Longevity Risk in Fair Valuing Level-Three Assets in Securitized Portfolios

Peter M. Mazonas, CPA Life Settlement Financial, LLC

P.J. Eric Stallard, ASA, MAAA, FCA, Advisor to Life Settlement Financial, LLC

Lynford Graham, CPA, PhD Bentley University

e-mail: peter.mazonas@lifesettlementfinancial.com

ABSTRACT

Level-three longevity valued assets pose unique valuation risks as securitized pools of these alternative asset classes come to market as investment vehicles for pension plans and individual retirement accounts. No uniform framework is applied to assure consistent fair market valuation and transparency for investor decision-making. Applying existing international auditing standards, IAS §540, and analytical procedures, IAS §520, offers a platform upon which fund managers, their auditors and actuaries can agree upon uniform valuation and presentation guidelines. Application of these quasi-governmental standards will bring future liquidity to otherwise illiquid capital market instruments.

This paper presents a valuation methodology consistent with fair value accounting and auditing standards. The methodology incorporates the longevity predictive modeling of Stallard, *NAAJ* 2007, and is compatible with Bayes Factor weighted average valuation techniques from Kass and Raftery, 1995. Securitizers continue to hide behind the theory of large numbers, despite the work of Milevsky, 2006. The size of securitized portfolios coming to market do not have credibility, Longley-Cook, 1962. This is observed in life settlement portfolios where the combination of too few large death benefit policies and large variances in life expectancy estimates challenge accurate valuation and periodic re-valuation.

INTRODUCTION

Level-three longevity valued assets pose unique valuation risks as securitized pools of these alternative asset classes come to market as investment vehicles for pension plans and individual retirement accounts. No uniform framework has yet to be established to assure consistent fair market valuation and transparency for investor decision-making. Recently both the International Accounting Standards Board (IASB) and the U.S. Financial Accounting Standards Board (FASB) have issued draft standards requiring these assets to be accounted for and reported using fair value accounting.[1] [2]Auditors have also developed auditing standards that apply to assets that are fair valued. Applying these clarified auditing and accounting standards, IAS §540 and AU §328[3], analytical procedures, IAS §520 and AU §329 and Accounting Standards Codification, ASC §820, Fair Value Measurement and Disclosure, offer a platform upon which fund managers, their auditors and actuaries can agree upon uniform valuation and presentation guidelines. Reconciling the variations in longevity estimates and then applying these professional standards will increase confidence in the valuation of these securities, bringing further liquidity to these rapidly expanding capital market instruments.

FASB ASC §820.10, formerly FAS 157 and 159, lays out the framework for the fair value hierarchy and prioritizes the inputs to valuation techniques. The inputs (variables and methods used) are determined by the lowest level, level-one through level-three, in which the fair value measurement falls. Life settlements, annuities and reverse mortgages are level three assets because their fair value is determined by an unobservable future event – the death of the insured, annuitant or borrower.

The relevant AICPA/PCAOB standards are rules-based, known as SAS101, and effective since June 2003. The IASB standards are principles based, but the thrust of the accounting and auditing standards vary only slightly and have the same intent of routinely marking portfolios to fair value while providing clear, understandable and transparent investor reporting. The open question is one of implementation in the murky world of individual level life expectancy estimates. This issue is relevant for life settlements, life settlement derivative and leveraged debt products, annuities and reverse mortgages.

---

[1] On May 26, 2010, the Financial Accounting Standards Board (FASB) issued an Exposure Draft of a Proposed Accounting Standards Change: *Accounting for Financial Instruments and Revisions to the Accounting for Derivative Instruments and Hedging Activities*, Financial Instruments (Topic 825) and Derivatives and Hedging (Topic 815). The topic sections refer to the Board's new Accounting Standards Codification (ASC) and are included in sections 820.10, Fair Value Measurements and Disclosures. And are available at: http://asc.fasb.org/topic&trid=2155941&nav_type=left_nav&analyticsAssetName=home_page_left_nav_topic

[2] Board Comment number 40 in the Exposure Draft states: "The Board decided that life settlement contracts should be included in the scope of the proposed guidance. The Board observed that requiring fair value measurement would, in effect, eliminate the option to use the investment method described in Subtopic 325-30."

[3] AU 328 *Auditing Fair Value Measurement and Disclosures* superceded AU 342 *Auditing Accounting Estimates*. AU 328 expanded auditor responsibilities when fair values were involved. See "The Auditor's Approach to Fair Value" (Susan Menelaides, Lynford Graham and Gretchen Fischbach). *Journal of Accountancy* (June, 2003), pages 73 – 76.

Large population-based mortality tables, starting with Gompertz (1825), are known to be flawed (Iachine et al. 1998)[4]. Actuaries apply the law of large numbers (LLN) under the assumption that individual lives are stochastically independent to try to compensate. Moshe Milevsky and colleagues, in their paper *Killing the Law of Large Numbers, Mortality Risk Premiums and the Sharpe Ratio*, presented at the 2006 meeting of this conference, clearly differentiated between the deterministic case of a known survival curve and the more realistic case where the survival probabilities are unknown but can be modeled using stochastic hazard rates[5]. For the latter case, Milevsky demonstrated that there are inherent standard deviations (i.e., errors) in mortality tables that do not converge to zero, but instead converge to a positive constant (which may be material), when the number of lives in the table is increased without bound. The latest life insurance industry mortality tables, Valuation Basic Table 2008 (VBT 2008), are known to have largely extrapolated mortality over the age of 75 because of the lack of sufficient historic data[6]. Combined, these results call into question the accuracy of longevity-based valuations employing standard table-based methodologies.

Longley-Cook (1962), using Poisson distribution statistics, quantified the number of claims from among a portfolio of policies required to meet various *credibility standards* for accuracy and reliability[7]. He showed that it requires 384 claims (deaths) to achieve ±10% relative accuracy (i.e., the actual number of claims falls within ±10% of the expected number) with 95% probability and 1,537 claims (deaths) to achieve ±5% relative accuracy with 95% probability. Unfortunately few policy portfolios are sufficiently homogeneous with respect to policy size (face value) and seldom are there sufficient numbers of policies to achieve satisfactory credibility levels. This further amplifies the need to work first to reconcile the life expectancy estimates before applying stochastic principles and Monte Carlo simulations to arrive at the discounted cash flows and thus the value of longevity valued assets.

