Long Term Care Insurance Underwriting
Building a Risk Scoring Model for
John Hancock Insurance & Financial Services

_____   _____
Joanna B. Kluza   Joseph F. Kraynak

_____   _____
William J. Treese   Jaris B. Wicklund

Date: April 26, 2007

Approved by:

_____   _____

Professor Jonathan P. Abraham   Professor Arthur C. Heinricher

Abstract

The project goal was to test a previously designed adaptive risk scoring model for long term care insurance underwriting using actual applicant data from John Hancock. A data filtering method for removing applicants who should not be used to train the model was developed. The model provides accurate risk class assignment, based solely on the medical conditions, when trained on the filtered data. In addition, the model identifies errors in the risk points assigned to individual medical conditions.

Table of Contents

List of Figures

List of Charts

1. Introduction

This paper develops an adaptive risk scoring model to aid in underwriting for long term care insurance (LTCI). Two previous projects at WPI (see (Arsenault 2006) and (Chen 2006)) have developed methods for assigning risk scores to LTCI applicants and this project will use the algorithm developed by the Research Experience for Undergraduates team in summer 2006 (Chen 2006). The goal is to test the models using real applicants provided by John Hancock.

LTCI is a relatively new product and as such, it is currently difficult for any company to identify the risk associate with an applicant. John Hancock's LTCI team would like to be the first with a tool to assist underwriters in their decision making process. This tool would provide John Hancock's LTCI underwriters with a consistent, objective measure of an applicant's risk and John Hancock would have the ability to develop new products for specific categories of applicant. The model developed in this project is the next step in the path to accomplishing these goals.

1.1 Objectives

The primary objective of this project is to test a previously developed risk scoring model for underwriting LTCI applicants using real data provided by John Hancock. Previous groups had designed risk scoring models for underwriting applicants; this project will test the previous models and refine them to handle real applicant data.

The main contributions of this project are

- the mapping of the condition-severity matrix developed by Avon Long Term Care Leaders to the medical condition indicators used by John Hancock and

- filtering methods used to remove applicants who are either obviously uninsurable or should be underwritten based on factors other than medical conditions.

This dataset will be used to compute model coefficients and replicate underwriting decisions.

2. Background

This section will provide background on Long Term Care Insurance (LTCI) as well project work completed previously at WPI.   This is the first project that works with real applicant data from John Hancock's LTCI division.

2. 1 Long Term Care Insurance

Long term care (LTC) is defined as "services that are needed when you can no longer perform normal activities of daily living that healthy, fully-functioning people do on their own" (John Hancock, 2007). LTC can be provided in a number of places including nursing homes, the patient's residence, or in special communities, and each service can range from assistance in daily tasks to full-time nursing care. LTCI provides a way for patients or families that would otherwise be unable to afford the high cost of LTC to have the appropriate care.

The majority of those that purchase LTCI in the United States are individuals earning between $32,500 and $100,000 annually (American, 2007). Wealthier people are able to afford the costs of care, and poorer individuals cannot afford the premiums associated with LTCI. LTCI is becoming more important not only because modern medicine is able to provide a quality of life never before seen in chronically ill patients, but also because of changing family life. Fewer children are able to take care of their parents who require help accomplishing tasks that are taken for granted by healthy individuals.

For most LTCI policies, benefits are triggered when two or more of the following *Activities of Daily Living* (ADL) can no longer be performed: *bathing*, *toileting*, *dressing*, *continence*, *transferring, eating, ambulating*, and *mobility* (Long Term Care, 2007). Continence is the ability to refrain from using the toilet until an appropriate time. Transferring is moving

from a bed to a chair and vice-versa. Ambulating is the ability to move from place to place (typically walking) and the others are self-explanatory.

LTCI is a relatively new product, only being offered for roughly the last forty years as a supplement to Medicare (History, 2007). The market for LTCI is quite large, as expenditures on LTC services in the United States were approximately $135 billion in 2004 and this number is expected to grow significantly as the overall population ages; however, only 10% of elderly individuals currently have private LTCI policies. The likely reason for this small percentage is the fact that for a middle class individual 60-75% of premiums for a LTCI policy would go towards benefits that Medicaid would provide anyway. American citizens over the age of 65 are unlikely to purchase private LTCI policies (Market, 2005).

The cost of LTCI ranges from $55 per year to $12,000 per year, depending on the individual's age and the level of care the policy provides. Benefits can vary based on setting and type of care, conditions covered, maximum daily benefit, benefit period, elimination period, and inflation adjustment, with each of these factors affecting the cost of the policy (Long Term Care, 2007). The major difference between LTCI and a regular life insurance policy is that the duration and cost of long term care are unknown when the benefit is triggered. (A death benefit is much simpler to model.)   An individual may need LTC services for brief periods separated by intervals in which benefits are not required.   Medical conditions that are life threatening may make a person a bad risk for life insurance but the same medical conditions will not lead to significant LTC expenses. LTCI policy is meant to protect an individual's family from high expenses while that person is still alive, whereas a life insurance policy is meant to protect an individual's family in the event of that person's death (Johnson, 2007).

