SOCIETY OF ACTUARIES

Article from:

Product Matters!

June 2012 – Issue 83

Individualized Mortality Projection and Product Pricing **with Laboratory and Physical Measurement Data**

By Brian Lanzrath, Jim Palmier and Ammon Dixon

In any underwritten life insurance product, establishing the relationship between underwriting standards and expected mortality experience is the very core of the product design process. Risk stratification criteria, though, tend to be relatively unsystematic, and often consist of little more than simple variations on clinical guidelines as applied to certain well-studied biomarkers (particularly serum cholesterol and the broader lipid panel), supplemented by laboratory-supplied "reference ranges" (usually the middle 95 percent of an analyte's observed distribution) for tests where clinical significance is less well-established. Following an often ad hoc definition of underwriting classes, empirical mortality projections are developed from historical data, or established industry expectations. Credits and debits may sometimes be attached to individual test results on the basis of published clinical studies, virtually none of which will control for the full suite of laboratory and physical data available during insurance underwriting. The final result, of course, is the familiar three to five rate classes (plus table ratings) of most existing preferred underwriting systems.

Recent applications of modern data analytics methods to the extremely large (more than 8 million complete records since 2001) laboratory and physical measurement database of one insurance testing laboratory (ExamOne) have enabled a fully individualized approach to mortality projection—one which could in principle assign a unique pure premium to any given applicant. As might be expected from so granular an appraisement of mortality risk, this method identifies substantial numbers of significantly mispriced policies; including preferred-qualified individuals who represent a greater claims risk than most standard policies, and standard policies that can be confidently projected to perform at least as well as the majority of more favorably underwritten cases.

Analytical Methodology and Outputs

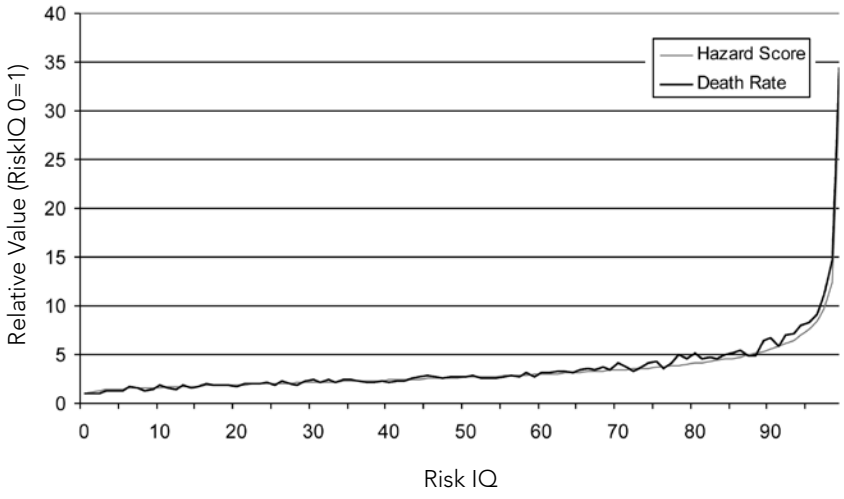
The development process for the mortality risk assessment model used in this study (Risk IQ) has been detailed elsewhere.1 Briefly, it is a multivariate proportional hazards regression model developed from laboratory and physical measurements, as matched to the Social Security death master file (SSDMF). There are two calculated final values: a hazard score, which represents the relative mortality risk of an applicant as normalized to a median value of 100 for the relevant peer group (defined as curtate age, sex, and smoking (cotinine) status), and the Risk IQ, which is simply a percentile ranking of the hazard score, again relative to age, gender, and smoking status. By definition, Risk IQ scores are bounded by 0 and 99, but hazard scores can be arbitrarily high (in very rare cases exceeding 100,000); hazard scores below 25 are uncommon. The hazard score, but not the Risk IQ, is directly proportional to mortality risk; in males 40 – 49, for instance, the mean hazard scores for Risk IQs of 25, 50, and 99 are 76.2, 100.6, and 1359.7, respectively.