A real world example of this problem begins with the variations between commercial life expectancies (LEs) used in the life settlement marketplace to price purchase offers. These individual insured's LEs are the output from trained professionals from different

---

4 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 *Research* 1: 196–205.
5 M.A. Milevsky, S.D. Promislow and V.R. Young. Killing the Law of Large Numbers: Mortality Risk Premiums and the Sharpe Ratio. *Journal of Risk & Insurance* 73(4):673-686, 2006.
6 Mike Fasano, Morris Fishman, Charlotte Lee, Phil Loy and Kevin Malone. The Evolving Approach to Developing Life Expectancy Reports. A panel discussion by leading life expectancy provider/underwriters at the Life Settlements Conference, September 25, 2008, Las Vegas.
7 L.H. Longley-Cook. *An Introduction to Credibility Theory*. Casualty Actuarial Society, 1962.

firms all interpreting the same insured's medical records. They apply positive and negative factors and weightings to the items/issues in the medical records to develop a multiplier. This multiplier is applied to a single or multiple modified large population mortality tables based upon the insured's medical conditions to derive the estimated life expectancy for that single individual. These life expectancies and their related survival functions are then used to price the bundled policies.

From data compiled since January 2009 where three or more commercial LEs were prepared from each insured's medical records, the range between the low LE and the high LE per insured, relative to the low LE, averaged 31%[8]. For the same data, where the relative range between the low and the high LE was greater than 30%, the average size was 56%[9]. Herein lies the problem for individual policy pricing and portfolio valuation.

All of this required taking the facts, standards and tools at hand and crafting a fair valuation methodology that would satisfy the fund managers, auditors, and regulators and provide investors transparency. In November 2009, Peter Mazonas was invited to testify before the SEC on the fair valuation of level-three assets, specifically life settlements. This request grew out of paper a colleague, P.J. Eric Stallard, and Peter prepared and presented to the SEC's Life Settlement Task Force[10]. The SEC's concern was the lack of standardized methodology for fair valuation and the historic lack of transparency. Their concern grew out of the proliferation of securitized life settlement financial instruments being sold or being designed to be sold to pension plans as rated and unrated portfolios of policies[11]. What was presented to the SEC and is being described in the current paper is a careful melding of accounting and auditing standards, actuarial science, Bayesian statistics and a relatively new, but peer reviewed and published, approach to predicting mortality in single individuals. The goal is to reconcile the substantial variation in life expectancy estimates to establish an ongoing basis for fair valuation of individual portfolios.

---

8 Results are based upon data compiled from settlement brokers and tabulated by the lead author. EMSI, AVS and others have conducted studies that yield similar results.

9 Ibid.

10 https://www.lifesettlementfinancial.com/pdf/Life%20Settlement%20Porfolio%20Valuation%20for%20Securitization%20LSF%202Nov09.pdf

11 On July 22, 2010, the SEC's cross-divisional Life Settlement Task Force issued recommendations that the Commission recommend to Congress that it amend the Definition of Security under the Federal Securities Law to include life settlements. It further recommended Federal regulation of settlement brokers and provider to ensure legal standards of conduct are being met. It also called for consistent regulation of life expectancy underwriters. http://www.sec.gov/news/studies/2010/lifesettlements-report.pdf

LONGEVITY COST CALCULATOR AS A POLICY AND PORTFOLIO VALUATION TOOL

Background

In the fall of 2007, Life Settlement Financial (LSF), initiated work on a computer model for life settlement valuation which, in addition to provisions for accommodating three or more commercial LEs, also included LE calculations for the model described in Eric Stallard's *Trajectories of Morbidity, Disability and Mortality Among the U.S. Elderly Population*, as published in the Society of Actuaries' *North American Actuarial Journal*[12]. LSF wrote and validated the new life settlement valuation model in robust web-based code and named it the Longevity Cost Calculator (LCC). In mid 2009 LSF launched this model as a settlement policy selection tool that provides low cost individual level survival functions based on "Grade of Membership" scores generated from data gathered in a web-based or telephonic interview.
https://www.lifesettlementfinancial.com/LSFSAP/SAPRegisterPage.aspx

The completed Longevity Cost Calculator assessment questionnaire, in addition to providing an LE, scores each insured using a four level Grade of Membership (GoM) system. As per the earlier referenced material on the LSF web site, the interrelationship of an insured's activity of daily living (ADL) impairments, instrumental activity of daily living (IADL) impairments, 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 model, which for the current implementation is age-group 65–69.

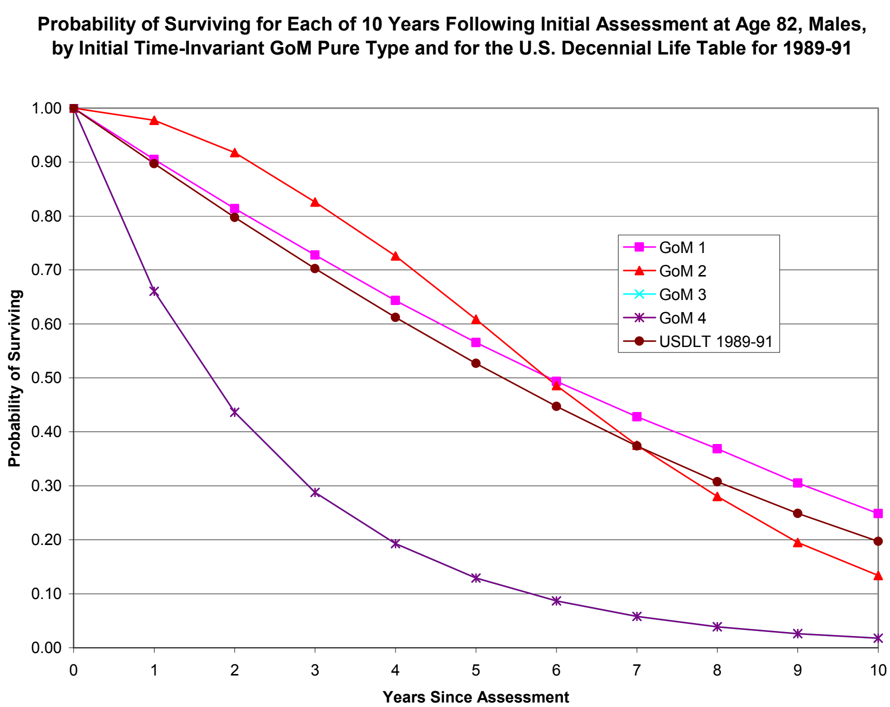
GoM 1 (also referred to as "Pure Type I" or "Type I", with Roman numerals designating the rank ordering of the states by health status; we frequently use the shorter GoM 1–4 reference in the remainder of this paper) refers to the healthiest component of the population. GoM scores 2–4 capture a range 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