2.2 Prior Work

This section will describe the previous works that are the basis for this project. A Major Qualifying Project (MQP) by Nicholas Rackliff and Elizabeth Arsenault from 2006 (Arsenault 2006) developed an expert system to underwrite ideal LTC applicants. Much of their work was focused on the expert system itself and creating a user-friendly interface, not on the applicants. In summer 2006, a group of students working on a Research Experience for Undergraduates (REU) project (Chen 2006) continued the MQP work by developing an adaptive risk scoring model for LTCI applicants. Their work consisted primarily of developing ways to estimate coefficients needed in the model as well as algorithms that allow the model to adapt and use data from actual underwriting decisions. Both projects used idealized sets of applicants to test the model. This project tests these previous models using real data from John Hancock.

2.2.1 Expert System Design MQP

This section describes the previous work completed in the area of Long Term Care Insurance (LTCI) for a senior project, or Major Qualifying Project (MQP), at Worcester Polytechnic Institute (WPI) (Arsenault 2006). The main project objective was to develop an expert system to calculate risk scores for LTCI applicants based on their medical conditions. An expert system in the field of underwriting is a program that learns from underwriters based on their procedures so that others can use it for evaluating underwriting decisions. The MQP team worked to develop an expert system in the field of LTCI underwriting and for the assignment of applicants to a risk class. Applicants' total risk scores fall into risk classes that have predetermined ranges and identify the risk of the applicant, or the probability of the applicant making a claim.

The model developed by the 2006 MQP team calculated risk scores for each of the following categories: *medical conditions*, *comorbidities*, *functional*, and *lifestyle*. The risk scores of the separate categories are summed to obtain total risk score, R_T, for an applicant:

$$R_T = R_C + R_{CM} + R_F + R_{LS} \qquad \text{[Eq 2.1]}$$

For example, an applicant with only one medical condition with R_C equal to 250 would result in a R_T of 250.

The 2006 MQP worked with Avon LTCL to develop a matrix *medical condition risk points* for up to 7 levels of severity for almost 500 medical conditions. The group computed a medical condition risk score, R_C, by first sorting the medical condition risk points, cr_i, associated with each medical condition in descending order. The medical condition risk score is a weighted sum of the condition risk scores where the weights are used "to diminish the effect of an applicant having a large number of minor conditions" (Arsenault 2006, 15). The risk score for medical conditions is given by the following equation:

$$R_C = \sum_{i=1}^{n} \frac{cr_i}{b_c^{(i-1)}}, \qquad \text{[Eq 2.2]}$$

where *n* is the total number of medical condition an applicant has and cr_i is the *i*th medical condition's risk points, and $\frac{1}{b_c^{(i-1)}}$ is the weight assigned to the *i*th condition.

The 2006 MQP team developed risk scores for comorbidities, functional, and lifestyles. Comorbidities are negative interactions between two or more medical conditions. The presence of comorbid medical conditions raises the risk for an applicant. To determine a total comorbidity risk score, R_{CM}, the following is used:

$$R_{CM} = \sum_{k=1}^{m} cmr_k \cdot b_{cm}^{(k-1)}, \qquad \text{[Eq 2.3]}$$

where n is the number of comorbid pairs, cmr_k is the risk score for one of the comorbid pairs, and $b_{cm}^{(k-1)}$ is the weight assigned to the kth comorbid pair. The weights were increasing in the model, indicating that comorbidities add more risk than just the sum of the separate medical conditions.

The functional risk score, R_F, calculation is similar to the comorbidity equation. This score represents the applicant's information about Activities of Daily Living (ADL) and if assistance (medical equipment or personal) is required. This functional risk score increases the total risk score of an applicant; difficulty or failure to perform ADLs, which are vital activities needed to be preformed daily such as bathing, eating, transferring, or dressing, are scored higher. In fact, failure for a single ADL is sufficient cause to decline coverage. The functional risk score, R_F, is found by the following:

$$R_F = \sum_{j=1}^{q} fr_j \cdot b_f^{(j-1)} ,$$ [Eq 2.4]

where q is the total number of functions, fr_j is the functional factor assigned to each element, and b_f is the compounding factor similar to the one used in obtaining the comorbidity risk score, R_{CM}.

The final component of the total risk score model is lifestyle elements. Depending on the lifestyle, this risk score can increase or decrease the total risk score of an applicant. Exercise can decrease the risk of an applicant while activities like smoking and drinking heavily can increase the risk. The lifestyle risk score is found by summing each lifestyle activities lsr, or risk factor. No interactions between activities were assumed and the following is the resulting formula:

$$R_{LS} = \sum_{l=1}^{r} lsr_l ,$$ [Eq 2.5]

where r is the total number of lifestyles.

