# Longevity Risk and Fair Value Accounting

Peter M. Mazonas, CPA

The technical parts of this chapter are extracted from testimony given by this author on November 2, 2009, before the Securities and Exchange Commission's Life Settlement Task Force. The original copywrited documents are available at https://www.LSF-llc.com

Reverse mortgages and life settlements are sister asset classes. Both reverse mortgages and life settlements are longevity-dependent assets where the value and thus initial pricing is dependent upon an unobservable future event – the death of the borrower or the insured. This chapter investigates and demonstrates a better way of predicting life expectancy, "moveout", for an individual reverse mortgage loan. The methodology described can easily be built into the loan origination process with significant advantages to both senior borrowers, while affording investors and the American taxpayers more accurate individual loan pricing, portfolio valuation and transparency.

There has been little incentive to reevaluate and then alter the pricing components of the existing Home Equity Conversion Mortgage (HECM) reverse mortgage. These Federal Housing Administration (FHA) loans are offered by mortgage brokers, backed by a Ginnie Mae guarantee and securitized by large institutions.

Reexamination of reverse mortgage moveout determination is important today as regulators, rating agencies and the accounting profession meet the requirements of fair value accounting. Every securitized portfolio of reverse mortgages that is either issued as a security or where investors in the portfolio issue audited financial statements will have to annually comply with fair value accounting. Specific methodologies to revalue each loan in the portfolio and mark the portfolio to market are defined by the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB).

Basing moveout on values derived from large population mortality tables that have been known for thirty years to be flawed[1] creates significant liability for bankers selling the securities, auditors and actuaries who attest to portfolios valuation, and the lawyers structuring the transactions.

Proper reverse mortgage design requires an accurate understanding and prediction of the three main variables used to price each loan:

---

[1] *This was first described in* Vaupel, J. W., K. G. Manton, and E. Stallard. "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16(3): 439–454, 1979. *Further evidence was provided in* Iachine, I. A., N. V. Holm, J. R. Harris, A. Z. Begun, M. K. Iachina, M. Laitinen, J. Kaprio, and A. I. Yashin. 1998. "How Heritable Is Individual Susceptibility to Death? The Results of an Analysis of Survival Data on Danish, Swedish and Finnish Twins." *Twin Research* 1: 196–205.

1. Appraised home value at origination,
2. Interest rate risk, especially for variable rate loans, and
3. An accurate assessment of loan term based upon the non-institutionalized life expectancy of the borrower, or the survivor in the case of a couple.

Historically the estimated loan term, typically called “moveout”, for a reverse mortgage has been determined using large population mortality tables, typically the Social Security Administration Mortality Tables. These are derived from the U.S. Decennial Life Tables (USDLT) published every ten years. By definition average life expectancy (LE) is that point in the future when fifty percent of a large population will have died and fifty percent will still be alive. The use of these tables simplified the determination of the term for each loan to a simple table lookup and added uniformity to product design.

For each loan to be accurately priced, and thus a portfolio to be accurately valued, two longevity factors are critical:

1. The applicability and accuracy of the methodology for determining morbidity (onset of sickness) and mortality (death), and
2. Determination of the future point in time when morbidity will force the senior to move out because of their need for institutional care, thus terminating the loan.

A number of problems exist using large population mortality tables to define the term of each loan. This easy solution has been a disservice to both senior borrowers and to investors who purchase interest in securitized pools of reverse mortgages. Using average life expectancy to predict the loan term means that half of the borrowers will live longer than the average. Someone or something has to pay for this risk if the product design is going to be valid and yield the expect rate of return. Early product designers hedged their bet and balanced the “who” and “what” that would compensate for error in product design and execution. Loan payments to borrowers were reduced so the “underpayments” to the fifty percent who died early or on time offset the extra payments made to the long-lived borrowers. The “what” was requiring borrowers to make an upfront mortgage insurance premium payment from their first loan draw a mandatory part of the popular HUD/FHA Home Equity Conversion Mortgage (HECM). It was anticipated that this would solve the longevity risk element of product design.

Controlling longevity risk in this way has now been shown to no longer work. The two percent (2%) mortgage insurance premium charged at origination on each loan and added to a risk premium pool has proven to be inadequate to cover the combination of longevity risk and interest rate risk. For HECM loans, when the principal paid out, and the accumulated accrued interest equals the maximum loan amount, each loan is “put” back to FHA under the Ginnie Mae guarantee. To the extent losses exceed the value of the mortgage insurance pool the American taxpayers absorb the loss with more national debt.

This shortcoming or design flaw was anticipated by product designers including this author who designed and patented Transamerica HomeFirst's lifetime reverse mortgage, "HouseMoney". Using outside actuarial consultants, HomeFirst relied upon new data gathered from the National Long Term Care Survey (NLTCS). In 1984 the United States Government, as part of the new Medicare Program, began gather assessment data on more than 32,000 seniors between the ages of 65 and 69. More specific data was gathered by Cologne Reinsurance on seniors in assisted living and skilled nursing facilities across the country.

The Cologne Re data was interpreted by actuary Phillip Barckman. By 1989 Barckman was convinced there was a correlation between an individual's loss of activities of daily living (ADLs) and the need for institutional care. With some exceptions, most seniors would like to age in place in their home and familiar surroundings. Based upon observation of data, Barackman hypothesized that at the loss of two ADLs a senior would only be able to remain non-institutionalized with the help of a cohort. By the time they lost three ADLs, they would require institutional care or in-home health care. Look at the list below and image your condition having lost two and then three of these capabilities.

- Personal hygiene
- Dressing and undressing
- Eating
- Transferring from bed to chair, and back
- Voluntary bladder / bowel control
- Elimination
- Moving around (as opposed to being bedridden)

Approximately the first 1,000 lifetime reverse mortgages written by Transamerica HomeFirst from 1993 through 1994 based moveout on large population age-based statistical data that extrapolated when the borrower would have lost 2.5 ADLs. Moveout for each loan pricing was at that point. Actuaries from Ernst & Young, Transamerica's auditors, objected to this methodology and forced HomeFirst to conform to conventional wisdom and use Social Security mortality table data to estimate a borrower's moveout.

In 1980 and 1981 a design committee created an individual level assessment tool for the Social Security Administration that collected medical, activity of daily living (ADL), instrumental activity of daily living (IADL), cognitive, and other socio-demographic/behavioral variables. A subset of these questions was developed into a questionnaire assessment tool for the Center for Medicare and Medicaid Services (CMS). CMS later started requiring quarterly assessment of Medicare recipient residents in assisted living and skilled nursing facilities. One company that developed a computer-based version of this assessment tool similar to the questionnaire developed for the NLTCS was Vigilan Systems, in Wilsonville Oregon. Individual resident level data gathered using the Vigilan *Administrator* assessment tool has subsequently confirms Barckman's hypothesis of when seniors move from

their homes.[2] Vigilan data shows that the average resident stays in assisted living for 25 months, after which they move out because of increased morbidity or death.

Longevity valued assets (LVA), specifically reverse mortgages and life settlements, are today being mispriced because (1) the reference tables being used to predict the insured's life expectancy for pricing are population or sub-population based tables and (2) the multiplicative rating factors applied to those tables are biased. Understanding the origins and workings of these reference tables is important to understanding their flaws.

The force of mortality exhibits exponential growth above about age 20. This was discovered by Gompertz (1825) who published the famous "law of mortality" that now bears his name.[3] The Gompertz law can be written with three elements: (1) a continuous exponential function of age, (2) with a growth constant, and (3) a proportionality constant, (detail below).[4]

The Gompertz law has been evaluated in numerous studies.[5] Several problems have been identified. (1) The fitted Gompertz function tends to overestimate mortality above about 80–90 years of age. (2) The Gompertz "constants" are not really

---

[2] Under a HIPAA Business Associate Agreement, Life Settlement Financial, LLC, (LSF) has examined approximately 1,000 resident sets of assessment data spanning up to seven calendar years. This evaluation and data compilation was done primarily to revalidate LSF's Longevity Cost Calculator™ (LCC). Thirty-one of the questions in the Vigilan assessment tool directly map to the seventy-six questions in the LCC. A byproduct of this revaluation was the observation that seniors entering assisted living have typically lost between two and three ADLs. Statistically, the average stay by a senior in an assisted living community is twenty-five months. They either leave because of death or are moved to a skilled nursing facility where they pass.