[12] E. Stallard. Trajectories of Morbidity, Disability, and Mortality Among the U.S. Elderly Population: Evidence from the 1984-1999 NLTCS. *North American Actuarial Journal* 11(3):16–53, 2007.
http://www.soa.org/library/journals/north-american-actuarial-journal/2007/july/naaj0703-2.pdf

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. Note that GoM 3 and GoM 4 have converged by age 82.

Figure 1



According to the USDLT, the male LE at age 82 was 6.2 years. This value was less than the LEs of 6.9 and 6.3 years for GoM 1 and 2, but was substantially higher than the

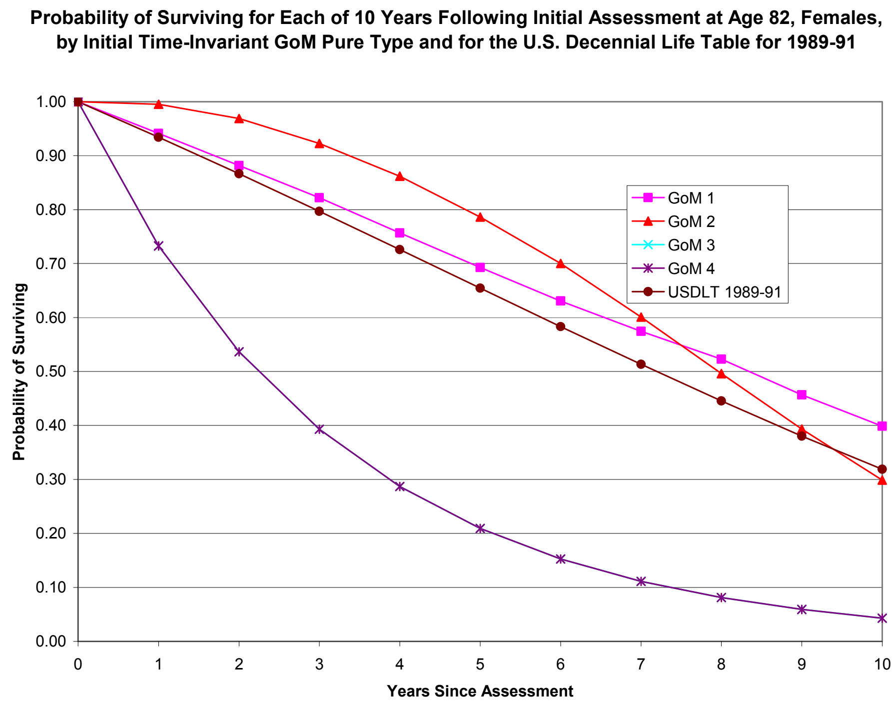
LEs of 2.5 years each for GoM 3 and 4 (the values were the same because an initial GoM 3 “converted” to GoM 4 prior to age 82).

Importantly, given that both the survival curve and its slope are used to price an offer (the former for premium costs, the latter for death benefit offsets), it is important to accurately estimate these quantities. Armed with this knowledge, an underwriter can price an offer to outbid the competition, win the policy, but not pay the full price indicated by the MAPS (Milliman) or other proprietary pricing model.

In Figure 1, an initial GoM 2 had approximately a 16% greater likelihood of surviving through years 1 through 5 until the lines converged at year 6, near the 6.2 year LE. On the other hand, an initial GoM 3 or 4 had approximately a 75% lower likelihood of surviving through years 1 through 5, with corresponding reductions for persons who had 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 for up to 76 questions on the web-based or telephonic interview (selected from 95 questions in the original analysis).

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 was less than the LEs of 9.2 and 8.2 years for GoM 1 and 2, but was substantially higher than the LEs of 3.2 years each for GoM 3 and 4 (as for males, the values were the same because an initial GoM 3 “converted” to GoM 4 prior to age 82).

Figure 2



For both sexes, the differences in the survival curves and their slopes illustrate the potential for substantial overpayment for a policy without this additional knowledge regarding the level and slope of the relevant survival curves.

It is interesting to note that when the 76 predictor variables in the LCC assessment questionnaire were ranked in order of importance by the ratio chi-squared/d.f., it was not until number 20 of 76 that one found a variable typically reported on medical records.[13] The best predictors involved ADL and IADL impairments and functional limitations.

The Longevity Cost Calculator web based model contained 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.

---

[13] See pages 49-50 in: E. Stallard. Trajectories of Morbidity, Disability, and Mortality Among the U.S. Elderly Population: Evidence from the 1984-1999 NLTCS. *North American Actuarial Journal* 11(3):16–53, 2007.

Table 1 and Figures 3–4 were originally presented in the *North American Actuarial Journal* paper. Table 1 compares the total and age-specific predicted probabilities of death within each year to the corresponding observed probabilities of death of individuals in the assessment population. The age-specific comparisons are shown by 5-year age groups, where age was determined at the start of each 1-year follow-up. Also shown are the total and age-specific predicted probabilities of death for each of the four GoM pure types represented in the model.