The finished product consisted of VBA macros in Excel that include tables of conditions and their associated risk scores, a comorbidity matrix, functional limitations, and lifestyles. It also contained a macro which uses a distribution specified by the user to create a test data set of LTCI applicants.

Another macro allows a new applicant's information—conditions, comorbidities, functional, and lifestyle—to be added into the applicant pool by the user. The total risk score for individual applicants can also be viewed and verified by the user, while another macro can calculate the total risk scores for all the applicants and record the results in a spreadsheet. This macro also analyzes the risk scores by providing graphics of the distribution and the total risk scores. The last macro transforms the applicant data in the spreadsheet into a matrix form that can be used in Matlab to determine the optimal weights for the model.

The parameters b_c, b_{cm}, and b_f used for discounting or compounding in the formulas described above were found using the least absolute deviations method (as apposed the a least squares method used in the REU project (Chen 2006) and in this project). For the medical condition risk score, the target score was based on the sum of the two highest (riskiest) condition risk scores. By summing the risk scores for the comorbid pairs and multiplying by a constant, a target score was obtained. The functional target score was obtained in a similar way. All three target scores were than summed to determine the total target score for an applicant. A matrix containing applicant information (condition risk scores, comorbidities, functional, and lifestyle data) and a column vector of the total target scores were used to determine the unknown b's. These parameters are used to calculate each applicant's total risk score. The goal is to have each applicant assigned to one of the five risk bins (Preferred, Select, Class I, Class II, or Declined) based on the applicant's total risk score.

Varying the parameters, or b's, had a small impact on the total risk score and in turn on the risk bin assignment of the applicants. When the condition discounting factor, b_c, was increased by 20%, it was seen that 93% of the applicants had their total risk scores change by less than 10%.

2.2.2 Adaptive Risk Score Assignment Model REU Project

The REU project (Chen 2006) built on the framework from the MQP to create and test an adaptive model for assignment of a Total Risk Score for an applicant of LTCI based on the applicant's medical conditions and comorbidities. The objectives were to generalize the work done in the MQP and develop a scoring model that would "learn" from underwriter data.

The REU project focused on only medical conditions and co-morbidities. The Condition Risk Score uses risk point matrix developed in the MQP and the Comorbidity Risk Score uses the conditions that make up comorbid pairs. Applicant data are fed into the model in groups, between which the model updates itself in several ways, attempting to reduce assignment error.

The report describes three different methods for computing the Condition Risk Score: *Score Interpolation*, *Single-Plane Coefficient Interpolation*, and *Multi-Plane Coefficient Interpolation*. Both the *Score Interpolation Model* and the *Single-Plane Coefficient Interpolation Model* involve fitting a plane to known applicant information with pairs of conditions; while the *Score Interpolation Model* has a plane of Condition Risk Scores, the *Single-Plane Coefficient Interpolation Model* has a plane of coefficients that are then used to calculate the Condition Risk Scores. The *Single-Plane Coefficient Interpolation Model* is a special case of the *Multi-Plane Coefficient Interpolation Model*: the single plane referred to in the title is a plane that combines two different conditions, while the Multi-Plane model determines a new plane for each different number of concurrent conditions.

The final choice of model for estimating the Condition Risk Score (CRS) was the *Single-Plane Coefficient Interpolation Model* using least squares regression. The algorithm actually generates a (slightly) different model depending on the number of medical conditions in the applicants file. The mathematical models are

$$Y_{crs}^2 = b_0^2 + b_1^2 p_1 + b_2^2 p_2$$
$$Y_{crs}^3 = b_0^3 + b_1^3 p_1 + b_2^3 p_2 + b_3^3 p_3$$
$$Y_{crs}^4 = b_0^4 + b_1^4 p_1 + b_2^4 p_2 + b_3^4 p_3 + b_4^4 p_4$$
$$Y_{crs}^5 = b_0^5 + b_1^5 p_1 + b_2^5 p_2 + b_3^5 p_3 + b_4^5 p_4 + b_5^5 p_5$$

In these models, Y_{crs}^k is the condition risk score for an applicant with k medical conditions and $p_1, p_2, \ldots p_5$ are the risk points assigned to the applicant's medical conditions in decreasing order. The coefficients are different for each number of medical conditions. For example, b_1^k is the coefficient associated with the "worst" medical condition for a patient with k medical conditions. The coefficients are determined by a least-squares best fit for that group based on the applicants' underwriter decisions.

A test dataset was constructed in order to measure the model's ability to adapt by modifying its parameters. The model is capable of adjusting three different sets of parameters: condition risk points p, condition coefficients b_j^k, and comorbidity coefficients $w_{\alpha\beta}$. Five thousand and fifty simulated applicants were created to test the model's accuracy in assignment and the speed at which the model learns with respect to the amount of applicants.