Distribution of Mortality Risk in Applicant Populations

A direct comparison of Risk IQ, raw death rates, and hazard scores [Fig. 1] may be the simplest illustration of the risk segregation attained by this approach.

In Figure 1, death rates are not normalized by age or sex, as Risk IQ is itself demographically normalized. In all percentiles, the mean age is 40.65, 54.34 percent of applicants are male, and 9.3 percent are cotinine-positive.

FIGURE 1
Hazard Score and Raw Death Rates by Risk IQ



CONTINUED ON **PAGE 22**

Alternately, we can plot the fraction of all recorded deaths attributable to applicants in a given percentile range [Fig. 2], which highlights the disproportionate concentration of risk (as represented by actual deaths) among the upper quantiles of the ranking system.

Over the study period, fully 10.4 percent of all recorded deaths occurred among the 1 percent of applicants assigned scores of 99, and 31 percent among applicants in the highest Risk IQ decile. Mortality was commensurately represented among lower score ranges, with only 4.6 percent occurring in the zero to nine Risk IQ decile. Again, these results are in excellent accordance with the hazard score projections (the mean all-applicant hazard score for Risk IQ-99 individuals is 1172).

FIGURE 2

All ExamOne Applicants and Deaths by Risk IQ: Non-smoking Applicants, 2001-2008

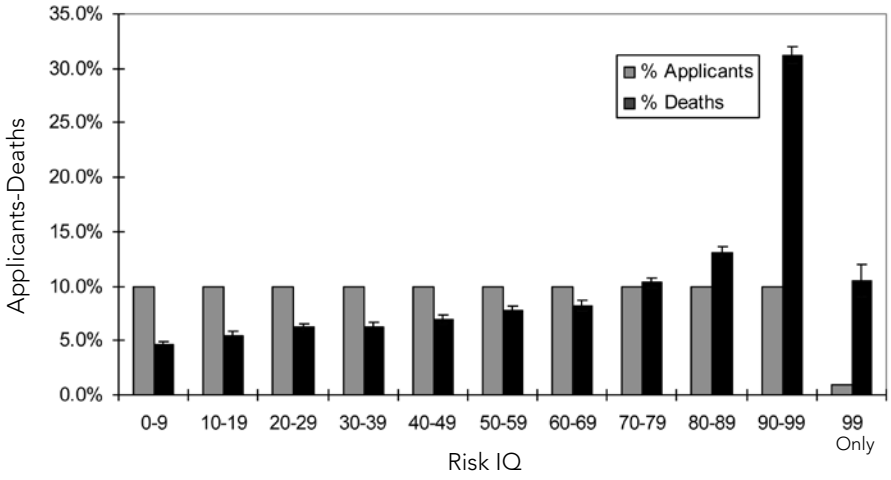
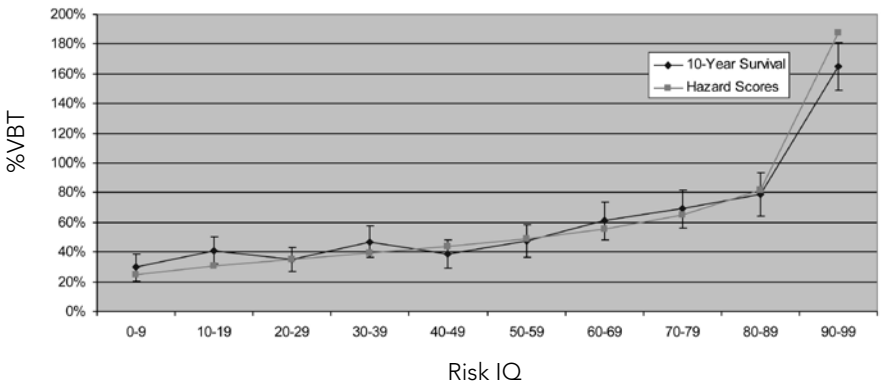


FIGURE 3

10-Year Death Rates: 45-Year Old Male Non-Smokers



Observed Absolute Death Rates by Risk IQ

When stratified by model demographic groups, sample sizes are too small for statistically significant empirical percentile-by-percentile estimates of death rates (as in Fig. 1), but decile-level aggregation remains feasible [Fig. 3 for males 40 – 49].