[3] Gompertz, B. 1825. "On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies." Philosophical Transactions of the Royal Society of London 115: 513–583.

[4] The Gompertz law can be written with three elements: a continuous exponential function of age, $a$, with a growth constant $\beta$, and a proportionality constant $\alpha$.[4] The Gompertz law has been evaluated in numerous studies.[4] Several problems have been identified. (1) The fitted Gompertz function tends to overestimate mortality above about 80–90 years of age. (2) The Gompertz "constants" ($\alpha$ and $\beta$) are not really constant; they differ from one population to another and over calendar time and cohort within a given population.[4] (3) The numerical values of $\alpha$ and $\beta$ are negatively correlated when calculated for populations with a broad range of mortality conditions.[4] Despite these problems, the Gompertz function provides a useful approximation to the age-specific mortality probabilities, $q_a$, and an even better approximation to the age-specific mortality "hazard rates," $\mu_a$, defined as $\mu_a = -\ln(1 - q_a)$, which can be solved for $q_a$, using $\mu_a = \alpha \exp(\beta a)$. Note that $q_a$ is the probability that a person alive at exact age $a$ will die prior to exact age $a + 1$. Typical increases in $q_a$ are in the range of 8–10% per year.

[5] *See* Wetterstrand, W. H. 1981. "Parametric Models for Life Insurance Mortality Data: Gompertz's Law over Time." Transactions of the Society of Actuaries 33: 443–468; *and* Olshansky, S. J., and B. A. Carnes. 1997. "Ever since Gompertz." Demography 34(1): 1–15.

constant; they differ from one population to another and over calendar time and cohort within a given population.[6] (3) The numerical values of two of the constants are negatively correlated when calculated for populations with a broad range of mortality conditions.[7] Despite these problems, the Gompertz function provides a useful approximation to the age-specific mortality probabilities and an even better approximation to the age-specific mortality "hazard rates". The typical annual increase in age specific mortality is in the range of 8–10% per year. All Gompertz functions are proportional to each other. Using the rule of 72, this implies a doubling time of 7.2 – 9 years of individual mortality probability. If the doubling times for individual mortality probability are constant in the range of 7.2 – 9 years, then it is reasonable to infer the underlying function is Gompertz. If doubling time is not constant, then the mortality is non-Gompertz, (Figure 1 below). This means the exponential tables use a constant multiplier to increase the probability of death at increased age irrespective of the individual's personal characteristics.

The problem with the multiplicative model is that individual mortality does not appear to follow the Gompertz function, even approximately.[8] If mortality is not proportional to Gompertz, then the multiplicative model is incorrect and should not be relied upon. The Social Security, USDLT and other exponential large population mortality tables are Gompertz.

In 2007 the *North American Actuarial Journal*, after two years of peer review, published a paper by P.J. Eric Stallard, Associate Director, Center for Population Health and Aging, Duke University, entitled *Trajectories of Morbidity, Disability, and Mortality Among the U.S. Elderly Population: Evidence from the 1984-1999 NLTCS*.[9][10] Unlike medical records' only mortality-predictive tools based upon large population reference tables, Stallard's model describes health and the probability of death at the individual level without assuming any specific functional form for the increase of mortality over age or of the ratios of mortality between individuals and the population to which they belong, (e.g. non-Gompertz). Stallard is the 2006 recipient

---

[6] *See* Strehler, B. L., and A. S. Mildvan. 1960. "General Theory of Mortality and Aging." Science 132(3418): 14–21; *and* Gavrilov, L. A., and N. S. Gavrilova. 2001. "The Reliability Theory of Aging and Longevity." Journal of Theoretical Biology 213: 527–545.

[7] Ibid

[8] *This was first described in* Vaupel, J. W., K. G. Manton, and E. Stallard. "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16(3): 439–454, 1979. *Further evidence was provided in* Iachine, I. A., N. V. Holm, J. R. Harris, A. Z. Begun, M. K. Iachina, M. Laitinen, J. Kaprio, and A. I. Yashin. 1998. "How Heritable Is Individual Susceptibility to Death? The Results of an Analysis of Survival Data on Danish, Swedish and Finnish Twins." *Twin Research* 1: 196–205.

[9] Eric Stallard, *Trajectories of Morbidity, Disability, and Mortality Among the U.S. Elderly Population: Evidence from the 1984-1999 NLTCS*.

[10] The above paper was first peer reviewed before presentation at *The Living to 100 and Beyond Symposium* sponsored by the Society of Actuaries, January 12-14, 2005. The conference committee referred the paper for further peer review and consideration for publication in the *North American Actuarial Journal*.

of the Edward A. Lew award from the American Society of Actuaries for his work in longevity-specific actuarial modeling.

The results of that analysis were surprising and illuminating. The rates of increase in the mortality probabilities for individuals were both faster and slower than the average rate of increase in the study population, with the slowest rates of increase occurring at the highest mortality levels and the fastest rates at the lowest mortality levels. However, there was substantial heterogeneity in the rates of increase at the lowest mortality levels making it difficult to generalize the pattern of increase for individuals without considering their individual risk profiles.

Simplistically but true, unhealthy people die at a faster rate and healthy people die at a slower rate. Ask any nurse in a senior care community and he or she will tell you that much about the mortality curve of an individual is not contained in their medical records.

Life Settlement Financial owns the commercial rights to this model, now named Longevity Cost Calculator™ (LCC)[11]. The LCC model clearly shows that impairments in Activities of Daily Living (ADLs), Instrumental Activities of Daily Living (IADLs), range of motion and cognition are greater predictors of morbidity and mortality than are individual or groups of medical conditions or table values. The LCC is the only current predictive model to scientifically do this.

Unlike large population-based mortality tables, the LCC uses a questionnaire assessment model to evaluate the overall health of an individual and predict the degradation of their health. This allows the model to distinguish between when the senior is able to live in the community independently and when they must move out permanently to receive institutional care.

Using linear regression analysis the actual vs expected accuracy of the Longevity Cost Calculator™ model is validated at 96% or greater. This mortality accuracy also applies to the important prediction of how long an individual will remain living in the community before they should move out to an institutional setting. This means that the predicted year-over-year annual mortality has this degree of accuracy. It is interesting to note that when the above 95 predictor variables are Chi-square ranked, not until number 26 of 95 do you find a variable typically found on medical records.[12]

This model has been in use by the Centers for Medicare and Medicaid Services (CMS) and The National Council on the Aging for three years as the Long-Term Care Planning Tool on the Medicare web site. For applications of the tool, see http://www.medicare.gov/LTCPlanning/Include/DataSection/Questions/SearchCriteria.asp?version=default&browser=Firefox|3|WinXP&language=English&defaultstatus=0&pagelist=Home

---

[11] http://www.lifesettlementfinancial.com/the-longevity-cost-calculator.html

[12] Eric Stallard, NAAJ paper, Appendix, pages 49-50, NLTCS Variables, Log-Likelihood Values, and Chi-Squared Statistics by Variable, by Sex

For documentation of the methodology, see http://www.medicare.gov/LTCPlanning/Static/ProjectedCosts_BriefDesc.asp?dest=NAV|Results|Report|ProjCostsBriefDesc|ProjectedCosts#TabTop

The mortality model has since been independently implemented in the Longevity Cost Calculator and a working web based model is available at https://www.lifesettlementfinancial.com .