The first stage consisted of initializing the model with correct parameters p, b_j^k, and $w_{\alpha\beta}$; this was done to test the "baseline" error in the model. Baseline error occurs because

underwriters assign applicants to (relatively large) risk classes while the model uses a target risk score with is a single point in the risk class interval. The baseline error is due to the loss of precision when assigning numerical risk scores ("target" risk scores) to applicants from underwriting decisions so that the mathematics of the model can properly handle the applicants. The large majority of the applicants tested in the model (~98%) were placed into the correct risk bin by the model.

The next tests involved initializing the model with all correct parameters except for one of the following: model coefficients, comorbidity weights, and condition risk points. 94.86%, 95.8%, and 57%, respectively, of each test's applicants were placed into the correct risk class. It is important to note that in each case, there were less incorrectly placed applicants in the last 1250 applicants than in the first 1250 applicants. This displays the model's ability to adapt.

The last test involved initializing the model with model coefficients, comorbidity weights, and condition risk points all incorrect. In this case, the model correctly placed 82% of the applicants. After the average risk point error per applicant stabilizes, comparing the average risk point error per applicant to the estimated baseline error determined in the first test shows that the model coefficients do not all stabilize around the correct values.

The foundation that the paper lays on creating a model for long term care underwriting allows for application of the theoretical model using actual underwriting data as a basis. The problems associated with analyzing actual data, such as incomplete or incorrect information, combined with the possibility of incorrect model assumptions, such as assuming that all condition information is encompassed within condition risk points and co-morbidities, need to be tested and addressed in order to determine the viability of this model.

2.2.3 Credit Risk Scoring Model

A well known risk scoring model used in the United States is the credit risk scoring model. Every adult in America has a credit score associated with his or her credit history, or lack thereof. The most commonly used credit scoring method is the one developed by Fair Isaac, called the FICO score. Their scores range from 300 to 850, with a higher score indicating a better risk. A person's credit score is the primary tool for lenders to determine whether or not that person should receive credit, and at what interest rate (Curry, 2006). A LTCI risk score can be used in a similar way; it would allow underwriters to determine how likely it is that an applicant will make a claim compared to other applicants and whether or not to approve the applicant for LTCI.

According to the Equal Credit Opportunity Act, a credit score must be fair and impartial, and as such the FICO score is not based on age, gender, race, national origin, marital status, or location (Federal Trade Commission, 1998). This is the opposite of an LTCI risk score, as information pertaining to an applicant's medical history is necessary for underwriting a policy. A person's credit score does include the following information, with the approximate weights supplied by Fair Isaac (Fair Isaac Corporation, 2006):

- Timeliness of past payments – 35%

- Ratio of current debt to current credit limits – 30%

- Length of credit history – 15%

- Types of credit used – 10%

- Number of recent inquiries – 10%

Defining weights in this manner is quite vague. For instance, while timeliness of past payments accounts for 35% of a person's credit score, it is unclear how much a single delinquent payment

would affect that person's score, or how many consecutive on-time payments would be necessary to negate the effects of a delinquent payment. Also, these definitions do not specify an ideal credit history length or ideal credit use is.

On July 22, 1999, Fair Isaac presented a small number of details to the Federal Trade Commission (FTC) about how their FICO score is calculated (Fair Isaac Corporation, 2000). One fact that can be drawn from this presentation is that Fair Isaac treats credit factors in the exact same manner that medical conditions have been treated by previous groups working on this project – by simply giving them a number. Though the FICO scoring method and the previously developed risk scoring models compute two very different scores, both are based on the idea of assigning a numerical value to important risk factors that are considered for each purpose. The FICO score is concerned with such information as debt ratio, time at present job, time at present address. The LTCI risk scoring model is concerned with information about medical conditions, mental health, and the ability to perform ADLs.

One portion of the Fair Isaac method for developing their credit score is similar to the condition risk scoring method. Fair Isaac does not apply exactly the same weight to each factor used in their calculation of a credit score for every person. They have developed different "scorecards" for different groups of people. One example demonstrated to the FTC is that Fair Isaac treats people with a "previous serious delinquency or derogatory public record" differently from those without one (Fair Isaac Corporation, 2000). Essentially, the scorecard method is declaring that if a Person A's credit history varies drastically from that of Person B then it makes sense to treat Person A's credit differently. This scorecard approach to scoring could easily be translated to the condition risk score method. If Applicant A has three medical conditions and

Applicant B has three medical conditions but at far more severe levels, Applicant A should be deemed less risky than Applicant B.

3. About the Data

The project goal is to test the risk scoring model developed by the REU team using data provided by John Hancock. This section describes the ideal dataset for modeling underwriting decisions and compares it to the one we actually received from the John Hancock LTCI team.