Death rates and confidence intervals were constructed according to standard life-table methodology using SAS/STAT. For convenience, death rates have been expressed as a percentage of 2001 VBT Select values. It is important to note that the absolute mortality estimates provided in this article are derived from SSDMF data, which is known to be an incomplete record of U.S. deaths. Upward adjustments of as much as 10 percent (assuming a ~90 percent completeness rate for the SSDMF) would be required for product-level implementation. We have little reason to suspect that Social Security record quality would vary with prior mortality risk (particularly among the life insurance applicant population), so relative risk levels should be highly consistent. The precise shape and magnitude of the Risk IQ/mortality curve vary somewhat by age and sex [Figs. 4 & 5 for results from select demographic groups].

A striking characteristic of all of these charts is the very favorable mortality experience among the lowest three Risk IQ quartiles (scores less than ~75). If, as is the case in many products, a preferred applicant is defined as one who's projected death rates fall below 60 – 70 percent of the 2001 Select VBT, then in many demographic groups as much as 75 percent of the applicant pool may qualify, according to this appraisal. Given that, in existing products, it is rare for more than 30 – 35 percent of policies to be issued at the best rate class, the potential for a substantial, actuarially justifiable, expansion of these classes is obvious. In general, low-Risk IQ applicants currently excluded from preferred pools (the hidden healthy) represent 25 – 40 percent of underwritten cases. The most common grounds for relatively unfavorable decisions in these cases are mild

obesity or elevated cholesterol—two conditions that, though undoubtedly correlated with premature claims in a univariate sense, have few to no marginal mortality implications when present in isolation (i.e., in the absence of common co-morbidities such as hypertension or diabetes).

The existence of an identifiably charged population naturally implies the existence of a subsidized high-risk group currently granted relatively favorable rates. This phenomena (cryptic risk) is in fact observed; defined as applicants with scores of 75 or above who are admitted to preferred pools, it represents ~4.5 percent of the total applicant population. In the paradigmatic cryptic risk case, all underwriting variables will lie within established preferred ranges, but several of these values will fall near the extreme upper or lower boundaries; the aggregate effect of several such high- or low-normal results can easily surpass that of a single more overtly abnormal value. As a group, cryptic risk applicants die at approximately twice the rate for which a preferred or preferred-best class is priced, with obvious financial implications. Under reasonable assumptions regarding discount and lapsation rates, the present value of claims in a $300,000 20-year term policy written on a 45-year-old male non-smoker subject to a preferred-level life table is $3,300; if this applicant were a representative cryptic risk case, the actuarial present value would rise to $7,400, a $4,100 expected loss on a present value basis.

As is evident from figures 1 and 2, mortality risk in Risk IQ 99 applicants is qualitatively different from that of lower scores; in most demographic groups, deaths are more than twice as common in 99s than even among 98s. Studies conducted with carriers in fully underwritten applicants have confirmed that, as might be expected, these individuals are substantially more likely to be declined in the course of conventional underwriting. The final decline rate, however, has not exceeded 50 percent in any study, and analysis of actual claims experiences reveals that issued 99s die at virtually the same disproportionate rates as those who are excluded (in approximate terms, Risk IQ-99 applicants represent 1 percent of the typical carrier's applicants, 0.5 percent of its issued policies, and 5 percent of its paid claims). [Fig. 6 summarizes one carrier study]

"The most common grounds for relatively unfavorable decisions…are mild obesity or elevated cholesterol …"