Table 1

Probabilities of Death within One Year in Four Pure Types GoM Models, 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 I	II	III	IV	Observed Probability	Predicted Probability
			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	0.032	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	0.003
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 age 65-69 in 1999.
Source: Stallard, NAAJ paper, 2007, table 8.

The differences between the observed and predicted age-specific death probabilities were very small and were statistically nonsignificant, with chi-squared values of 1.36 and 0.29, respectively, each with 6 d.f. (reference values were 12.59 and 16.81 at the conventional 5% and 1% significance levels). Figures 3–4 show the same data but add the corresponding death probabilities from the U.S. Decennial Life Table (USDLT) for 1989-91.

Figure 3

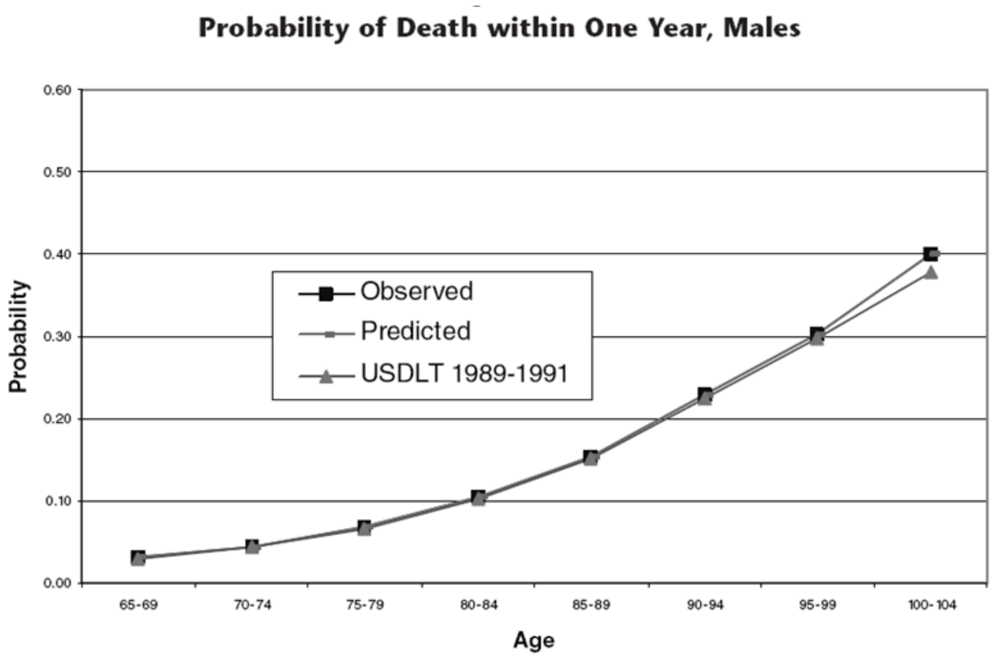
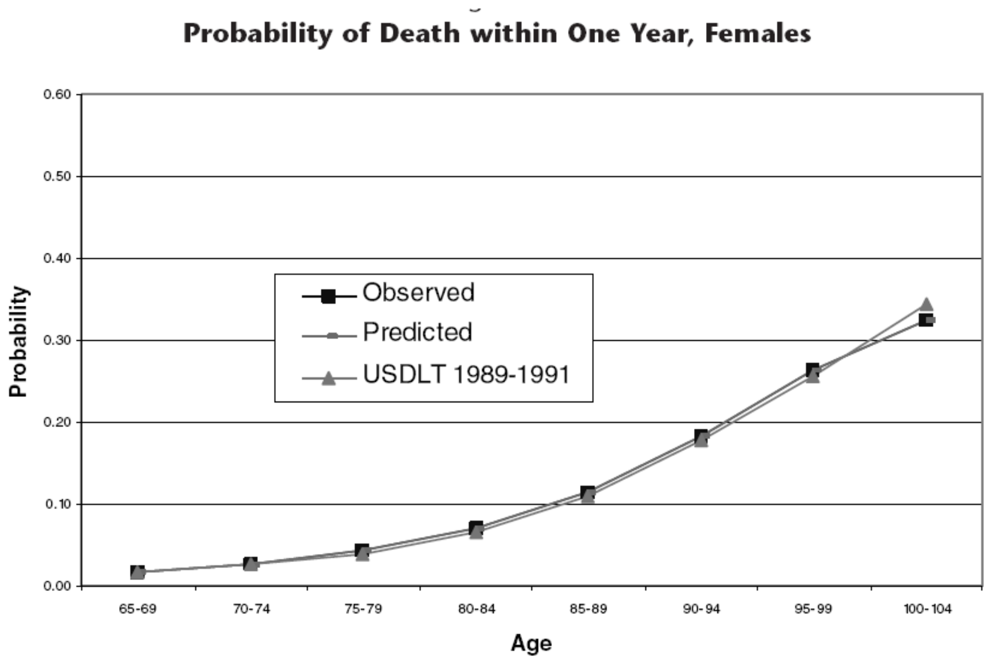


Figure 4



Additional Measures of Accuracy

Our consideration of additional measures of accuracy begins with an assessment of 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 extends the results of Longley-Cook (1962) to show the minimum number of expected events needed to meet various credibility standards for accuracy and reliability, with relative accuracy defined by the ratio: (observed no. of deaths − expected no. of deaths) / expected no. of deaths.

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		
ı/-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 *North American Actuarial Journal* analysis were based on 20,428 deaths (8,583 males and 11,845 females), indicating that the total rates were very stable but the stratifications by age and other variables may have been 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% probability levels, respectively, consistent with A.M. Best's recommendation that the collateral pool for life settlement portfolios consist of at least 300 lives[14].

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 *North American Actuarial Journal* 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 mortality-exposure data were 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.