3.1 The Ideal Dataset

The model requires a specifically formatted list of information about an applicant in order to calculate that person's risk score. It needs a unique identifier for the applicant, a list of medical condition indicators associated with the applicant's medical conditions and their severities, and the underwriter's assignment to a risk class of that applicant. For each condition, the severity can be any value between one and seven. The possible risk classes, ordered from lowest risk to highest risk, are Preferred, Select, Class 1, Class 2 and Decline. This information about an applicant must be listed in a horizontal vector such as a row of cells in Microsoft Excel. An example of data for applicants with three medical conditions that the model can handle is given below:

Applicant ID	Cond1	Cond1 Sev	Cond2	Cond 2 Sev	Cond3	Cond3 Sev	Risk Class
1	10	4	155	6	27	3	Decline
2	211	1	15	1	120	2	Preferred
3	75	2	89	3	166	5	Class 1
4	44	1	10	1	33	5	Select
5	7	2	100	2	234	7	Decline

Figure 3.1: Ideal Model Input

The risk scoring model's current input requires that all applicants have the same number of medical conditions. The entire set of applicants is placed into subsets based on the total number of medical conditions. Applicants can have 0, 1, 2, 3, 4, or 5 medical conditions.

The data received from John Hancock contained the necessary information, including applicant identifiers and medical condition indicators, as well as height, weight and the results of a cognitive screen test.

3.2 Data from John Hancock

John Hancock provided a set of actual LTCI applicants underwritten since January 1, 2005. This data set consists of 66,737 records, where each record represents a different LTCI applicant. There are 16,999 applicants with one medical condition, 14,505 with two, 10,639 with three, 6,680 with four, and 6,399 with five. If an applicant has more than five medical conditions, the five most severe are included in the data.

Previous groups working on this project had communicated primarily with a team from Avon LTCI, who had been assisting John Hancock with their LTCI underwriting. The Avon matrix assigns risk points to each medical condition. For a medical condition with different levels of severity, the risk points increase with every level. If an applicant had only one medical condition on which they were underwritten, the risk points associated with that condition would place that applicant into one of the risk class ranges. Each risk class represents a different level of coverage that John Hancock would be willing to provide an applicant, if any, with Preferred being the best and Class 2 being the worst possible coverage. The ranges for each risk class are given below:

JH Risk Class	Risk Score Range
Preferred	0 - 200
Select	200 - 300
Class 1	300 - 400
Class 2	400 - 500
Declined	500 - 1000

Figure 3.2: Risk Class Ranges

Recently, however, John Hancock ended their connection with Avon, so the data is not in the form that the model requires. For instance, John Hancock does not include severities with each medical condition as Avon had; rather, John Hancock combines the condition and a description of its severity into a single condition indicator. As an example, to Avon, the medical condition of alcoholism is given an identifying value of 8, and it has severities one to seven, each with different risk points associated to them, as shown below:

ID	Medical Condition	1	2	3	4	5	6	7
8	Alcoholism	100	111	134	182	288	514	1000

Figure 3.3: Example Avon Medical Condition Indicator

To John Hancock, alcoholism looks like this:

Indicator	Condition	Description	Risk Class
11	Alcoholism	Abstinence >18 months, chronic treatment with an an	Individual consideration
12	Alcoholism	History of alcoholism with cirrhosis	Decline
13	Alcoholism	History of alcohol abuse with current fatty liver or abn	Decline
14	Alcoholism	Alcohol abuse with multiple (three or more) hospitaliz	Class 1 w/180 EP or Class 2
15	Alcoholism	Alcohol abuse with multiple hospitalizations and/ or d	Postpone 36 months
16	Alcoholism	Alcohol abuse with detox admission, current abstinen	Postpone 18 months
17	Alcoholism	Abstinence >18 months, normal blood studies, no cor	Select
18	Alcoholism	Abstinence <18 months, normal blood studies, no cor	Postpone 18 months
19	Alcoholism	Alcohol abuse with detox admission, current abstinen	Select

Figure 3.4: Example John Hancock Medical Condition Indicator

Among the information for an applicant that must be analyzed before being used in the model is a value for the person's Minnesota Cognitive Assessment Screening (MCAS) score. For John Hancock's purposes, a score of zero or greater is considered passing, while anything less than zero is a failing score. Anyone with a failing score is automatically declined LTCI, regardless of their medical conditions. All applicants with failing MCAS scores must be removed from the data before it is input into the model. If these people are not removed, the model may learn improperly. An applicant that would otherwise be accepted for long term care

insurance if not for a failing MCAS score would indicate to the model that the medical

conditions for that applicant are more serious than they actually are.