In addition to proper longevity valuation at the individual loan and portfolio levels, the size of pooled loan portfolios is also important and often misunderstood by investors. Assuming there has been consistent and proper loan underwriting, not until a portfolio contains at least 384 loans can one expect to have an accuracy of +/-10.0% at a 95% probability level.[13] Not until at least 1,537 loans can you expect an accuracy of +/- 5.0% at the same 95% probability level.[14] This means that at a minimum, a reverse mortgage loan portfolio must contain $100 million of loans. A more appropriate portfolio size would be in excess of $400 million of loan amount. This dollar size is the reason this author asserts that these portfolios, now or during their life, will be audited under GAAP and may be under the oversight of the Public Company Accounting Oversight Board (PCAOB) because issuers of audited financial statements are holding these portfolios as investments. At the PCAOB's Standing Advisory Group Meeting on October 14-15, 2009, staff called for a standards-setting project to revise its existing standards on auditing fair value measurement.[15]

## Longevity Cost Calculator™ as a Loan Underwriting and Pricing Tool

The completed Longevity Cost Calculator™ questionnaire, in addition to providing an LE and move out date, scores each insured using a four level Grade of Measurement (GoM) system. As per the earlier referenced material on the Life Settlement Financial web site, the interrelationship of an insured's ADLs, IADLs and possible cognitive impairment affect those GoM scores and the trajectory of the individual's survival curve used to price a settlement offer. To facilitate the description of the health changes, the model generates time-invariant GoM scores that characterize the predicted health status of each person at the time they are/were in the youngest age-group in the calibration dataset, which for the current implementation is age-group 65–69.

GoM 1 (also referred to as "Pure Type I" or "Type I", with Roman numerals designating the rank ordering of the states by health status; we use the shorter GoM 1–4 reference throughout this document) refers to the healthiest component of the

---

[13] L.H. Longley-Cook. *An Introduction to Credibility Theory*. Casualty Actuarial Society, 1962, p. 9.

[14] L.H. Longley-Cook. *An Introduction to Credibility Theory*. Casualty Actuarial Society, 1962, p. 9.

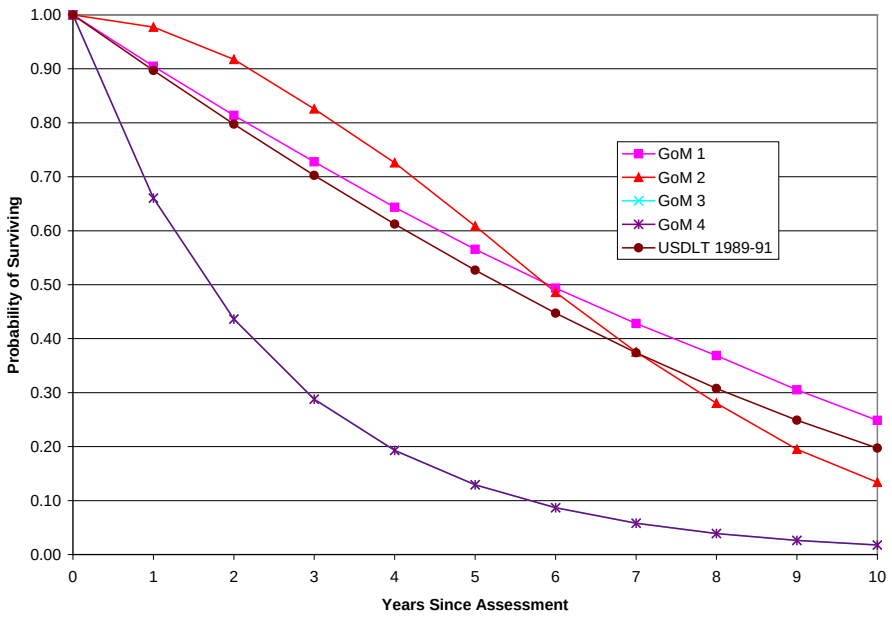
[15] PCAOB Standing Advisory Group Meeting, October 14-15,2009, Auditing Fair Value Measurements and Using the Work of a Specialist, paper developed by the staff of the Office of the Chief Auditor.

population. GoM scores 2–4 captures a range of health problems that occur at different ages, with progressive and graded transitions from GoM 2 to GoM 3 and 4. GoM 2 scores refer to people who have numerous medical problems, but few, if any ADL or other functional problems, or cognitive impairment. Persons with initial strong scores (i.e., close to 1.0, or 100%) on GoM 2 will live longer than traditional LE providers estimate; although this changes at older ages where these persons transition to strong scores on GoM 4. Persons with strong scores on GoM 3 have minor medical problems, but mild/moderate cognitive impairments, usually not indicated in their medical records, although this also changes at older ages where these persons transition to strong scores on GoM 4. Strong scores on GoM 4 identify people who have more serious medical problems, combined with serious ADL and/or cognitive impairments, usually not indicated in their medical records. People with initial strong scores on GoM 3 and GoM 4 will have shorter LE's than those issued by traditional LE underwriters.

Recall that the LE is the area under the relevant survival curve for the person or population for which the LE is being calculated. Thus, differences in LE between persons or groups of persons are best understood by examining the associated survival curves. This is illustrated in the following graph (Fig. 1) which plots the survival curves for males assessed at age 82 (i.e., age at last birthday is 82) for the next 10 years following the assessment, with a separate curve shown for each of the four time-invariant GoM pure types and also for comparison the survival curve from the U.S. Decennial Life Table (USDLT) for 1989-91.

**Figure 1**

**Probability of Surviving for Each of 10 Years Following Initial Assessment at Age 82, Males, by Initial Time-Invariant GoM Pure Type and for the U.S. Decennial Life Table for 1989-91**



According to the USDLT, the male LE at age 82 is 6.2 years. This value is less than the LEs of 6.9 and 6.3 years for GoM 1 and 2, but is substantially higher than the LEs of 2.5 years each for GoM 3 and 4 (the values are the same because an initial GoM 3 "converts" to GoM 4 prior to age 82).

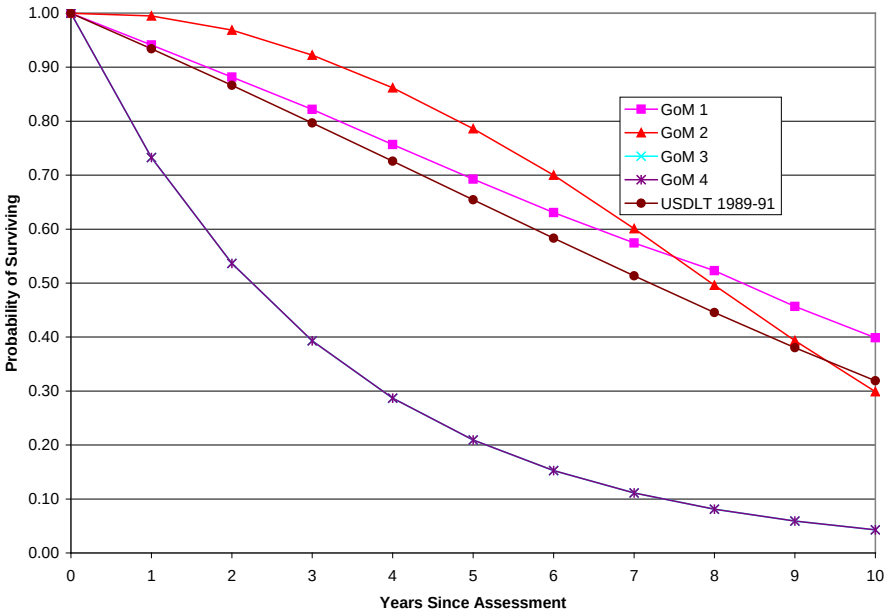
Importantly, given that both the survival curve and its slope are used to price a loan (the former to determine the duration of loan payments, the latter for loan maturity offsets), it is important to accurately estimate these quantities.

On the above graph, you will note that an initial GoM 2 has approximately a 16% greater likelihood of surviving through years 1 through 5 until the lines converge at the year 6, near the 6.2 year LE. On the other hand, an initial GoM 3 or 4 has approximately a 75% lower likelihood of surviving through years 1 through 5, with corresponding reductions for persons who have initial *fractional* scores on GoM 3 or 4, with complementary *fractional* scores on GoM 1 and/or 2. The sum of all fractional scores must equal 1.0 (100%), with the scores for any given individual derived from his/her answers to the 76 questions on the application form.