---

[14] Emmanuel Modu, Life Settlement Securitization, Best's Rating Methodology, A.M. Best Company, March 24, 2008. www.ambest.com/debt/lifesettlement.pdf

Figure 5

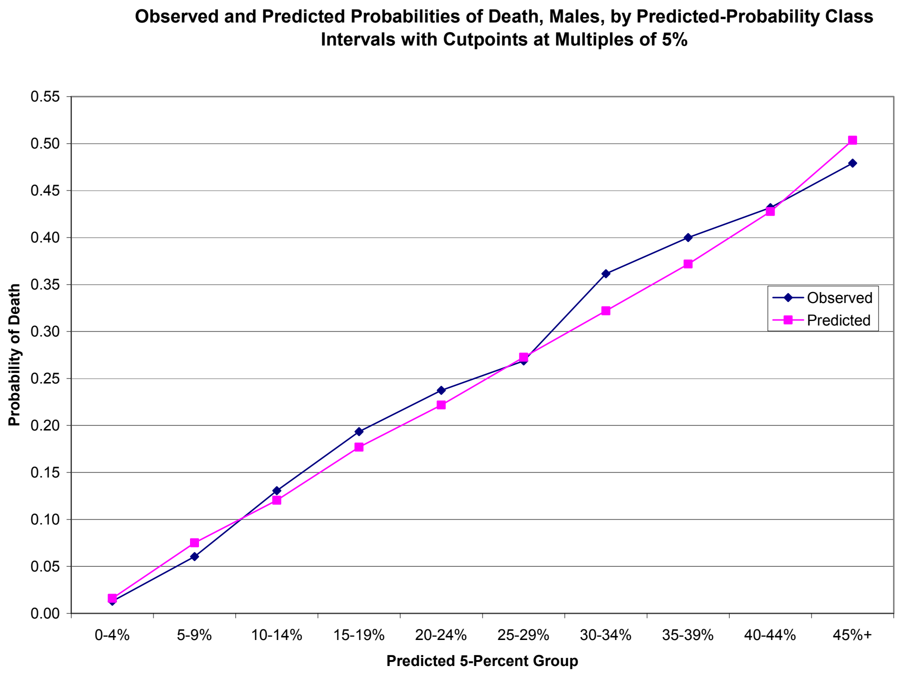
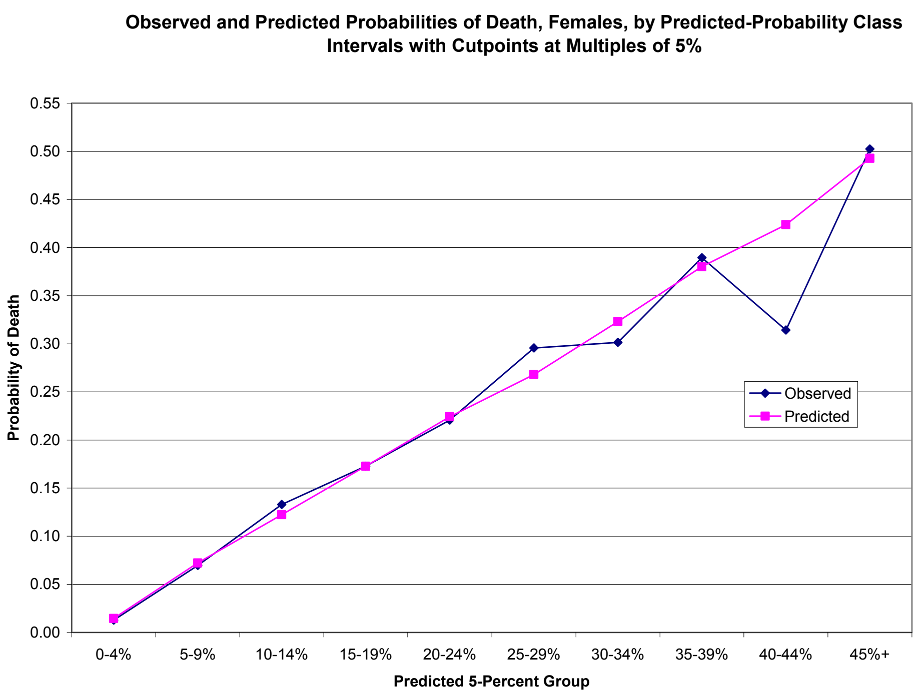


Figure 6



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-squared test of the deviation for that one point indicated that the difference was statistically nonsignificant, with a chi-squared value of 3.44 with 1 d.f. (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 produced 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 indicated that the models displayed in Figures 5 and 6 did *not* fit the data. This means that at least some of the deviations of the observed from predicted numbers of deaths were larger than expected by chance.
2. These were 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 were for the four lowest probability groups 0–19% and 30–34%; for females, the three significant deviations were 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 exceeded the 1,082 cutpoint for highly credible data in Table 2; all eight exceeded the minimum level of 300 lives recommended by A.M. Best for life settlement collateral pool sizes.

We concluded that there were factors operating in these data that were not represented in our model. This should not be surprising given that the model used four GoM scores to summarize data on the original set of 95 questions concerning medical conditions, activities of daily living (ADLs), cognitive and behavioral impairments. Moreover, there may be other influential factors not included in the set of 95 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 suggested that the deviations from the models were not predictable, at least at the higher probability levels which are of greatest concern to life settlement providers.
2. Nonrandom deviations in the GoM models can be tolerated if they are sufficiently small, relative to the errors that would occur in the absence of the GoM models.

To quantify the size of the nonrandom deviations, we applied linear regression analysis with the observed probabilities regressed on the predicted probabilities, which are 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 was 0.964 which may be interpreted as a measure of the accuracy of the GoM models: 96.4% of the variance of the observed probabilities was accounted for by the expected probabilities produced by the GoM models. The remaining 3.6% of the variance constituted a tolerable level of nonrandom deviation in the GoM 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 was 0.991 which may be interpreted as an alternative measure of the accuracy of the GoM models: 99.1% of the variance of the logarithm of the observed probabilities was accounted for by the logarithm of the expected probabilities produced by the GoM models. The remaining 0.9% of the variance constituted an even more tolerable level of nonrandom deviation in the GoM models.