The model can compute a risk score based on an applicant's medical conditions. Medical

conditions can be assigned risk points based on their levels of severity. However, an underwriter

looks at more than the information that can be quantified. Lifestyles such as running marathons

or smoking a pack of cigarettes per day are reviewed by underwriters; a decision is based on the

whole application rather than just medical conditions. We observed the correlation between

underwriter guidelines for a medical condition and the underwriting decisions by isolating the

medical condition from the other factors the model takes into account. Applicants that had

exactly one condition, acceptable BMI, and an acceptable MCAS score were examined to see

how frequently the actual underwriter decision matched the decision given in John Hancock

underwriting guidelines. The following table displays the twenty most common conditions for

these applicants, the underwriter guidelines for that condition, and the most frequent decision

made for the applicants:

Indicator	JH Condition Name	JH Decision	Decision Mode
341	Hypertension	Preferred	Preferred
40335	Hypercholesterolemia	Preferred	Preferred
491	Osteopenia	Preferred	Preferred
40342	Hypertension	Select - Class 1	Select
2161	Depression	Select - Class 1	Select
4667	Tobacco Use	Select	Select
668	Tobacco Use	Select	Select
342	Hypertension	Select - Class 1	Select
335	Hypercholesterolemia	Preferred	Preferred
120	Cancer - Internal	Select w/365-day EP	Select
493	Osteoarthrits / DJD	Select	Select
4059	Anxiety	Rate 1 to Rate 2	Select
287	General Muskuloskeletal Disorder	None Given	Preferred
551	Prostatic Hypertrophy - Benign	Preferred	Preferred
348	Hypothyroidism	Preferred	Preferred
484	Osteoarthrits / DJD	Select	Select
20	Allergies	Preferred	Preferred
4064	Arthritis	Select	Select
83	Asthma	Select	Select
4338	Hypertension	Preferred	Preferred

Figure 3.5 – Most Frequent Medical Conditions

4. Results

This section details all of the results of the project. It describes the filters applied to the applicants, how risk scores were assigned to each medical condition, the data input and output to and from the model, a brief overview of how to run the model, and model coefficients from initial and later trials.

4.1 Filtering the Applicants

The following characteristics describe the applicant pool used to initialize and train the model:

- All applicants have MCAS scores greater than or equal to zero

- All applicants have complete underwriting decisions

- All applicants have medical conditions corresponding to unique John Hancock condition indicators

- All applicants have medical conditions with risk scores less than 750

- All applicants have no comorbid medical conditions

- All applicants have Body Mass Indexes (BMI) less than 35

- All applicants have medical conditions with risk scores greater than zero

- All applicants have medical conditions where maximum condition risk score is greater than or equal to the underwriting decision

The model can be initialized and trained using a subset of the database containing 55,220 applicants with 1, 2, 3, 4, or 5 medical conditions. However there were several complications that could not be accounted for and certain applicants needed to be removed from the useable applicant pool.

The following chapter describes how each of these characteristics contributed to an applicant pool with which a model could be developed. Each filter was applied to a subset of the 55,220 applicants; each subset consisted of applicants with the same number of listed medical conditions.

4.1.1 MCAS Score

All of the applicants have passing Minnesota Cognitive Acuity Screen (MCAS) scores. By keeping only applicants with positive scores, 876 applicants with negative scores were removed from the applicant pool. This filter was performed on the entire applicant pool, including applicants that have conditions with 0 risk points.

MCAS results are widely used by John Hancock for LTCI underwriting to reduce the likelihood of applicants making claims based on mental health issues. The test has been shown to be 98.1% accurate at determining mild to moderate cognitive impairment (John Hancock, 2005).

An MCAS score of less than zero constitutes an automatic decline and the applicant's medical conditions would not be considered for further review by the underwriter. Since the underwriter would not have to determine a total risk score for the applicant, the use of the model is not needed and such an applicant would not be valuable in training the model.

4.1.2 Complete Underwriting Decisions

The applicants that can be used in training the model have complete underwriting decisions. An underwriting decision is complete if the applicant was ultimately approved or declined for LTCI. The possible underwriting decisions are A (Approved), AM (Approved with Modifications), D (Declined), DM (Declined with Modifications), and DP (Decision Postponed).

Decisions of D, DM, or DP were all treated identically. Since the final decision of an applicant

with underwriting decision DP is not known, it is assumed that the applicant was declined.

However, some applicants were listed as having incomplete (I) decisions or with the

decision field blank. So the underwriting decision is not known. The model needs to know how

each applicant is assigned in the end because the model compares the computed total risk score

to an underwriter's risk class assignment. With out the underwriting decision, the applicants can

not be used to train the model to pick up on condition risk scoring errors, and the applicants were

removed. The following table shows how the useable applicant pool changed after the

application of this filter.

No. of Conditions	Applicants Received	Applicants Removed	Useable Applicants
1	16999	433	16566
2	14505	319	14186
3	10637	277	10360
4	6680	199	6481
5	6399	172	6227
Total	55220	1400	53820

Figure 4.1 – Applicant Counts after U/W Decision

Only 1,400 applicants had incomplete or blank underwriting decisions.

4.1.3 John Hancock Condition Indicators

All of the applicants have condition indicators that specify the medical conditions and the

corresponding level of severity. Applicants' medical conditions are identified with these

indicators, for example medical condition indicator 341 represents hypertension. To obtain the

risk score associated with each medical condition, the condition has to appear in the Avon

condition matrix, which consists of seven levels of severity for 486 conditions. Using the most

recent (April 2006) underwriting guidelines, each John Hancock medical condition is matched

with a risk score from the Avon matrix. Underwriting condition guidelines from February 2005

and July 2005 were also available and used to find medical condition descriptions and check condition indicators.

