A Value-at-Risk framework for longevity trend risk

Abstract of the London discussion

[Institute and Faculty of Actuaries, 26 November 2012]

Contact

Stephen Richards, F.F.A., E-mail: stephen@longevitas.co.uk

This abstract relates to the following paper: Richards S. J., Currie, I. D. and Ritchie, G. P. A Value-at-Risk framework for longevity trend risk. *British Actuarial Journal*, doi:10.1017/ S1357321712000451

Mr P. G. Scott, F.I.A. (Chairman): I will invite Stephen Richards, co-author of the paper for discussion this evening, to introduce the paper.

(The full text of Dr Richards' introduction is given in the report of the Edinburgh session held on 19 November 2012.)

Mr S. McDonald, F.I.A. (opening the discussion): I would like to thank the authors for producing and presenting an interesting, thought-provoking and topical paper.

The paper gives an overview of two existing ways of considering longevity risk, or at least its trend component, before going on to set out a new practical approach for assessing longevity trend risk in a manner consistent with the requirements of a one-year value-at-risk based regulatory regime such as proposed under Solvency II.

The different approaches are illustrated by providing some results for particular model parameterisations. However, the main proposal in the paper is the framework itself rather than any particular capital numbers. A benefit of the proposed framework is that it can be used with a number of stochastic models, rather than requiring concurrent commitment to one particular model.

An additional application of the framework is also proposed – it is suggested that it can be used as a robustness test in the selection of a projection model. That this test can be conducted for simulations from other models as well as one's own makes for a strong test and I would agree that it could be a useful addition to the tests described in the Cairns, Blake and Dowd papers.

The authors make no claim that a one-year view is an especially natural way of assessing a long-term risk such as longevity. The proposed framework answers a specific regulatory question particularly relevant for pillar one of Solvency II. I can envisage that under such a regime, an insurer's management view might be informed by both an assessment using a one-year view such as this, and an assessment of the liabilities under run-off. The latter might be based on the same stochastic projection models, models with a causal driver, or both.

157

In common with a number of related papers, the data used in the paper is population data for England and Wales, provided by the Office of National Statistics. This data set has been much discussed, so I will make only some brief observations.

Whilst the death counts can be considered quite reliable, the exposures are estimates. We have had a timely reminder this year of the potential consequences of reliance on these estimates when it became apparent that a material restatement of the older age population estimates for the past decade is very likely. Richard Willets' articles from September describe the issue and potential consequences well. More recently still, Barnet Waddingham have suggested an alternative approach by which old age populations can be estimated, which would imply a larger departure again from the 2001 census. However the data is updated should not affect the validity or otherwise of the method, but will mean changes to the numerical results, and the apparent relative strength of the different models, because of the way that different models project the future from the past data.

My second observation as regards data surrounds the inclusion of the years prior to 1971. As noted by several authors on the subject, there are particular issues with exposure estimates at the oldest ages before 1971. Some models are particularly sensitive to whether this data is included, for example moving the data start date back from 1971 to 1961 in one Lee-Carter model increases the capital requirement for a 70-year-old male by more than 50%.

The authors truncate the data at age 105, and provide a table which seems to indicate that the simplification is not material. It might be helpful here to indicate whether or not the difference in annuity factors becomes more material under stressed improvement assumptions, particularly since longevity stress is defined as the ratio of two annuities, so it would only require a small change in the stressed annuity to increase the liabilities by around 0.1%, which might be considered material in a valuation. Also, if working with a smaller data set, perhaps the commonly used data set to age 90, the method might introduce additional approximation error.

Finally on data, a reminder that use of population data introduces a degree of basis risk for an annuity portfolio or pension scheme. A recent sessional paper gave some insight into the extent to which recent mortality improvements have differed among subsets of the population. An actuary looking to use any method based on population data in practice will need to determine whether and how to quantify and manage basis risk.