The corresponding graph (Fig. 2) for females aged 82 years at assessment displays similar patterns, but with a somewhat longer LE, 7.8 years in the USDLT, which is less than the LEs of 9.2 and 8.2 years for GoM 1 and 2, but is substantially higher than the LEs of 3.2 years each for GoM 3 and 4 (as for males, the values are the same because an initial GoM 3 "converts" to GoM 4 prior to age 82).

**Figure 2**

**Probability of Surviving for Each of 10 Years Following Initial Assessment at Age 82, Females, by Initial Time-Invariant GoM Pure Type and for the U.S. Decennial Life Table for 1989-91**



For both sexes, the differences in the survival curves and their slopes illustrate the potential for substantial mispricing without this additional knowledge concerning the level and slope of the relevant survival curves. Individual moveout level pricing would correct this problem and accurately match loan payments to the senior's anticipated tenure in their home.

**Longevity Cost Calculator – Actual vs Expected Accuracy:** The Longevity Cost Calculator™ web based model is a validated replication of the original peer reviewed model published in the *North American Actuarial Journal*. The original model was calibrated using 120,832 male person-years of consecutive assessment data and 196,270 female person-years over an 18-year period commencing with the 1984 National Long Term Care Survey. The annual mortality probabilities were based on 20,428 deaths (8,583 males and 11,845 females) among 32,389 participants in the survey (12,974 males and 19,415 females) in the 18-year period.

Below are a table (Table 1) and two graphs (Figs. 3–4) reproduced from the 2007 *NAAJ* paper. Table 1 shows the predicted probability of death within each year contrasted to the observed death of individuals in the assessment population. These are shown in age brackets by the 5-year age groups (age at start of each 1-year follow-up) used for model estimation as well as the total over age of the observed vs predicted mortality for males and females.

**Table 1**

**Probabilities of Death within One Year in Four Pure-Type GoM Model, Adjusted for Declines in Vitality, by Sex and Attained Age at Time of Exposure**

| Exposure Age | No. of Person-Years at Risk[1] | Annual Probability by Type | | | | Observed Probability | Predicted Probability |
|---|---|---|---|---|---|---|---|
| | | I | II | III | IV | | |
| Males | | | | | | | |
| 65–69 | 20,323 | 0.000 | 0.000 | 0.132 | 0.138 | 0.031 | 0.032 |
| 70–74 | 38,255 | 0.002 | 0.005 | 0.194 | 0.246 | 0.043 | 0.044 |
| 75–79 | 31,291 | 0.038 | 0.013 | 0.319 | 0.319 | 0.067 | 0.067 |
| 80–84 | 19,170 | 0.095 | 0.023 | 0.340 | 0.340 | 0.105 | 0.106 |
| 85–89 | 8,117 | 0.127 | 0.202 | 0.330 | 0.330 | 0.154 | 0.155 |
| 90–94 | 2,728 | 0.198 | 0.323 | 0.323 | 0.323 | 0.228 | 0.229 |
| 95–99 | 793 | 0.226 | 0.434 | 0.434 | 0.434 | 0.301 | 0.301 |
| 100–104 | 155 | 0.372 | 0.528 | 0.528 | 0.528 | 0.400 | 0.401 |
| Total | 120,832 | 0.041 | 0.033 | 0.253 | 0.271 | 0.071 | 0.072 |
| Females | | | | | | | |
| 65–69 | 25,424 | 0.000 | 0.000 | 0.081 | 0.140 | 0.017 | 0.017 |
| 70–74 | 52,008 | 0.001 | 0.003 | 0.108 | 0.223 | 0.027 | 0.027 |
| 75–79 | 48,498 | 0.018 | 0.003 | 0.249 | 0.249 | 0.043 | 0.043 |
| 80–84 | 35,563 | 0.059 | 0.005 | 0.267 | 0.267 | 0.070 | 0.070 |
| 85–89 | 20,404 | 0.089 | 0.110 | 0.271 | 0.271 | 0.115 | 0.115 |
| 90–94 | 9,577 | 0.127 | 0.272 | 0.272 | 0.272 | 0.183 | 0.184 |
| 95–99 | 3,804 | 0.168 | 0.388 | 0.388 | 0.388 | 0.264 | 0.264 |
| 100–104 | 992 | 0.274 | 0.499 | 0.499 | 0.499 | 0.325 | 0.324 |
| Total | 196,270 | 0.036 | 0.037 | 0.201 | 0.239 | 0.060 | 0.061 |

[1] Includes up to four observations per respondent; excludes respondents aged 65–69 in 1999.
*Source:* Author's calculations based on data from the NLTCS.

The differences between the observed and predicted age-specific death probabilities are very small and are statistically nonsignificant, with chi-squared values of 1.36 and 0.29, respectively, each with 6 d.f. (reference values are 12.59 and 16.81 at the conventional 5% and 1% significance levels). The graphs show the same data but add the corresponding death probabilities from the U.S. Decennial Life Table (USDLT) for 1989-91.

**Figure 3**
**Probability of Death within One Year, Males**

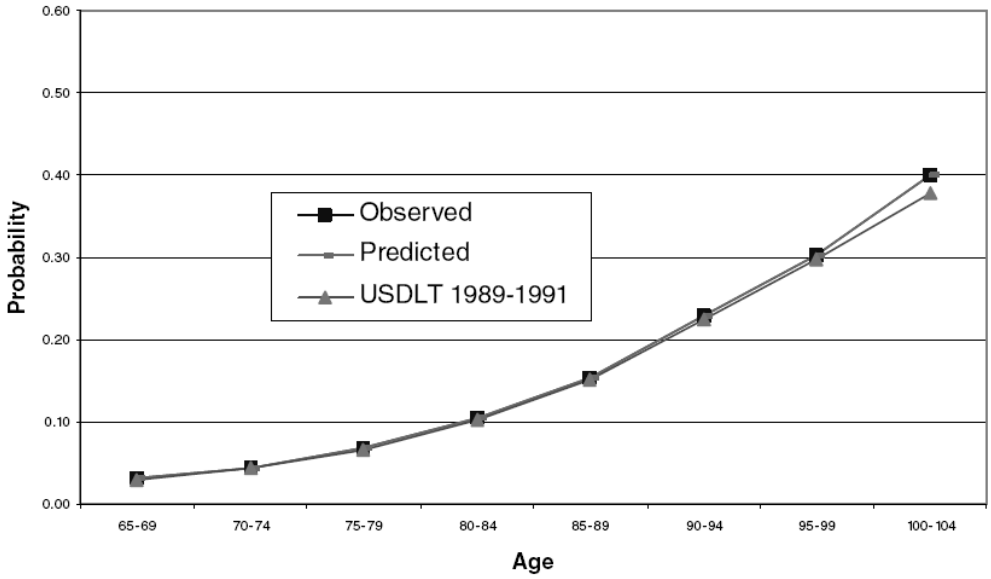
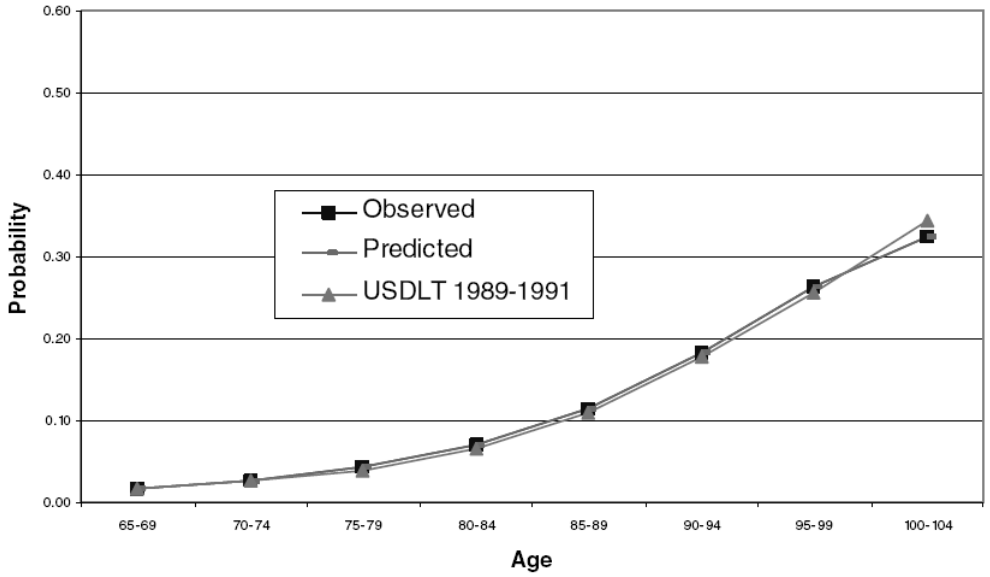



**Figure 4**
**Probability of Death within One Year, Females**



Below we attempt to translate these values into the numbers presented by commercial LE providers who represent that they are accurate 86% to 96% of the time. They do not provide nor publish the basis for their estimates. Hence, we cannot be certain that our measures of accuracy are fully comparable with theirs.