However, there are conditions that appear in the John Hancock underwriting guidelines but do not appear in the Avon matrix. The applicants with these medical conditions, such as lymphoma, dwarfism, and panic disorder for which risk scores could not be obtained, were removed.

The following table illustrates how the useable applicant subset was affected after the application of this filter.

No. of Conditions	Applicants	Applicants Removed	Useable Applicants
1	16566	814	15752
2	14186	1287	12899
3	10360	1360	9000
4	6481	1079	5402
5	6227	1164	5063
Total	53820	5704	48116

Figure 4.2 – Applicant Counts after Medical Conditions

4.1.4 Automatically Declined Medical Conditions

There are medical conditions whose risk scores are 750 or greater. These conditions are defined as automatic declines. When the standard underwriting practice is to decline any applicant with one of these conditions, the applicant does not provide new information about underwriting auto declines.

The following table shows the useable applicants after auto declines were removed.

No. of Conditions	Applicants	Applicants Removed	Useable Applicants
1	15752	2012	13740
2	12899	2886	10013
3	9000	3165	5835
4	5402	2758	2644
5	5063	3609	1454
Total	48116	14430	33686

Figure 4.3 – Applicant Counts after Automatic Declines

A large subset of applicants with auto decline medical conditions was removed and 33,686 applicants remain.

4.1.5 Comorbid Medical Conditions

Comorbid medical conditions were removed only because we have not developed a method of assigning risk scores to comorbid pairs. The existence of comorbidities affects the risk scores associated with the medical conditions negatively, increasing the risk associated with the medical conditions.

The following table shows how removing comorbid medical conditions affected the useable applicant pool. Note that for applicants having only one medical condition this filter made no difference; there have to be at least two medical conditions for a comorbidity to exist.

No. of Conditions	Applicants	Applicants Removed	Useable Applicants
1	13740	0	13740
2	10013	890	9123
3	5835	1406	4429
4	2644	1050	1594
5	1454	813	641
Total	33686	4159	29527

Figure 4.4 – Applicant Counts after Comorbidites

4.1.6 Body Mass Index

Applicants with a Body Mass Index (BMI) great than or equal to 35 were removed. The Centers for Disease and Control and Prevention (CDC, 2006) defines BMI as an indicatory of "body fatness." The number is calculated using an applicant's height (inches) and weight (pounds), along with a conversion factor in the following manner:

$$BMI = weight \div [height]^2 \cdot 703 \qquad\qquad \text{[Eq. 4.1]}$$

The CDC also defines the following ranges and weight statuses.

BMI	Weight Status
Below 18.5	Underweight
18.5 to 24.9	Normal
25.0 to 29.9	Overweight
30 or Higher	Obese

Figure 4.5 – Body Mass Index Figure

A high BMI can be comorbid with other medical conditions; obesity can cause or worsen hypertension, stroke, sleep apnea, and some cancers (CDC, 2006). Since all comorbidities are removed from the useable applicant pool, a potentially comorbid BMI also has to be removed. It was determined that a BMI greater than or equal to 35 will be removed. Since 35 is a very high BMI, it seemed quite likely that this score would be comorbid with other medical conditions.

The following table shows the results of removing applicants with high BMIs.

No. of Conditions	Applicants	Applicants Removed	Useable Applicants
1	13740	541	13199
2	9123	399	8724
3	4429	1197	3232
4	1594	88	1506
5	641	215	426
Total	29527	2440	27087

Figure 4.6 – Applicant Counts after BMI

4.1.7 Zero Point Conditions

The applicants in the useable pool also have no zero point conditions. Conditions such as allergies, high cholesterol, gout, migraines, and hypertension/white coat syndrome all have condition risk scores of zero when in the lower levels of severity. Zero point conditions do not contribute to the total risk score of an applicant; regardless of the weight assigned to the zero point condition that will be equal to zero in the total risk score equation.

This filter had the unique property of adding applicants to the subsets of the applicant pool with less than five medical conditions. An applicant with 4 non-comorbid medical conditions but 2 zero point conditions is now moved to the 2 condition subset.

No. of Conditions	Applicants	Applicants Removed/Added	Useable Applicants
1	13199	+333	13532
2	8724	-2679	6045
3	3232	-1374	1858
4	1506	-923	583
5	426	-321	105
Total	27087	-4964	22123

Figure 4.7 – Applicant Counts after Zero Point Conditions

4.1.8 Maximum Condition Score ≥ Underwriting Decision

A complete underwriting decision not only indicated whether or not the applicant got approved but identified which risk class the applicant was placed into. The risk class decision can be 1 (Preferred), 2 (Select), 4 (Class I), or 5 (Class II). Applicants who were declined were assigned to rate class 6. The number 3 is not used to denote a risk class in John Hancock underwriting procedures.