Section 3 of the paper gives a good example of a classification system for the components of longevity risk, which is helpful. Section 3.3 reminds us that the paper only sets out to address the longevity trend risk component, although some of the results in the paper are potentially helpful in informing our view on model risk.

Table 3 presents an illustration of the results for a single life male annuitant aged 70 in run-off, using a selection of stochastic models in fairly common use. It is important not to take false comfort from too brief a look at the results in this table. The capital requirement – that is, the difference between the 99.5th percentile of a model and its own central estimate – only varies between 4.3% and 6.3%. However, we can see that the central estimate of the 2D Age-Period model is actually outside of the 99.5% confidence envelope of the Cairns-Blake-Dowd (CBD) model. This is a clear indication of the materiality of model risk within longevity projections.

158

Figure 2 expands on the single age results of table 3 with an illustration of the way in which capital requirements vary by age for the four models shown. I think this is helpful, and guards against inappropriate quick conclusions about the models' relative strength that might not hold over the whole age range. I was surprised that the authors did not include a similar graph for their main value-at-risk results (in table 5) – as the patterns by age with that approach can be at least as interesting.

The discussion in paragraph 4.5 applies equally to the stressed trend and value-at-risk approaches. As the authors remind us, the numerical results are for a specimen age only and I would add that they also apply only to single life, male annuitants. Considerable care would be needed if looking to assess capital requirements directly using the specimen age approach described. I would agree with the authors that portfolio valuations are preferable in practice.

In paragraph 4.6 we are reminded that capital requirements under all approaches discussed depend critically on the level and shape of the yield curve. However, for the reasons just discussed, a practitioner might be more likely to extend the value-at-risk approach by undertaking a portfolio run, using the mortality improvements from the 99.5th percentile simulation. In that case what is more relevant is whether small changes in the level or shape of the yield curve lead to different ranking of the simulations. When I looked briefly at this point I found the method to be stable in this regard under small changes to the yield curve.

Paragraph 4.7 gives us a useful example of the difficulties of demonstrating the equivalence of any run-off based approach with a requirement, such as that in Solvency II, to measure risk over a one-year time horizon. More sophisticated approaches are possible, but still require work to be done in order to demonstrate equivalence with the requirements.

Section 5 considers the shock approach to longevity risk, as proposed in the Solvency II standard formula. The authors cite Börger's (2010) paper, which sets out well the shortcomings of the shock approach, and supplement it with figure 4, a powerful illustration of the inappropriate shape of the SCR requirement. Note that the dramatic increase in capital requirement by age would lead to further issues with the risk margin calculation for a portfolio where the average age is increasing over time. My own view is that the comment that "In theory [across the portfolio, the inappropriate shape] might not matter" is generous – this would imply a significant cross-subsidy by age, which could unwind as the portfolio mix changed.

The mechanics of the value-at-risk approach are set out in section 6, with a more technical expansion detailing the generation of the sample paths in section 8. The key question that an actuary will be looking to answer as regards the method is whether it adequately deals with both components of the one-year risk. These components are the risk that next year's mortality will be lighter than expected so that future projections start from a lower base, and the risk that our expectation of mortality for future years changes during the year, for example due to a medical innovation not yet reflected in the data.

I think it will be clear to most readers how the framework sets out to handle the first risk – including as it does both adverse mortality trend and trend volatility – such as caused by an especially benign external environment. Binomial variation in the number of deaths is added, reflecting that the number of deaths observed in the population may not correspond to the theoretical mortality rate. This point will be more relevant when the exposures are small, for example at older ages with the population data used here.

The point more likely to be subject to challenge is whether the method fully reflects the second risk – a new innovation such as progress towards treating a major cancer. This is a challenge frequently levelled at any pure projection-based approach.

The method as described, using all-cause mortality rates, does not consider specific causes of mortality improvement. Indeed the purpose of such a stochastic projection model is to project a general trend, without focusing on the causal drivers (that is, acknowledging that we can't know now what will be driving improvements in 20 years' time).