FIGURE 4
10-Year Death Rates: 35-Year Old Female Non-Smokers

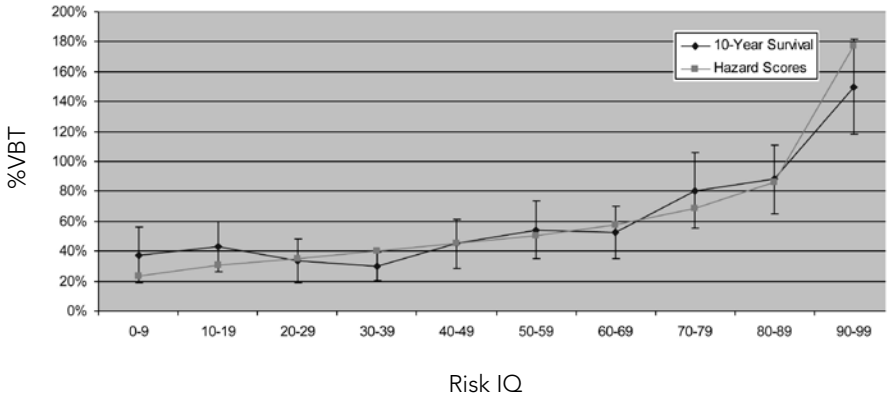
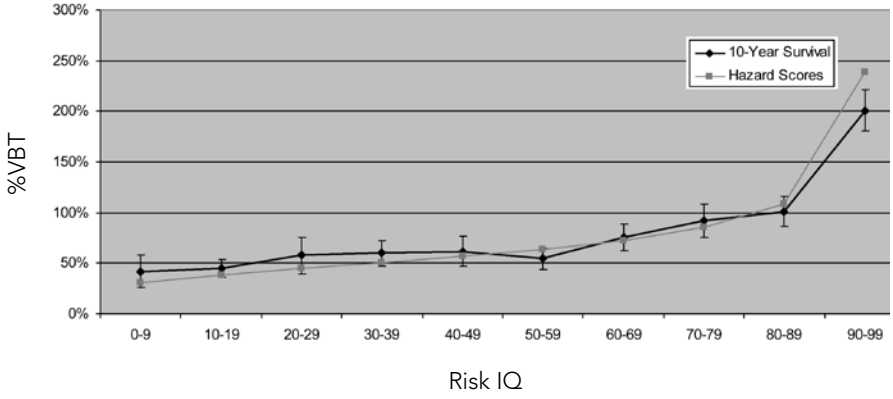


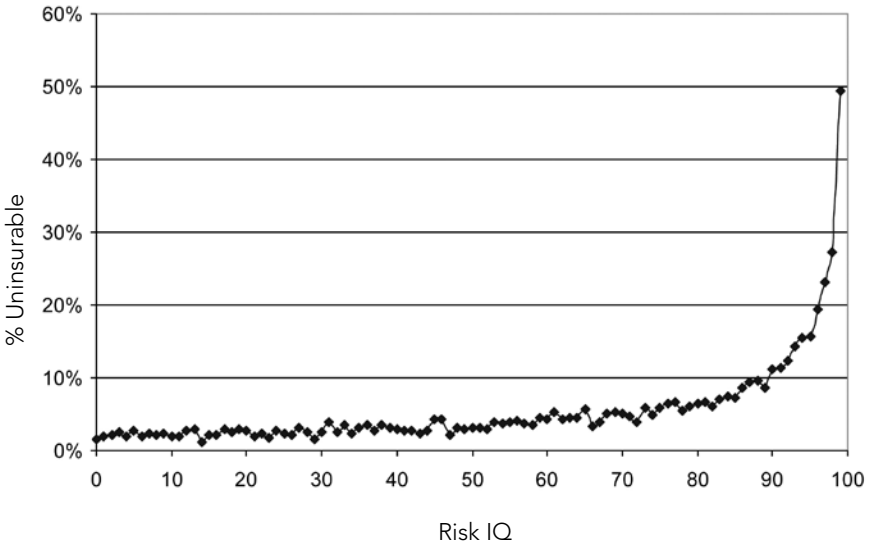
FIGURE 5
10-Year Death Rates: 65-Year Old Male Non-Smokers



CONTINUED ON **PAGE 24**

Although, as mentioned above, individual Risk IQ scores cannot be associated with specific death rates on a strictly empirical basis, decile-level results can be interpolated (using hazard scores are in principle directly proportional to risk) to produce mortality tables specific not only to each Risk IQ score, but potentially to each individual applicant. Fig. 7 provides expected life-table multiples (as percentages of the 2001 Select VBT) for selected scores and demographic groups.