To begin consideration of measures of accuracy, it is useful to consider the random statistical fluctuations that are expected from estimates based on the different numbers of events likely to be observed in different sized samples, assuming fixed underlying event rates. The table below shows the minimum number of expected events needed to cap the maximum relative error (e.g., [observed number of deaths minus expected number of deaths] ÷ [expected number of deaths]) with one of three levels of confidence.

**Table 2**
**Credibility and Event Counts**

| Maximum Acceptable Departure from the Expected Count | Probability of Observed Count Falling Within the Acceptable Range | | |
|---|---|---|---|
| | 90% | 95% | 99% |
| | Minimum Required Expected Count | | |
| +/–2.5% | 4,329 | 6,146 | **10,616** |
| +/–5.0% | **1,082** | 1,537 | 2,654 |
| +/–7.5% | 481 | 683 | 1,180 |
| +/–10% | 271 | 384 | 663 |
| +/–20% | 68 | **96** | 166 |
| +/–30% | 30 | 43 | 74 |
| +/–40% | 17 | 24 | 41 |
| +/–50% | **11** | 15 | 27 |

Source: Based on Longley-Cook (1962).

To be 99% confident that the maximum relative error is less than 2.5%, the sample size needs to be large enough to produce 10,616 deaths (in bold in the table). The annual mortality probabilities in the *NAAJ* analysis were based on 20,428 deaths (8,583 males and 11,845 females), indicating that the total rates are very stable but the stratifications by age and other variables may be affected by random statistical fluctuations. To be 90% confident that the maximum relative error is less than 5%, the sample size needs to be large enough to produce 1,082 deaths, which is the standard size used for full credibility in the actuarial literature. To be 95% confident that the maximum relative error is less than 20%, the sample size needs to be large enough to produce 96 deaths, which rounds to about 100. To be 90% confident that the maximum relative error is less than 50%, the sample size needs to be large enough to produce 11 deaths, which rounds to about 10. Thus, as the expected number of deaths falls from 10,000 to 1,000 to 100 to 10, the relative error increases from about 2.5% to 50%.

Practical considerations often dictate sample sizes less than that needed for full actuarial credibility. Table 2 indicates that sample sizes of 271 and 384 can yield relative errors of ±10% at the 90% and 95% confidence levels, respectively.

Random statistical fluctuations are inherently unpredictable. Hence, our measures of accuracy must focus on our ability to generate accurate values for the expected number of deaths among any selected set of insured lives.

The tables and graphs from the *NAAJ* paper (Table 1; Figs. 3–4) show that this can be done for groups of insured lives when the groups are defined on the basis of age and sex.

The next two graphs (Figs. 5–6) show the performance of the model when the NLTCS mortality-exposure data are grouped into 10 categories according to the predicted probability of death, based on the use of fixed cutpoints at multiples of .05 (5%), separately for males and females. Chi-squared statistical tests of fit of the models are presented separately in Tables 3–4.

**Figure 5**

**Observed and Predicted Probabilities of Death, Males, by Predicted-Probability Class Intervals with Cutpoints at Multiples of 5%**

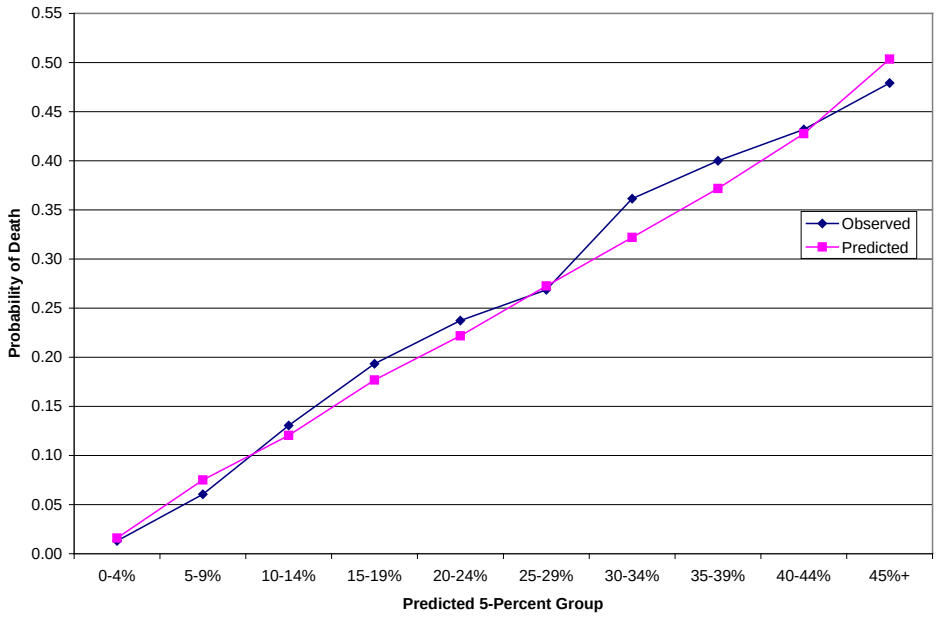
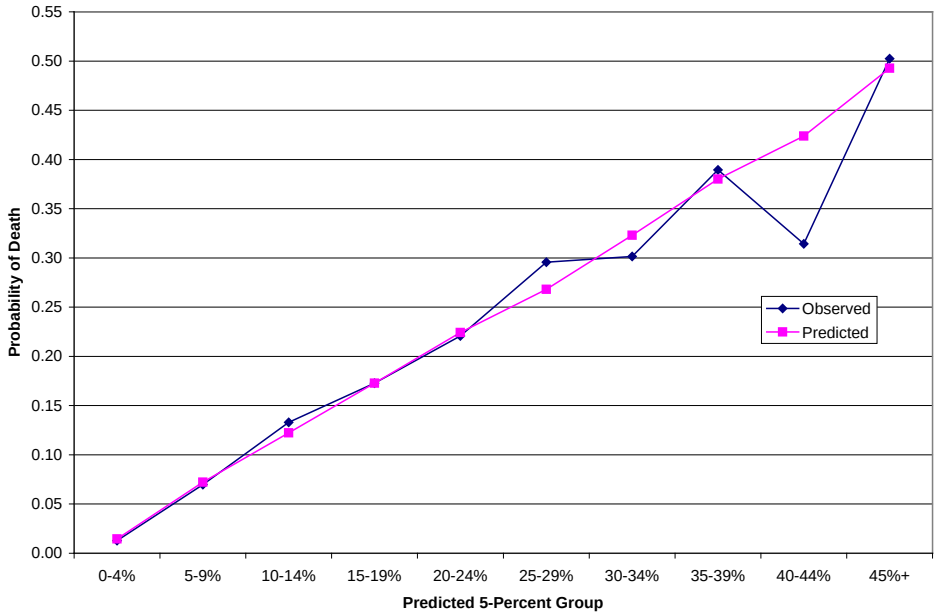


**Figure 6**

**Observed and Predicted Probabilities of Death, Females, by Predicted-Probability Class Intervals with Cutpoints at Multiples of 5%**




Visually, one can see that the observed probabilities increase across the 10 categories for both sexes, except for category 9 for females. However, the chi-square test of the deviation for that one point indicates that the difference is statistically nonsignificant, with a chi-squared value of 3.44 with 1 df (Table 4; reference values are 3.84 and 6.63 at the conventional 5% and 1% significance levels).