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

The first question was whether the GoM scores added 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 was the simplest model. It assumed that the sex-specific mortality probability was constant over age and GoM scores. Model 2 assumed that the sex-specific mortality probabilities increased over age but not over GoM scores, following the observed values displayed in Figures 3–4. Model 3 assumed that the sex-specific mortality probabilities increased over GoM scores but not over age, following the values displayed in the Totals row of Table 7 of the *North American Actuarial Journal* paper. Model 4 assumed that the sex-specific mortality probabilities increased over GoM scores and over age, following the values displayed in the age-specific rows of Table 7 of the *North American Actuarial Journal* 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 was the one that had the minimum value of AIC or BIC (indicated with boldface font in Table 5); the relative performance of each model was 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 were regarded as strong evidence in support of the model with the lower AIC or BIC value.

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

To determine whether the GoM scores added significant information beyond the information already available using the age-specific mortality probabilities, we needed 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 exceeded the reference values by a factor of 746–874, indicating that the additional information provided by the GoM scores was huge.

Comparison of Models 2 and 3 provided additional confirmation of the power of the GoM model. The ΔAIC and ΔBIC for Model 2 were 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 were 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 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 was whether the excellent calibration displayed in Figures 5 and 6 continued when the predictions were stratified by age-groups. The results of these stratifications 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 combined 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 were ordered by increasing predicted probabilities, and within each probability group, by increasing age.

Figure 7

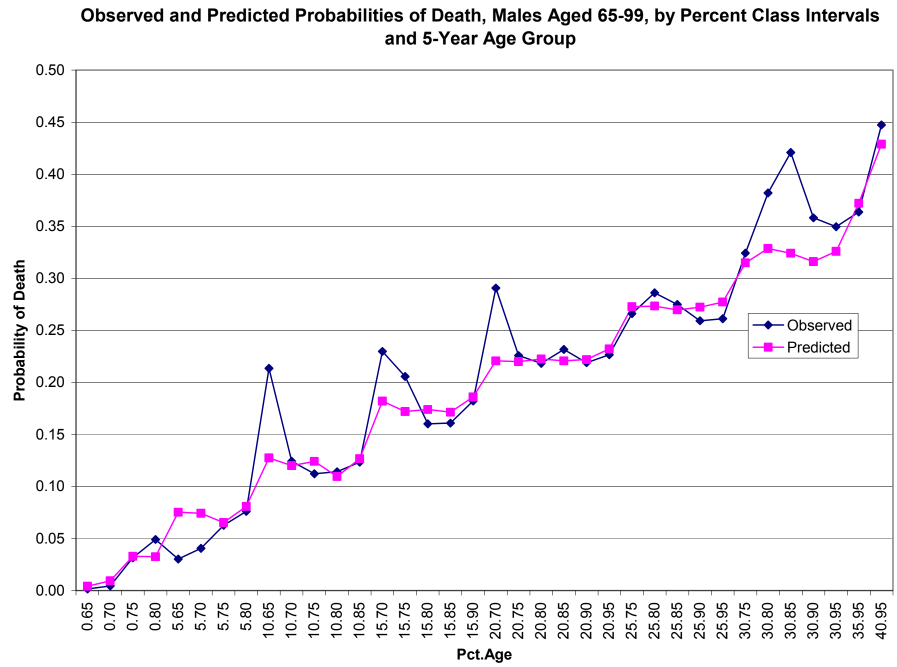
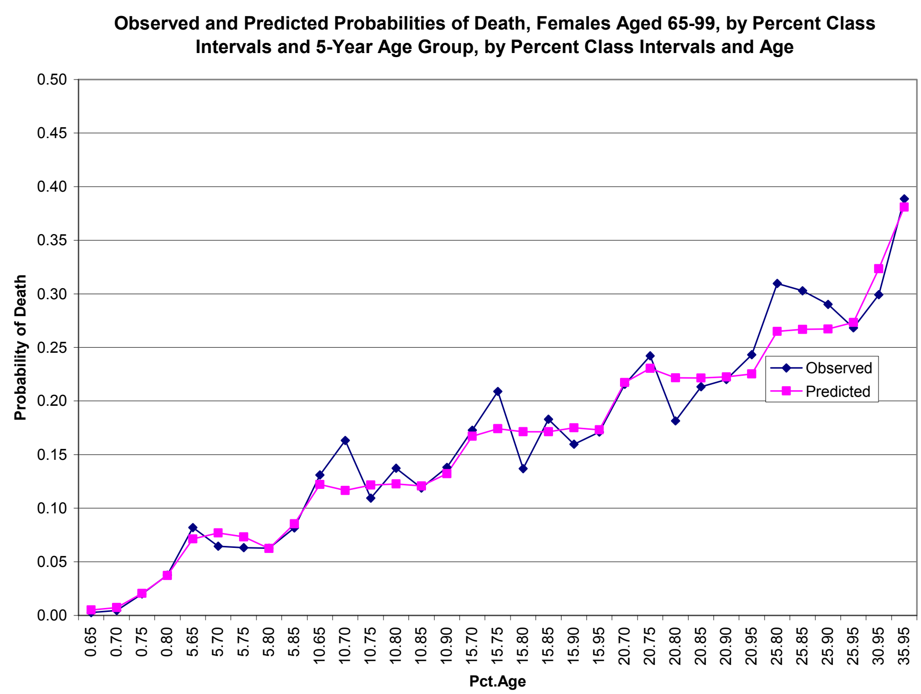


Figure 8



Examination of the results for males and females in Figures 7 and 8 showed that the largest deviations were 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 were not replicated for females.

One possible explanation for the male result involved the design of the NLTCS in which persons who did not have ADL or IADL impairments at the time of the survey did not receive the in-person assessment; such persons "screened 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-related questions in the original model (or the 76 questions in the LCC implementation).

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 was thus in the range 95–97%, depending on the form of the regression.

<u>Ongoing Model Validation and Recalibration</u>

LSF has been conducting retrospective validation tests using data sets with hundreds of assessments collected over the last seven years in assisted living communities. The questions in these assessments have been mapped to the 76 questions in the LCC.

Complementing these retrospective validation tests, LSF is in the process of collecting new prospective data, for use in additional validation activities, from the answers to the 76 questions on the LCC applications provided by cooperating life settlement brokers, providers and funders.