In other words, the method attempts to model the impact of future medical innovations and changes in behaviour via the projection of past rates of improvements, which were driven by events in the past. So it is intended to do more than just capture the effect of past innovations and behavioural changes diffusing through the population.

The question may then simplify to a consideration of whether the data incorporates changes as significant as those we consider plausible during the period in which projected liabilities will be financially material. Individual actuaries will reach their own conclusions here, but I would note that someone aged in their 60s or older in 1961, the start point of the data, was born in the 19th Century. So in considering the innovations feeding in to the mortality improvements in the data set I would potentially include:

- The development of vaccines and the introduction of antibiotics such as penicillin after World War Two
- The introduction of the National Health Service, and the more recent introduction of NHS screening programmes
- And the rapid decline in smoking after 1970.

Section 7 contains some detail on models that might be suitable for use within the framework. There is brief mention here that some of the models will be sensitive to the time period and/or age range chosen. Decisions such as these are key parameterisation choices; and the paper might benefit from some sensitivity analysis. This would likely demonstrate that often model parameterisation is as important as choice of model.

Table 5 in section 9 presents the value-at-risk results, again for a single life male annuitant aged 70. As before, we can see that the central estimate of some models is outside of the 99.5% confidence envelope of others, reiterating the model risk point. Indeed, there is a 15% difference between the central estimate of the CBD model and the 99.5th percentile of the 2D Age-Period model.

Given these differences, it might be best practice to consider more than one model, either as part of the core internal model or for validation.

Note that, unlike the stressed trend where the percentiles were determined mathematically from the fitted model, here they are estimated from 1,000 simulations. One approach to dealing with this issue using Harrell-Davis estimates is discussed in paragraph 9.3 and illustrated in figure 6. Another would be to increase the number of simulations. I would suggest that this is an area that a practitioner needs to consider, as I have seen differences in excess of 10% across all ages between SCR estimates from 1,000 samples and those based on larger samples.

160

An issue arises in practice if there is material divergence between the average value, or central projection, from a model and an actuary's best estimate view of future mortality improvement. Solvency II requires consistency between the approach used in calculating the probability distribution and the methods used in valuing the liabilities – so this point may merit consideration. This issue applies equally to the stressed-trend approach.

Since the method deals with each gender separately then consideration must be given to the way in which the results for each gender are combined. To make direct use of the 99.5^{th} percentile outcomes may assume that male and female mortality results are perfectly correlated. Historically, this has certainly not been the case as the underlying causes of improvement have affected the genders at different times. Arguably, though, the correlation might increase somewhat in the tails of the distributions.

As regards the paper's conclusion, I would make two comments. The first is about the observation that the framework tends to produce lower capital requirements than the stressed-trend approach. I am not sure that this follows directly from the information provided in the paper since there is no like-for-like comparison. This is because both table 3 and table 5 present the 99.5^{th} percentile but, as noted, it is likely that a different percentile would be considered under the stressed-trend approach.

Secondly, I am not sure that the one-year value-at-risk results in the paper lead naturally to the proposal of 3.5% as a floor for longevity trend risk. As has been noted, these results apply only to a small sample for particular model parameterisations at a particular age. To my mind, more results would be needed if there is a desire to present a numerical conclusion in addition to the valuable framework the authors propose.

Mr M. G. White, F.I.A.: This is quite a simple point. The new European vision of Solvency II sets out a "look one year ahead and do not think further" approach to calibration as well as some interesting assumptions on the assets. In part, this is to avoid recognising higher capital requirements than capital levels currently held.

But as the authors clearly believe, a one year view does not capture properly the presently committed risk, a particularly dramatic problem for annuity business, but not trivial for long tail non-life, either. Is there not a danger that the regulatory approach being prescribed in this way will become the way that companies start to think and to communicate to their shareholders? If this does happen without bodies such as ours speaking out to say how silly it might be, it may lead to a general under-appreciation of these risks rather in the way that inappropriate models with inappropriate parameters helped to achieve the banking mess.