**Table 3**
**Observed and Predicted Probabilities of Death, Males, by Predicted-Probability Class Intervals with Cutpoints at Multiples of 5%**

| Percentage Group | Number of Person-Years at Risk | Observed Number of Deaths | Expected Number of Deaths | Observed Probability | Predicted Probability | Hosmer-Lemeshow Chi-Squared |
|---|---|---|---|---|---|---|
| 0-4% | 61,463 | 801 | 981 | 0.013 | 0.016 | **33.60** |
| 5-9% | 23,256 | 1,407 | 1,747 | 0.061 | 0.075 | **71.73** |
| 10-14% | 20,100 | 2,622 | 2,420 | 0.130 | 0.120 | **19.20** |
| 15-19% | 7,705 | 1,490 | 1,363 | 0.193 | 0.177 | **14.29** |
| 20-24% | 4,557 | 1,082 | 1,011 | 0.237 | 0.222 | 6.44 |
| 25-29% | 2,095 | 563 | 571 | 0.269 | 0.273 | 0.16 |
| 30-34% | 1,361 | 492 | 438 | 0.361 | 0.322 | **9.73** |
| 35-39% | 115 | 46 | 43 | 0.400 | 0.372 | 0.39 |
| 40-44% | 132 | 57 | 56 | 0.432 | 0.428 | 0.01 |
| 45%+ | 48 | 23 | 24 | 0.479 | 0.504 | 0.11 |
| Total | 120,832 | 8,583 | 8,655 | 0.071 | 0.072 | **155.65** |

**Table 4**
**Observed and Predicted Probabilities of Death, Females, by Predicted-Probability Class Intervals with Cutpoints at Multiples of 5%**

| Percentage Group | Number of Person-Years at Risk | Observed Number of Deaths | Expected Number of Deaths | Observed Probability | Predicted Probability | Hosmer-Lemeshow Chi-Squared |
|---|---|---|---|---|---|---|
| 0-4% | 111,425 | 1,434 | 1,614 | 0.013 | 0.014 | **20.33** |
| 5-9% | 44,837 | 3,124 | 3,239 | 0.070 | 0.072 | 4.42 |
| 10-14% | 18,004 | 2,396 | 2,202 | 0.133 | 0.122 | **19.46** |
| 15-19% | 10,245 | 1,770 | 1,770 | 0.173 | 0.173 | 0.00 |
| 20-24% | 6,536 | 1,443 | 1,465 | 0.221 | 0.224 | 0.43 |
| 25-29% | 3,413 | 1,009 | 915 | 0.296 | 0.268 | **13.17** |
| 30-34% | 617 | 186 | 199 | 0.301 | 0.323 | 1.32 |
| 35-39% | 914 | 356 | 347 | 0.389 | 0.380 | 0.35 |
| 40-44% | 70 | 22 | 30 | 0.314 | 0.424 | 3.44 |
| 45%+ | 209 | 105 | 103 | 0.502 | 0.493 | 0.08 |
| Total | 196,270 | 11,845 | 11,885 | 0.060 | 0.061 | **62.99** |

The Hosmer-Lemeshow chi-squared test produces statistically significant total chi-squared values of 155.65 and 62.99, respectively, for males and females, each with 8 d.f. (reference values are 15.51 and 20.09 at the conventional 5% and 1% significance levels).

Several comments are in order:

1. The tests indicate that the models displayed in Figures 5 and 6 do *not* fit the data. This means that at least some of the deviations of the observed from predicted numbers of deaths are larger than expected by chance.
2. These are identified by the boldface font in the rightmost columns of Tables 3 and 4 using a cutpoint equal to the critical value of 6.63 using the conventional 1% significance level.
3. For males, the five significant deviations are for the four lowest probability groups 0–19% and 30–34%; for females, the three significant deviations are for the 0–4%, 10–14%, and 25–29% probability groups.
4. The expected counts for the eight groups with significant deviations ranged from 438 to 2,420, and five of the eight exceed the 1,082 cutpoint for highly credible data in Table 2.

We conclude that there are factors operating in these data that are not represented in our model. This should not be surprising given that our model uses four GoM scores to summarize data on the full set of 76 questions concerning medical conditions, activities of daily living (ADLs), cognitive and behavioral impairments. Moreover, there may be other influential factors not included in our set of 76 questions. Given a sufficiently large sample one would expect to identify significant deviations from *any model* using the statistical procedures described above.

Two additional comments provide additional perspective:

1. Only two groups had significant deviations for both sexes: the 0–4% and 10–14% probability groups. This suggests that the deviations from the models are not predictable.
2. Nonrandom deviations in the LCC models can be tolerated if they are sufficiently small, relative to the errors that would occur in the absence of the LCC models.

To quantify the size of the nonrandom deviations, Stallard applied linear regression analysis with the observed probabilities regressed on the predicted probabilities, as shown in Tables 3 and 4, obtaining R-squared values of 0.985 and 0.942, respectively.

The average of these two R-squared values is 0.964 which may be interpreted as a measure of the accuracy of the LCC models: 96.4% of the variance of the observed probabilities is accounted for by the expected probabilities produced by the LCC models. The remaining 3.6% of the variance constitutes a tolerable level of nonrandom deviations in the LCC models.

We considered the possibility that the linear regression analysis may not fully represent the impact of small deviations at the lower probability levels in Tables 3 and 4. This was motivated in part by the chi-squared tests which indicated that most of these deviations were statistically significant. To deal with this issue, we generated a second set of regressions with the logarithms of the observed probabilities regressed on the logarithms of the predicted probabilities, obtaining R-squared values of 0.994 and 0.989, respectively.

The average of these two R-squared values is 0.991 which may be interpreted as an alternative measure of the accuracy of the LCC models: 99.1% of the variance of the logarithm of the observed probabilities is accounted for by the logarithm of the expected probabilities produced by the LCC models. The remaining 0.9% of the variance constitutes an even more tolerable level of nonrandom deviations in the LCC models.

Two additional questions are important to our assessment of the accuracy of the model.

The first question is whether the GoM scores add any significant information beyond the information already available using the age-specific mortality probabilities displayed in Figures 3–4; and if so, how much? This question can be directly addressed using the log–likelihood-ratios for the four sex-specific models listed in Table 5 to generate the corresponding AIC and BIC statistics typically used for model assessment.

**Table 5**

| # | Model Description | Log-Likelihood-Ratio | d.f. | AIC | BIC | ΔAIC | ΔBIC |
|---|---|---|---|---|---|---|---|
| | | | **Males** | | | | |
| 1 | Constant Probability | 0.00 | 1 | 2.00 | 9.06 | 10,878.66 | 10,659.87 |
| 2 | Age-Specific Probabilities (no GoM) | 1,632.10 | 8 | -3,248.19 | -3,191.73 | 7,628.46 | 7,459.08 |
| 3 | GoM-Specific Probabilities (no Age) | 5,278.87 | 4 | -10,549.75 | -10,521.52 | 326.91 | 129.30 |
| 4 | Age&GoM-Specific Probabilities | 5,470.33 | 32 | **-10,876.66** | **-10,650.82** | 0.00 | 0.00 |
| | | | **Females** | | | | |
| 1 | Constant Probability | 0.00 | 1 | 2.00 | 9.38 | 16,276.04 | 16,047.27 |
| 2 | Age-Specific Probabilities (no GoM) | 3,776.30 | 8 | -7,536.59 | -7,477.56 | 8,737.45 | 8,560.34 |
| 3 | GoM-Specific Probabilities (no Age) | 7,873.29 | 4 | -15,738.58 | -15,709.06 | 535.46 | 328.83 |
| 4 | Age&GoM-Specific Probabilities | 8,169.02 | 32 | **-16,274.04** | **-16,037.89** | 0.00 | 0.00 |

Model 1 is the simplest model. It assumes that the sex-specific mortality probability is constant over age and GoM scores. Model 2 assumes that the sex-specific mortality probabilities increase over age but not over GoM scores, following the observed values displayed in Figures 3–4. Model 3 assumes that the sex-specific mortality probabilities increase over GoM scores but not over age, following the values displayed in the Totals row of Table 7 of the *NAAJ* paper. Model 4 assumes that the sex-specific mortality probabilities increase over GoM scores and over age, following the values displayed in the age-specific rows of Table 7 of the *NAAJ* paper.