Prospective data collection by life settlement providers will allow comparisons of the LE and survival predictions based on the LCC model to the corresponding predictions of commercial LE providers, and comparisons to the actual outcomes, based on the tracking of each insured until death. As sufficiently credible experience is accumulated, the parameters of the model will be recalibrated to better capture unmeasured aspects of insured survival that are unique to life settlement participants.

The 32,389 individuals observed in consecutive assessments over the term of the original NLTCS calibration data were a random sample of Medicare enrollees who participated in the study. Assessments began either in 1982 when the NLTCS started, or at a later date when the participants were aged 65–69. Because life insurance data were not collected, we do not know how many of these seniors either had life insurance in-force or may not have been medically insurable at age 65. What we do know from Census Bureau data is that more than 80% of seniors age 65 are homeowners. Moreover, there is nearly a one-to-one correlation between homeownership and the purchase of life insurance.

Thus it was reasonable to expect that the general population experience represented in the NLTCS would be reasonably representative of the experience of life insurance policy holders, given that the answers to most or all of the 76 health-related questions were known. This latter conditioning is important because it controls for most of the health-related selection biases that may distinguish the insured from the non-insured population.

Application of the same model to the life settlement population must consider that life settlement participants may exhibit further health-related selection biases that may distinguish them from the non-settlement insured population, given that the former group is a small fraction of the latter. Again, it is reasonable to expect that most of these biases will be controlled if the answers to most or all of the 76 health-related questions are known.

The purpose of the prospective data collection is to assess the size and direction of biases that may exist, and when necessary to recalibrate the model parameters as needed to remove them. Thus, periodic assessment and recalibration are important parts of the proposed methodology.

VALUATION METHODOLOGY FOR SETTLEMENT PORTFOLIO MANAGERS

A settlement portfolio's management is required to prepare fair valuation estimates and have these audited by independent auditors. The standards require that the valuation be determined using methodologies to corroborate and reconcile the results. The LCC tool is a low cost, peer reviewed, and published methodology that can be incorporated in a defined process of reconciling the life expectancies as part of the portfolio fair valuation process.

The proposed methodology for annually establishing fair value is compliant with ASC §820.10 and is straightforward and drawn from ISA §540 / AU §328 and applies accepted Bayesian statistical methodology to arrive at a weighted average fair value per policy and thus the sum will be a fair value of the portfolio. Annually, the difference between the purchase value and the newly established value will be clearly documented, transparent and available to be consistently used year after year in fair valuation. Policy and portfolio cash flows can then be stress tested based upon this reconciliation of life expectancy to arrive at the portfolio's current fair value.

Proposed steps at underwriting and portfolio origination:

1. Each settlement policy at the point of underwriting will have three (3) commercial LE's, ideally from the same three commercial LE providers for the entire portfolio. LE providers will each provide a table of calculated survival functions for each insured case.
2. The settlement application will contain the LCC assessment questions or a telephonic assessment will be completed for automated LCC scoring.
3. Each policy will be priced independently for each commercial LE and the LCC LE using the same probabilistic pricing model. Typically, the funder will require that the offer be priced based upon commercial LE's only, no matter how many LE's may have been discarded at the time of pricing.
4. If the offer is accepted, the policy and its four (4) related LE data will become part of the portfolio valuation data base.
5. Initially at the time of portfolio formulation, equal weighting (1/4 to each) will be assigned to the four life expectancy providers (i.e., the three commercial LE's plus the LCC LE).

Proposed steps at quarterly and annual fair valuation:

1. Compare the actual to expected mortality results for each of the four LE providers over the examination period. The expected survival counts will be generated by summing the survival functions (S_x) by LE provider for that period for each policy in the portfolio for which an LE value was provided. The change from one period to the next in the sum of S_x will constitute the expected number of deaths for the period as predicted by each of the four LE providers.

Identical Analysis Performed for Each LE Provider

	EXPECTED SURVIVAL AND DEATHS						ACTUAL SURVIVAL AND DEATHS					
	Expected Survival			Expected Deaths			Actual Survival			Actual Deaths		
Portfolio Member Code	Year-End 1	Year-End 2	Year-End 3	Year-End 1	Year-End 2	Year-End 3	Year-End 1	Year-End 2	Year-End 3	Year-End 1	Year-End 2	Year-End 3
0001-F	0.9577	0.9095	0.8564	0.0423	0.0905	0.1436	1.0000			0.0000		
0002-F	0.7824	0.6428	0.4987	0.2176	0.3572	0.5013	1.0000			0.0000		
0003-F	0.8974	0.8021	0.7895	0.1026	0.1979	0.2105	1.0000			0.0000		
0004-F	0.8547	0.7952	0.6587	0.1453	0.2048	0.3413	1.0000			0.0000		
1000-F	0.9211	0.9033	0.8524	0.0789	0.0967	0.1476	1.0000			0.0000		
Totals	883	811	731	117	189	269	852			148	Actual	
Less Expected Deaths										117	Expected	

2. Compare the actual number of deaths for the period against the expected number to ascertain the accuracy of each LE provider relative to their prediction.
3. The portfolio valuation for each of the four LE providers will define the bounds of the valuation. For example, assume the independently determined portfolio values are $255 million, $274 million, $316 million and $329 million[15]. Absent better precision the portfolio has a value between $255 million and $329 million.

At Portfolio Origination

LE Provider	NPV of Cash Flows Value	Assumed Probability	Probable Weighted Cash Flows	Probable NPV of Cash Flows
1	$329 Million	25%	$82.3 Million +	
2	$316 Million	25%	$79.0 Million +	
3	$274 Million	25%	$68.5 Million +	
4	$255 Million	25%	$63.8 Million =	$293.6 Million

4. Use Bayesian analysis as detailed by Kass and Raftery[16] to determine the weighted average total value of the portfolio.