Prof. A. D. Wilkie, F.F.A. F.I.A.: I made a number of points at the meeting in Edinburgh last week, but there are a few other, smaller points that I should like to make this evening.

Last week there was mention – and Dr Richards said it again this evening – that the different models respond differently to an extra year's data and are more or less stable in their responses.

One way of looking at this would be to go back to 1992, say, 10 years before the latest data, and add a year at a time within each model and see what the variability in the future forecast is, starting at that time.

The authors do not mention much about the difference between the central point and the scatter of the one-year forecasts of mortality. The different models are starting from slightly different places in their central assumptions, but over one year their central assumptions and the scatter may vary. It would be quite interesting (if laborious) to forecast one year from each model and then apply each other model to the updated history as generated by each model – 36 different calculations rather than just six.

In reality, the next year's data does not come from any of the models at all and is just the same for all models by the time you reach the end of the year. But, starting at this point, it might be interesting to see whether some of the changes come as a result of variability in the one-year forecast or variability in response of the model to that.

A further point I would like to make is this. In paragraph 8.5 the authors mention that for certain models that do not produce automatically the volatility from the past, you can use the past residuals.

You choose one of the years that they have used from 1961 to 2010 at random, look at the variability, age by age, and apply those residuals to a new forecast, using an indicator variable which they say can either be zero or one, so you can either use this method or not use this method.

I have never been happy about using just the residuals or the experience from the past if it means the most extreme future event is going to be the same as one of the past events in both directions. You have no chance of going worse than the worst in the past or better than the best in the past.

Just a suggestion: one could vary this indicator variable to see what would be the result by giving it a value of, say, 1.5 or 2 or 0.5. So, for all the variability you have from the past, you take some bigger or smaller fraction of the values, so as to at least give the opportunity of some more extreme value occurring than happened to have occurred over the 40 years of your sample.

Dr I. D. Currie: Perhaps while the audience is thinking of some points, I could respond to one of the points made by our opener. There was a discussion of how the improvements that we have observed in mortality over the last 20 or 30 years have come about.

Our opener mentioned vaccines, the introduction of the NHS, and the dramatic changes in smoking habits, which have all had an obvious and positive impact on the course of mortality. There was an implication that these kinds of improvements might not happen in the future. However, the improvements that we have seen in the past, despite the fact that we have had these substantial positive shocks, have been remarkably steady. There is no indication over the last 20 or 30 years that they are tailing off.

I should also like to mention a famous paper that was published in Science in 2002 by Oeppen & Vaupel as it addresses this particular point directly. They plotted the maximum female life expectancy, or record life expectancy, over all the countries of the world. They did this from 1840 to the year 2000. In 1840 the record female life expectancy was just 45 years. It was held by Swedish women. In 2000 the record was 85 years. That was held by Japanese women. Incidentally, since the paper was published it has gone up by another two years – still Japanese.

What is quite extraordinary is not so much this amazing 40 year improvement taking place over those 160 years, but the plot of the maximum life expectancy is chillingly linear.

162

I quote from the paper: "The linear climb of record life expectancy suggests that reductions in mortality should not be seen as a disconnected sequence of unrepeatable revolutions but rather as a regular stream of continuing progress." This conclusion responds directly to our opener's point.

Finally, I should like to say that the authors of this evening's paper completely agree with Oeppen and Vaupel's conclusion.

Mr White: This is a trivial response to the last point but, given it is so obvious, how did we miss it? That is the question the outside world has asked.

The Chairman: It seems to me that the authors have done a great job addressing the question, which was, to use my words: how do you use a value-at-risk framework for assessing the capital requirements for longevity risk?

The question that was not asked was: should you use a measure of this type for assessing capital adequacy? Personally, I remain quite unconvinced, given the complexity of the stochastic modelling and the nature that one year is not what you are worrying about, but the potential significant movements into the future.

The only thing that you can draw from history over the last 20 years is that insurers have consistently under-estimated the amount of extra capital necessary to allow for improvements in mortality.