The log–likelihood-ratios were generated as the difference in the value of the log–likelihood for each model and the log–likelihood for Model 1. The degrees of freedom (d.f.) were defined as the number of parameters in each model. The AIC (Akaike Information Criterion) was calculated as the log–likelihood-ratio plus 2 times the d.f. The BIC (Bayesian Information Criterion) was calculated as the log–likelihood-ratio plus the product of the d.f and the logarithm of the number of deaths. The best model is the one that has the minimum value of AIC or BIC (indicated with boldface font in Table 5); the relative performance of each model is assessed by the difference between its value of AIC or BIC and the minimum value of these statistics (labeled ΔAIC or ΔBIC in Table 5). Differences of 10 or more points are regarded as strong evidence in support of the model with the lower AIC or BIC value.

For both sexes and both criteria, Model 4 is overwhelmingly selected as the best model.

To determine whether the GoM scores add significant information beyond the information already available using the age-specific mortality probabilities, we need to compare the value of ΔAIC or ΔBIC for Model 2 with the reference value of 10. For both sexes and both criteria the values exceed the reference values by a factor of 746–874, indicating that the additional information provided by the GoM scores is huge.

Comparison of Models 2 and 3 provide additional confirmation of the power of the GoM model. The ΔAIC and ΔBIC for Model 2 are each about 30.0% smaller than the

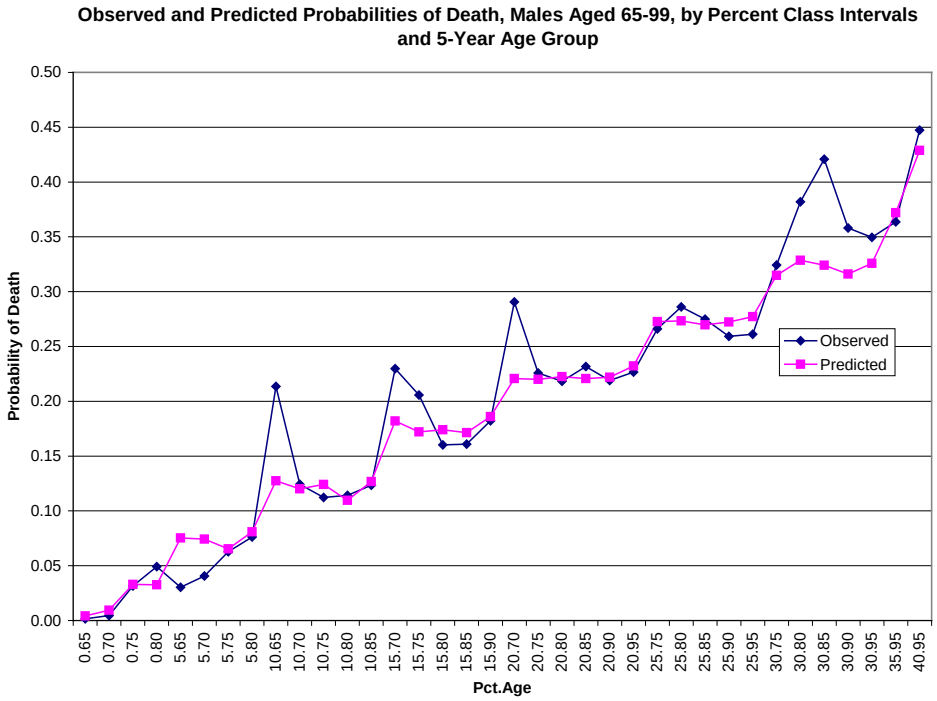
corresponding value for Model 1 for males and about 46.5% smaller for females. The ΔAIC and ΔBIC for Model 3 are each about 97.0% smaller than the corresponding value for Model 1 for males and 95.0% smaller for females. This means that if one were forced to choose between Models 2 and 3, then Model 3 would be selected as the better model and that the improvement offered by Model 3 would be huge. Model 3 would offer 95–97% of the improvement over Model 1 that could ultimately be obtained using Model 4. This would be far in excess of the 30–46% improvement offered by Model 2.

The second question is whether the excellent calibration displayed in Figures 5 and 6 continues when the predictions are stratified by age-groups. The results of this stratification are displayed in Figures 7 and 8 for males and females aged 65–99 years old. The death counts for males at age 100+ fell below the standard CMS cutoff of 11 events and hence were suppressed.

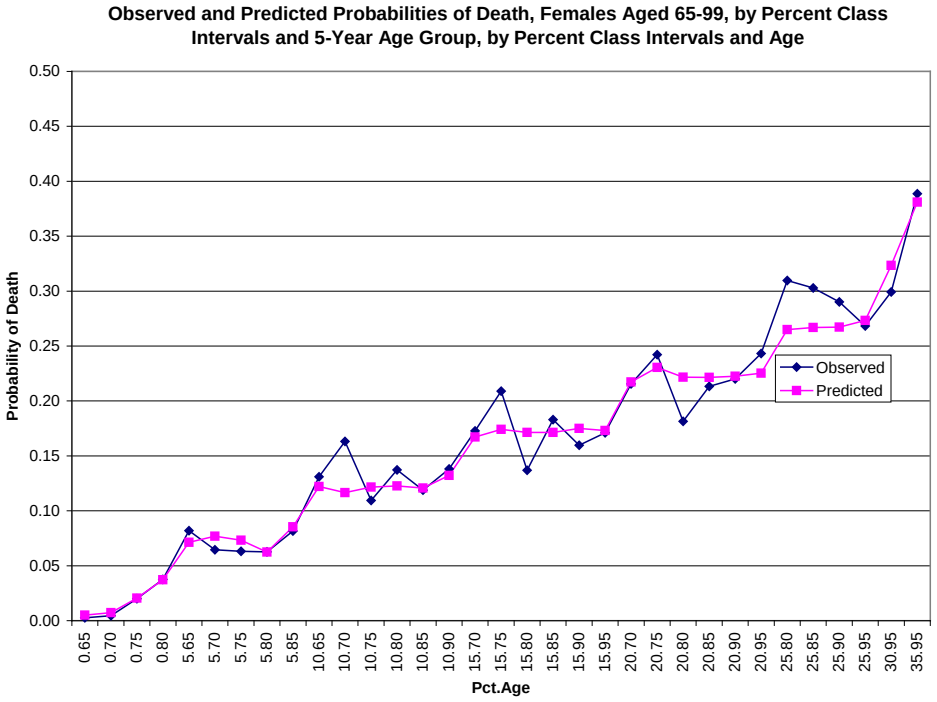
For comparability, females were restricted to the same age range. The aberrant point in Figure 6 for the 40–44% group turned out to be solely for females aged 100+ which meant that this point was excluded from Figure 8.

The labeling of the groups (Pct.Age) in Figures 7 and 8 combines the lower bounds of the 5-Percent labels in Figures 5 and 6 with the lower bounds of the age-groups in Figures 3 and 4. Thus, 0.65 identifies persons aged 65–69 years with predicted probabilities in the range 0–4%; similarly 35.95 identifies persons aged 95–99 years with predicted probabilities in the range 35–39%. The groups are ordered by increasing predicted probabilities, and within each probability group, by increasing age.

## Figure 7




## Figure 8




Examination of the results for males and females in Figures 7 and 8 shows that the largest deviations are for males aged 65–69 (with offsetting deviations for 5–9% and

10–14%) and aged 70–74 (with offsetting deviations for 5–9% and 15–19%), and that these same deviations are not replicated for females.

One possible explanation for the male result involves the design of the NLTCS in which persons who do not have ADL or IADL impairments at the time of the survey do not receive the in-person assessment; such persons "screen out" without answering any of the detailed health questions. Thus, the estimates of their GoM scores have substantially larger errors than would be the case for persons who answered most or all of the 95 health questions in the *NAAJ* model, or the 76 questions in the LCC.