FIGURE 6
'Uninsurability' Rates by Risk IQ: 2009-2010



Implications for Product Design

As with other innovations in risk assessment, the effects of individualized mortality modeling will likely emerge incrementally. Early adopters of this system have devoted most of their efforts to date on the identification and (in most cases) decline of the half of Risk IQ-99 cases not detected by conventional underwriting. In combination with cryptic risk detection, this constitutes the most immediately quantifiable value proposition of scoring algorithms. A more fundamental transformation of product design will require a willingness to offer preferred-level premiums to hidden healthy applicants, which for many carriers will require close coordination with reinsurers. The benefits of this latter approach are highly dependent upon the elasticity of carrier market share relative to preferred placement rates, which creates particularly strong incentives for adoption in the brokered market.

It can be expected that adverse selection will become a progressively more powerful driver of adoption as market penetration progresses (by the end of the second or third quarter of 2012, it is expected that roughly 25 per-

FIGURE 7

Gender	Age	40-49			60-79		
	Risk IQ	Hazard Score	%VBT	UW Class	Hazard Score	%VBT	UW Class
Female	0	42.2	17.8	Preferred or Better	32.7	21.4	Preferred or Better
	50	100.6	42.5	Preferred or Better	101.0	65.9	Preferred or Better
	80	157.2	66.5	Preferred or Better	165.0	107.7	Standard
	90	214.4	90.7	Standard	230.5	150.5	Table 2
	95	296.3	125.4	Table 1	329.4	215.1	Table 5
	99	1332.1	563.6	Table 19	1427.6	932.0	Table 33
Male	0	40.5	17.5	Preferred or Better	35.6	20.7	Preferred or Better
	50	100.6	43.4	Preferred or Better	100.9	58.6	Preferred or Better
	80	156.8	67.6	Preferred or Better	159.8	92.8	Standard
	90	213.3	91.9	Standard	221.0	128.3	Table 1
	95	295.4	127.3	Table 1	309.7	179.8	Table 3
	99	1359.7	586.2	Table 19	1163.9	675.8	Table 23

cent of laboratory panels conducted at North America's largest insurance testing lab will be accompanied by mortality scores). Many cryptic risk applicants denied admission to the preferred pools of early adopters will seek out coverage among non-adopters, and hidden healthy applicants (who constitute the least claim-prone half of existing standard pools) will tend to migrate to carriers able to recognize their comparatively low risk.

In the intermediate term, individual mortality risk scoring should greatly facilitate the transition to straight-through processing, and mitigate the need for the additional requirements (such as attending physician statements, 80 percent of which are requested for applicants generating preferred-level risk scores), which do most to slow and complicate the policy issue process. In one study with a large carrier, Risk IQ proved to be a more accurate predictor of short-duration (two to three year) claims than the actual human-made underwriting classification, despite drawing upon a more restricted set of raw data. In the future, it is possible that rate classes may become obsolete, replaced by applicant-specific premiums calculated from the unique mortality probability vector of individual insureds (as has been the strong tendency in the property and causality fields). □

END NOTES

[1] Lanzrath, Brian, et al. "A Comprehensive Multivariate Approach to the Stratification of Applicant-Level All-Cause Mortality Risk." On the Risk. Vol. 27, No. 1 (March 2011): 56-61.

Brian Lanzrath is Actuarial Biostatistician at ExamOne.

Jim Palmier, MD, MPH is Medical Director at ExamOne

Ammon Dixon, MAOA is Head of Risk Analytics at ExamOne. He can be reached at ammon.j.dixon@examone.com.