Do the authors have any views on the subject? If you were a regulator, would you be using a model of this type for assessing regulatory capital?

Dr Richards: Longevity trend risk is, I think, an example of one that should not be viewed in that particular way. If I were a regulator, I would probably want to see an insurer's attempt at a one-year value-at-risk for trend risk. But I would not be unhappy if an insurer stated clearly that this was not the best way to view this risk and I would not penalise an insurer for clearly stating that this risk does not fit this framework. The underlying objective is to manage the business properly, not to just fit particular risks into a particular regulatory framework.

Dr Currie: The only point I would add is that the value-at-risk approach and the stressed-trend approach both seem to require a stochastic approach. It does not seem possible to answer this question of a 99.5% extreme event in anything other than a stochastic framework.

Prof Wilkie: I want to pick up on Dr Currie's point. You have these linear trends in various other circumstances, too. One needs to look more closely at the detail. Rather than use the expectation of life, another useful statistic is the peak age at death. That is, taking the derivative of the l_x table, the d_x column, and seeing where in adult life it reaches a maximum. According to ELT 1, in 1841 in England and Wales the peak age was 72. It has gone up now to something like 89, which is not nearly as big an improvement as from 45 to 85.

Plenty of people in the far distant past are known to have lived into their 80s and 90s. Occasionally, people reached 100. The big change in Britain and elsewhere from the 1840s onwards was a huge reduction in mortality at young ages. The population mortality in 2000 at ages 5 to 14 is about $\frac{1}{2}$% of what it was in 1900.

163

So, far more people are reaching the biblical "three score years and ten", which must have been thought of as a decent age at one time.

That is quite different from many more people going on from the present older ages of, say, 70 upwards up to even older ages of 130 upwards. If you have an increase from 45 to 85, which is an increase of 40 years, over 170 years, then just projecting it over the next 170, taking us to 2180 – which probably none of us will be alive to see, but you are never quite sure with this improvement – it does seem implausible that the average age at death will reach 125.

Recent generations do not die in their 20s from consumption, accidents or diseases, but live until their 70s, 80s and 90s.

It was not mentioned, but there were also various improvements in the treatment of heart disease. But they will not keep people going forever. So, I am sceptical about taking long-term trends with a straight line view.

Another example is some figures I have on average earnings and prices in Britain. In the past 200 years since 1809, average earnings have increased by about 2% more than prices with a straight line on a logarithmic scale. Populations may have increased in straight lines on linear or logarithmic scales, but there must be some tailing off some time. You cannot have a continually increasing population in the world forever because ultimately there is not room. It is unlikely that real GNP per head would increase forever. It is more likely to stabilise some time. Similarly, longevity is likely to stabilise at advanced ages some time. It will not go on improving forever. That seems totally implausible.

The Chairman: Any more thoughts from the floor, please? Is everyone keen to have an average age at death of 125?

Mr T. W. Hewitson, F.F.A.: One of the tests that I think regulators would be quite interested in is the so-called back testing. Did the authors look at the possibility of applying the models using only some of the earlier data, and then testing the results? Albeit there will be less data available then and the results might statistically not be reliable.

My sense on this would be that you might well find most of the actual deviations, as compared to the model, moving in one direction. It does make me feel slightly uncomfortable if some of these models seem to be constantly erring one way rather than the other.

I also feel a bit nervous about some of these models and their instabilities. Small changes in the data are causing big changes in the results. Clearly, having several different models can help.

Coming back to the points made earlier, I think that trying to look behind the data at some of the potential causes for changes in longevity must be a sensible approach; that is, the changes that have been noted in the past: the reductions in smoking, the various antibiotics, and treatments of diseases and other conditions such as cancer. I am thinking about whether and to what extent some of those could be repeated in the future. That would be a good sense check.

Mr G. Ritchie (author): Prof Wilkie posed the question about whether our value-at-risk simulations could be driven by population forecasts from a different model.

164

The baseline model within the framework provides the sample paths that drive the populations underlying the simulations, but there is absolutely no reason why the simulation model needs to be the same as the baseline model.