To quantify the size of the deviations, we applied linear regression analysis with the observed probabilities regressed on the predicted probabilities, as done above for the data in Figures 5 and 6, obtaining R-squared values of 0.943 and 0.960, respectively, for males and females, with an overall average of 0.952. As done above, we generated a second set of regressions with the logarithms of the observed probabilities regressed on the logarithms of the predicted probabilities, obtaining R-squared values of 0.958 and 0.984, respectively, with an overall average of 0.971. The implied accuracy is thus in the range 95–97%, depending on the form of the regression.

The 32,389 individuals observed in consecutive assessments over the term of this study are a random sample of Medicare enrollees who participated in the National Long-term Care Survey. Assessments began either in 1982 when the NLTCS started, or at a later date when the participants were aged 65–69. Because homeownership data was not collected, we do not know how many of these seniors were homeowners. What we do know from Census Bureau data is that more than 70% of seniors age 65 are homeowners.

For those preferring medical study validity values, unpublished analyses using out-of-sample data from the Medicare Current Beneficiary Survey yielded a ROC score of 0.8, which for actuarial/medical studies represents a high degree of accuracy.

## Reverse Mortgage Loan Pricing Using the Longevity Cost Calculator

The 76 questions used in the Longevity Cost Calculator are today both built into an application form used for life settlements and into several automated web site applications used by settlement brokers, providers and financial professionals. At the end of the automated input all of the relevant morbidity and mortality data needed for a reverse mortgage pricing model are available. The use of this model to determine moveout in lieu of Gompertz type exponential mortality tables would allow individual loan level pricing of reverse mortgages. Homeowners' payments would be tuned to their tenure in the home. Investors would be buying securitized portfolios of loans where the cash inflows for loan maturities and outflows to make payments more closely match the tenure of individual seniors in their homes.

Annual mark-to-market valuations of these portfolios would produce more accurate values. A problem would exist under the auditing and accounting framework until a second methodology for valuation could be agreed upon.

For the life settlement sister asset class there is already an accepted methodology of using third-party life expectancy consultants to estimate an insured's life expectancy using their medical records. The underwriter determines debits and credits that are combined into a multiplier that is applied to a standard mortality table. These LE estimates, based upon Gompertz exponential large population mortality tables, are flawed for all the same reasons. But these commercial LEs offer a strong comparative starting point and the alternative valuation methodology needed for mark-to-market revaluation.

## Existing Accounting Framework

Established business asset valuation measurement is divided into three levels dependent upon the credibility and acceptance of the data available to establish the asset's value:

1. Level One measurement is determined by reference to quoted prices for identical assets in the active reference market such as stock exchange listed prices.
2. Level Two measurement is determined by valuation methods using observable data. The valuation methods are the cost approach, present value approach and the market approach. This method would only apply to valuing settlements after the death of the insureds.
3. Level Three measurement is determined by valuation methods using unobservable data. The valuation methods are the cost approach, present value approach and the market approach. Level Three is the only appropriate method of measurement for life settlements using the present value approach applied to each loan in a pool.

The framework exists today to create uniformity in reverse mortgage longevity underwriting and related disclosure by applying existing GAAP at the time new loans are originated, existing loans are pooled for securitization and annually at revaluation. These loan portfolios, whether rated or unrated, will require disclosure at the time they are pooled and sold, and subsequently when the pools are individually audited or because they are held for investment by issuers of audited financial statements. For level 3 assets where value is dependent upon a future unobservable outcome (the borrowers moveout) GAAP requires the use of the best information available, a first valuation method and then a second independent corroborating valuation methodology. The outcome of these two methodologies is then reconciled and disclosed. Specific relevant accounting doctrine includes AU §328, (previously SAS101), *Auditing Fair Value Measurements and Disclosures,* (June 2003); Accounting Standards Codification (ASC) 820.10.05 through 820.10.65, (previously FASB 157), *Fair Value Measurement and Disclosure*.

AU §328.03 states “ASC glossary term fair value is defined as the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants at the measurement date. Although GAAP may not prescribe the method for measuring the fair value of an item, it expresses a preference for the use of observable market prices to make that determination.” Clearly a life settlement is a “level 3” asset whose value is dependent upon an unobservable variable – the future death of the insured[16]. “In the absence of observable market prices, GAAP requires fair value to be based on the best information available in the circumstances.”

AU §328.06 states “Assumptions used in fair value measurements are similar in nature to those required when developing other accounting estimates. However, if observable market prices are not available, GAAP requires that valuation methods incorporate assumptions that marketplace participants would use in their estimates of fair value whenever that information is available without undue cost and effort. If information about market assumptions is not available, an entity may use its own assumptions as long as there are no contrary data indicating that marketplace participants would use different assumptions.”

AU §329.18 goes on to say “Management may have determined that different valuation methods result in a range of significantly different fair value measurements. In such cases, the auditor evaluates how the entity has investigated the reasons for these differences in establishing its fair value measurement.”

AU §328.23, under the heading *Testing the Entity’s Fair Value Measurements and Disclosures*, presents the following. “For example, substantive tests of the fair value measurements may involve (a) testing management’s significant assumptions, the valuation model[s], and the underlying data (see paragraphs .26 through .39), (b) developing independent fair value estimates for corroborative purposes (see paragraph .40), or (c) reviewing subsequent events and transactions (see paragraphs .41 and .42).”

AU 328.40 under the heading *Developing Independent Fair Value Estimates for Corroborative Purposes* states, “The auditor may make an independent estimate of fair value (for example, by using an auditor-developed model) to corroborate the entity’s fair value measurement.[17] When developing an independent estimate using management’s assumptions, the auditor evaluates those assumptions as discussed in paragraphs .28 to .37. Instead of using management’s assumptions, the auditor may develop his or her own assumptions to make a comparison with management’s assumptions. The auditor uses that understanding to ensure that his or her independent estimates takes into consideration all significant variables and to evaluate any significant differences from management’s estimates. The auditor also

[16] ACS 820-10-35-52 sets forth the parameters for level 3 assets where fair value is determined from unobservable future events.
[17] See AU §329, *Analytical Procedures*

should test the data used to develop the fair value measurements and disclosures as discussed in paragraph .39."

**Suggested GAAP Compliant Methodology to Standardize Life Settlement Underwriting**

Because AU §328 will require auditors to employ two valuation methodologies, the most cost-efficient way to comply will be to both use a standardized Longevity Cost Calculator™ compliant questionnaire to yield an LCC LE at the time of loan application or at the time of mandatory loan counseling. Once a second methodology has been agreed upon, that methodology will be used at origination to develop a "shadow pricing" for each loan. At the time of annual audit this will allow the actual versus expected values for each loan's moveout to be evaluated and ranked accordingly. The moveout evaluations will be conducted using Bayesian Information Criterion (BIC) measures of goodness of fit of the actual to expected probabilities of moveout computed separately for each moveout model using the same pool of loans. Kass and Raftery[18] showed how BIC values can be used (1) to rank the various moveout models and (2) to generate optimal weighted averages of the outputs of the various moveout models. The resultant weighted averaging can then be applied to each individual loan, and to the aggregate of all loans, in annually revaluing each portfolio. The BIC weights can be updated each year as additional information on the actual number of moveouts in that year becomes available. Over time, this will give greater weight to the better performing models. Disclosure of this methodology at portfolio formulation and annually will create transparency into these longevity valued asset transactions for investors.

.

Peter M. Mazonas
Co-Managing Member
Life Settlement Financial, LLC
(415) 526-5972
peter.mazonas@lifesettlementfinancial.com
https://www.lifesettlementfinancial.com

[18] See equations (9), (16), and (18) in .R.E. Kass and A.E. Raftery. Bayes Factors. *Journal of the American Statistical Association* 90(430):773-795, 1995. Note that BIC values can be generated using various approaches to measuring goodness of fit, including chi-squared statistics and regression-based $R^2$-statistics.