Results Oriented Weighted Valuation after Portfolio Year-One

LE Provider	NPV of Cash Flows Value	Assumed Probability	Probable Weighted Cash Flows	Probable NPV of Cash Flows
1	$329 Million	42%	$138.2 Million +	
2	$316 Million	16%	$50.6 Million +	
3	$274 Million	28%	$76.7 Million +	
4	$255 Million	14%	$35.7 Million =	$301.2 Million

5. This will allow the actual versus expected values for each policy's LE's to be evaluated and the LE providers to be ranked accordingly. The LE evaluations can be conducted via standard Bayesian methods or, more simply, by using Bayesian Information Criterion (BIC) measures of goodness of fit of the actual-to-expected probabilities of death computed separately for each LE model using the same pool of policies. Kass and Raftery[17] showed how BIC values can be

[15] Assumes a portfolio of $500 million of face amount with future net cash flows to maturity discounted at 2% to an NPV.
[16] Robert E. Kass and Adrian E. Raftery. Bayes Factors. *Journal of the American Statistical Association* 90(430):773-795, 1995.
[17] 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

used (1) to rank the various LE models and (2) to generate optimal weighted averages of the outputs of the various LE models, where each of the four weights is interpretable as the Bayesian posterior probability that the corresponding LE model is correct (assuming that one of them is correct). Weighted averaging can then be applied to each individual policy, and to the aggregate of all policies, in annually revaluing each portfolio. The weights can be updated each year (more frequently if the portfolio is large) as additional information on the actual number of deaths in that year becomes available. Over time, this will give greater weight to the better performing models. Disclosure of the chosen methodology at portfolio formulation and annual review will provide investors transparency into these longevity valued asset transactions.

6. Present the weighted average portfolio value as well as the individual portfolio values, and their associated weights, derived by using the LE values and survival curves from each of the four LE providers.

PREPARING FOR THE INDEPENDENT AUDIT

AU §328 directs independent auditors to first understand the process of developing the fair value estimates and the controls instituted by the entity to ensure the completeness, accuracy, and consistency of the methodology and data used in the computations. It will often be necessary for the auditor to verify that the data in the models are accurate and to evaluate the reasonableness of the assumptions used in the modeling.

The approach outlined in this paper will help provide documentation and support helpful to the auditor in assessing that a “reasonable basis” exists for the fair valuations. Another option open to auditors is to apply a different methodology to relevant fair value data and compare the results with the entity estimates. However, the estimates of the entity are generally only open to challenge when they fall outside of a “reasonable range” of outcomes as assessed by the independent auditor.

Because of the specialty nature of the application of fair value concepts to insurance valuations, the auditor is likely to engage a specialist to review the approach and computations underlying the entity estimates and to identify benchmark data that might be used to assess the reasonableness of the assumptions used in developing the estimate.

Recognizing these requirements, the entity can minimize its audit costs by creating clear and transparent documentation of the development of its estimates, establishing oversight and internal controls to ensure the quality of its estimates on a quarterly and

using various approaches to measuring goodness of fit, including chi-squared statistics and regression-based R^2-statistics.

annual basis, and having internal resources available to answer questions and explain the process during the audit.

Because consistency is an important accounting concept, fair value methodologies should be applied consistently from year-to-year, and thus should be chosen wisely. Departures in methodologies may indicate that an accounting change has occurred; the effects of the change should be quantified in financial reporting. Nonetheless, as more and better proven methodologies arise, they should be considered for possible implementation.

EXISTING ACCOUNTING FRAMEWORK

The framework exists today to create uniformity in life settlement policy underwriting and related disclosure by applying existing GAAP at the time new policies are originated, existing policies are pooled for securitization, and annually at revaluation. These policy portfolios, whether rated or unrated, will require disclosure at the time they are pooled and sold, and subsequently when the pools are individually valued and audited because they are held for investment by issuers. For level-three assets where value is dependent upon a future unobserved outcome (the insured's death) GAAP, Accounting Standards Codification (ASC) 820.10.05 through 820.10.65 (previously FASB 157), requires the use of the best information available in the corroboration of the valuation methodology. The outcome of these methodologies must then be reconciled and disclosed. Specific relevant auditing standards include AU §328 (SAS101), *Auditing Fair Value Measurements and Disclosures,* (June 2003).

AU §328.4 says "Management is responsible for making the fair value measurements and disclosures included in the financial statements. As part of fulfilling its responsibility, management needs to establish an accounting and financial reporting process for determining the fair value measurements and disclosures, select appropriate valuation methods, identify and adequately support any significant assumptions used, prepare the valuation, and ensure that the presentation and disclosure of the fair value measurements are in accordance with GAAP." We believe the straightforward rules-based methodology above supports these objectives.

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[18]. 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

[18] See AU §329, *Analytical Procedures.*

own assumptions to make a comparison with management's assumptions. The auditor uses that understanding to ensure that his or her independent estimates take into consideration all significant variables and to evaluate any significant differences from management's estimates. The auditor also should test the data used to develop the fair value measurements and disclosures as discussed in paragraph .39." If a well documented and controlled process, supported by published research or established as an industry practice, is used to develop the entity fair value estimates, the auditor generally will "audit" the entity process, controls and calculations rather than run alternative models that may need to be reconciled to the entity estimates[19].

Clear roadmaps that link to the data sources used and documented support for the key model assumptions will aid auditors in evaluating whether management has a "reasonable basis" for its estimates of fair value.

CONCLUSION

The proposed fair value methodology is credible, compliant with the accounting and auditing framework and, most importantly, doable by management. Management is responsible for implementing a supportable assumptions-based valuation methodology that is transparent and controlled. Provided they do this they can present a completed valuation to the independent auditors to critique and avoid costly additional modeling. The proposed methodology of valuation should be consistently applied year-after-year. As the portfolio valuation improves it adds income. Conversely, if the portfolio value were to decline in the future, consistently applying the methodology would identify the change in direction on a timely basis, and avoid lags in portfolio write-downs. This would mitigate the natural tendency to defer losses, pending the ugly last-minute, one-time write-down of asset values such as experienced recently, and expected to continue for the foreseeable future, in other collateralized asset classes.

[19] The testing requirements for fair value estimates are enumerated in paragraphs 23 to 42 of SAS 101 (AU 328).