In fact, it is a good robustness check. You might want to make a more volatile or responsive model generate those population forecasts just to see how your simulation model behaves in respect of those.

Mr G. L. Jones, F.I.A.: The authors identify many of the challenges we collectively face in setting appropriate capital levels for longevity, and ultimately they focus on one approach to set a floor on that capital within a specific regulatory context.

In an era of increasing regulatory intervention, it is important that we as a profession are active, and are seen to be active, in setting our own boundaries in a robust and transparent way, and potentially providing challenge to regulatory intervention where that may not be appropriate.

We have rehearsed from a number of viewpoints the challenges that stochastic modelling of long-dated mortality trends within a one-year view does not capture the underlying risks of the business, which is not so much a new data point as more widely new information.

That challenge is borne out by my colleague, the opener, and Dr Richards identifying that the key risk is probably the difference between two different models rather than within a single model. In general, those are the numbers which I find most interesting in papers that look at models: the inter-model, rather than the intra-model, risk.

I like the suggestion that one criterion for model selection should be stability under a sample path introduced from another model, but that does imply a prior view on what appropriate model behaviour might be to new data, and ultimately you must expect that models should behave in an unpredictable manner in the presence of new and potentially unpredictable data.

So to that end, I would certainly consider supplementing the results that are obtained from a single model with some additional intervention to recognise the model risks that are present.

As an aside, the model approach of adding one year's actual data, where it is impossible to disaggregate the components of volatility and trend, is a challenge with high age mortality data where I see high levels of anti-correlation between mortality rates and monthly average temperature. I question whether the regulator's question is really what level of capital you simply need to hold at the end of the year for a single model, but rather whether the intent is to capture the level of the capital that you need to hold to exit that line of business after a stress event.

As we have seen through the financial crisis, the most likely conclusion, following a capital event, is not simply that you have to recalibrate your models, but that your model was fundamentally wrong. So another way of framing the question of model risk is at what point you would discard the model that you are currently using. Historically, we have faced a couple of events of this nature. We are going through it in respect of longevity, recognising the accelerated pace of change that we have observed post-1992, and that is amply borne out by the data.

A previous challenge for mortality trend was 1984–85, following widespread awareness of what we now know as HIV. Prices within the market doubled over the year but there was minimal impact in

terms of explicit data to calibrate that. It also illustrates the capacity of markets and – I think it is fair to say – actuaries to overshoot sometimes following extreme events.

One of the fundamental challenges of the modelling is to reflect on what the nature of the underlying drivers are. I doubt any of us would identify a specific model as particularly strong at projecting mortality. To my mind, it is unclear whether we are seeing what should be regarded as a stochastic process. Perhaps what we are seeing is mortality as a chaotic response to a far larger system.

Ultimately, mortality and longevity lie at the heart of our existence as a profession. The paper has provided a welcome development in our consideration of the risks, but I suspect that as long as longevity continues to be a material risk, there will be a debate about an appropriate means of quantifying the uncertainty.

Dr Richards: I will take the opportunity to answer some of the points that have come up so far. To begin with, the most recent question from Dr Jones was to what extent volatility or trend uncertainty is contributing to the one-year value-at-risk approach. These are separately switchable. Section 8 of the paper describes how you can include either the volatility, or the trend risk, or both. Using the framework you are able to switch the different components on and off and see what impact they have on the value-at-risk capital.

Mr White raised a question in terms of the recent financial crisis: why did we not see what was coming in terms of changes in mortality risk? This touches on something which was raised in Edinburgh, namely the difference between this mechanistic one-year, value-at-risk approach and how a life office might operate in practice; in particular, there is a difference between recalibrating stochastic projection models and the real-world events that happen to life offices, and whether or not that squares with the capital requirements coming out of the value-at-risk approach.

