should perhaps be the willingness to challenge experts, regardless of their qualifications or status. For example, some projection models contain in-built limits to the extent of mortality reduction. What is the basis for such an assumption? Why should these limits be any different from previous failed claims? In a short and highly accessible article, Oeppen and Vaupel (2002) showed how “experts have repeatedly asserted that life expectancy is approaching a ceiling: these experts have repeatedly been proven wrong”.

A third tool is to challenge assumptions which look like they are made for convenience, rather than being justified by evidence. For example, if a projection model has various sub-categories for mortality drivers, how are the inevitable correlations between the sub-categories dealt with? A common assumption is to assume independence because it makes the calculations easier, but how sound is that assumption? And what are the consequences of it being invalid? The chapters entitled “Why clever people believe stupid things” and “Bad stats” in Goldacre (2008) are a useful reminder of how easy it is to fall into lazy ways of thinking.

A fourth tool is to rely on models which have been openly published, ideally in peer-reviewed journals, and where independent researchers have been able to scrutinise proposed models. As Sagan (1997) says, “claims that cannot be tested [...] are veridically worthless”. For example, Richards (2008) presented results on the use of postcode-based geodemographic profiles in modelling annuitant mortality. On its own this was interesting, but the usefulness of postcodes in mortality modelling was questioned in some quarters. However, argument ceased when Madrigal et al. (2009) confirmed these results using different data and different methods.

Openness is not about simply asserting transparency; it means publishing enough detail of a method to enable others to verify the results, and ideally build on them. The best-known projection model, that of Lee and Carter (1992), has been extensively studied and its limitations are well understood as a result. For example, the Lee-Carter model cannot ordinarily reproduce cohort mortality patterns, so Renshaw and Haberman (2006) extended it for this purpose. Similarly, the Lee-Carter model can produce long-term forecasts of mortality rates which cross over at adjacent ages, an issue which is minimised by the variants proposed by Delwarde et al. (2007) and Currie (2013). The Lee-Carter model is useful despite its limitations, because these limitations have been thoroughly researched. Scrutiny, criticism and improvement can only take place with sufficient openness, however. In contrast, to find out how some proprietary models work you have to sign a non-disclosure agreement (NDA). It is hard to see a place for such models in a world where assumptions must be carefully documented and parameter settings validated.

There are subtle differences between academics and actuaries, however, and to some extent these differences parallel those between academics and medics. For example, a medical trial will be stopped if it appears that a treatment is harming the patients. Equally, a medical trial will also be stopped if the treatment is proving to be particularly effective, so that those in the control group can benefit. In each case, the duty of patient care overrides the duty to complete the trial as planned. In a similar vein, an actuary has a professional — and often statutory — duty towards prudence, and this can override other considerations. For example, there is a lively debate as to the extent, nature and even existence of cohort effects in mortality in the UK. An academic is free to argue the case against the existence of cohort effects if he wishes — there is no harm in doing so, and indeed opposing cases must be argued if knowledge is to advance. An actuary, however, is professionally bound to acknowledge the possible existence of cohort effects in reserving if it is the prudent thing to do, even if the debate is still ongoing. Analogous to the medic's overriding duty of patient care, an actuary must follow the precautionary principle.

Perhaps the last tool in the box is to remind ourselves of another of Sagan's missives: "the more we want [something] to be true, the more careful we have to be". This has particular resonance for actuaries, because prudence has a price. Somebody is paying for the capital tied up in valuation margins, and actuaries are under perpetual pressure to release them. We must be constantly on our guard not to fall for a model just because it promises what our paymasters want, be that a weaker mortality-improvement basis or a false sense of comfort about the uncertainty over projections. As well as a duty of prudence, actuaries also have a duty to serve the public trust.

References

Currie, I. D. (2013). Smoothing constrained generalized linear models with an application to the Lee-Carter model. *Statistical Modelling to appear*.

Delwarde, A., M. Denuit, and P. H. C. Eilers (2007). Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting: a penalized likelihood approach. *Statistical Modelling 7*, 29–48.

Goldacre, B. (2008). *Bad Science*. Harper.

Goldacre, B. (2012). *Bad Pharma*. Fourth Estate.

Lee, R. D. and L. Carter (1992). Modeling and forecasting US mortality. *Journal of the American Statistical Association 87*, 659–671.

Madrigal, A., F. Matthews, D. Patel, A. Gaches, and S. Baxter (2009). What longevity predictors should be allowed for when valuing pension scheme liabilities? *British Actuarial Journal 16(I)*, 1–62 (with discussion).

Oeppen, J. and J. W. Vaupel (2002). Broken limits to life expectancy. *Science 296*, 1030–1031.

Renshaw, A. E. and S. Haberman (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics 38*, 556–570.

Richards, S. J. (2008). Applying survival models to pensioner mortality data. *British Actuarial Journal 14(II)*, 257–326 (with discussion).

Sagan, C. (1997). *The demon-haunted world: Science as a candle in the dark*. Headline.