After the Edinburgh meeting, I looked for a real-world example of this. I found one which is both recent and directly relevant: the 1999 paper by Richard Willets, "Mortality in the Next Millennium", in which he brought the cohort effect to the attention of the actuarial profession. This is a real-world example of a relatively sudden change in actuarial expectations as to the direction of future mortality improvements.

I did some calculations of annuity factors at age 70. If we take the 92 series mortality table, we can calculate annuity factors first using the CMR 17 projections and then the short- and medium- cohort projections as an example of the sort of real-world event that has an impact on actuarial work. The switch from CMR 17 to the short cohort resulted in a 3.6% increase in the annuity factor, whereas a switch to the more commonly used medium cohort caused a 4.2% increase. These increases are consistent with the capital requirements demonstrated in the value-at-risk process in the paper.

A few years after publication of Willets (1999), the change in expectation of mortality improvements caused one listed life office to have to make a stock exchange announcement because its annuity reserves had gone up significantly.

This example is recent, relevant and has nothing to do with stochastic models or the value-at-risk framework. Nevertheless, the capital increases seen are in line with the capital requirements being suggested by the value-at-risk approach.

166

The question we had in Edinburgh was whether or not the value-at-risk results are in any sense realistic. I think here we have an example which shows that the capital requirements produced here dovetail with what we have witnessed in the recent past. Of course, modern economic conditions are very different to what they were a decade ago. We know that the trend-risk capital requirements are linked to the shape of the yield curve, and the yield curve now has a different shape to what it had 10 years ago. However, the switch from CMR 17 to the short-, medium- and long-cohort projections produced reserve increases in line with the capital requirements from the value-at-risk approach.

This brings me to a question about models, in particular whether they are perhaps not sensitive enough to changes in the data, or are over-sensitive. A key point here is again to use multiple models to investigate their different responses. Dr Jones correctly assumed that the authors will not state which model is the best. This is because none of them is uniformly better than all of the others with respect to the kind of behaviours or tests to which you might want to subject these models. But they do share one key feature, namely they are all peer-reviewed and published in academic journals. Their respective flaws and drawbacks are therefore well known. For us it is not only important to use a number of different models, but to use peer-reviewed models which have been openly published for critical evaluation.

One other subject was limits to mortality improvements. Oeppen & Vaupel (2002) is a short and accessible paper, and it warns of the dangers that experts experience when trying to call a limit to total mortality improvements. Their more or less final words were: "Experts have repeatedly asserted that life expectancy is approaching a ceiling. These experts have repeatedly been proven wrong." This is a warning to any analyst who wants to assume that the worst of the mortality improvements are behind us. On this topic, Booth & Tickle (2008) reviewed the different approaches that might be used to project mortality, and they addressed the specific topic of expert opinion. They wrote that the advantage of expert opinion is "the incorporation of demographic, epidemiological and other relevant knowledge, at least in a qualitative way. The disadvantage is its subjectivity and its potential for bias. The conservativeness of expert opinion with respect to mortality decline is widespread in that experts have generally been unwilling to envisage the long-term continuation of trends, often based on beliefs about limits to life expectancy."

This is a cautionary note. We have seen already that the actuarial profession missed the cohort-based mortality improvements until Richard Willets brought them to our attention. As a general principle, we should be wary of any attempt to state that the strongest mortality improvements lie only in the past. Recent history has shown that this is a strong assumption to make, and not necessarily a sensible one to be making when setting capital requirements.

The Chairman: Thank you very much, Dr Richards. May I thank all of the authors for all the work that has been done? The combination of the paper and the discussion has been limited in quantity, but the quality has been high. The discussion has brought out some of the strengths and weaknesses of using these models for assessments of something that is fundamentally very difficult to assess.

During my career as an actuary, I have always tried to hold on to the view that models are useful for illustrating outcomes based on the assumptions that the models are using. What they can never do is make management decisions. They can inform management decisions – about a particular level of reserve, a particular price or a particular level of capital – but models cannot make decisions. This discussion has reminded me of that view.

So, on your behalf, may I thank the authors, the opener, the closer and those who participated this evening?