Applicants that have value in training the model have maximum condition scores that are greater than or equal to the overall underwriting decision. If Applicant A's highest (riskiest) medical condition would place them into Class I (risk score range: 300 to 399) and the underwriting decision rates that applicant as a 4 (Class I) or lower, the applicant would remain in the useable pool. Applicant B's highest medical condition also places them into Class I. However, the final decision for Applicant B was a 5 (Class II) or higher. Applicant B would be removed from the applicant pool.

Why was Applicant B rated higher (riskier) than Applicant A? The correct answer is not known and that is why some applicants are removed. More information is needed about the applicant than what is provided in the database. The model assigns total risk scores based on the medical conditions of each applicant; some applicants have other information in their application that is unknown and can negatively influence the underwriting decision. Applicants such as Applicant B who were placed higher are removed due to lack of information outside the

provided medical conditions. The following table shows the impact of this filter, as well as the final useable applicant pool.

No. of Conditions	Applicants	Applicants Removed	Useable Applicants
1	13532	2161	11371
2	6045	647	5398
3	1858	210	1648
4	583	77	506
5	105	16	89
Total	22123	3111	19012

Figure 4.8 – Applicant Counts after All Data Filters

4.2 Condition Mapping

The next step in preparing the data for use in the model was to map each John Hancock medical condition indicator to a matching condition and severity in the Avon risk point matrix. For instance, a John Hancock condition with a Preferred decision, such as Acoustic Neuroma, needs to be assigned a risk score that falls with in the Preferred range, or 0-200.   In the Avon matrix, Acoustic Neuroma has seven levels of severity, three of which fall within the Preferred range. The description of the medical condition in the underwriting guidelines was used to determine the severity of this specific case of Acoustic Neuroma and determine that a risk score of 100 is appropriate. Since John Hancock and Avon did not have an identical list of conditions, the John Hancock underwriting guidelines were used to update the Avon matrix.

Some conditions had a decision of Postpone in the John Hancock underwriting guidelines, rather than a clear decision of approved or declined. A Postponed decision was considered to be a temporary Decline; by this logic, the risk score associated with the medical condition should assign the applicant to the Declined risk class and have a risk point greater than 500.

4.3 Grouping the Applicants

The applicants were grouped into populations based on the underwriting rate class decision (1, 2, 4, 5, or 6). The first population included all the applicants that were placed into the Preferred class or rate class 1. A total of 67 or 5.72% of the three medical condition applicants are Preferred. The following table shows the distribution of applicants with three medical conditions grouped by underwriting rate class decision:

Three Condition Group	
UW Rate Class	**# of Applicants**
Preferred	67
Select	880
Class 1	103
Class 2	4
Declined	117
Total	1171

Figure 4.9 – Distribution of Three Condition Applicants by Underwriting Rate Class

There are five populations of applicants, each containing applicants with the same number of medical conditions. For each population, a set of model coefficients was found using the least squares fit. At first, applicants within a population were introduced by subgroups called generations to determine model coefficients. The first generation introduced was the Preferred applicants, and then the Select applicants were added on top of the Preferred applicants, and so on, until all 1171 applicants were being used in the last run and the final coefficients were recorded.

Generations were next formed by sorting the applicants by increasing condition scores and coefficients were recorded. Next, all of the applicants were sorted in ascending order based on their first medical condition risk point. The last method tried sorted applicants by the sum of their condition risk points. The applicants were then put into generations of 100 applicants each in each of the above methods. The first 100 applicants were introduced to the model, then the

first generation along with the next 100 applicants and so on until all 1171 applicants were used in the 12th generation.

4.4 Initial Trials

The Risk Scoring Model:

$$TrueRiskScore = b_0 + b_1 P_1 + b_2 P_2 ... + b_n P_n \qquad \text{[Eq. 4.2]}$$

where P_i is a medical condition risk point associated with the ith medical condition and b_i is a coefficient created by performing a least squares fit on the applicants. For applicants with four medical conditions, the intercept is denoted as b_0, and the coefficient for the first medical condition is denoted as b_1, the coefficient for the second condition is b_2, the coefficient for the third condition is b_3, and the coefficient for the fourth condition is b_4. The coefficients, including the intercept, are different depending on what applicant subset is used; applicants with 2 medical conditions have different coefficients than applicants with 3, 4, or 5 medical conditions.

Each group was run though to generate the coefficients for the four-condition applicants. The following tables display the results when four condition applicants were introduced by risk class assignment; first all the Preferred applicants, next all the Preferred and Select applicants, until the fifth run was all applicants.

Sorted by the U/W Rate Class

	Preferred	Select	Class 1	Class 2	Declined
b1	0	0.2277	0.3599	0.3914	1.0031
b2	0	0.1073	0.0606	0.0391	0.1631
b3	0	-0.0072	0.1062	0.1215	0.2897
b4	0	-0.0412	0.0013	-0.0093	-0.3343
b0	100	203.8498	195.6617	193.2046	134.7957

Figure 4.10 – Coefficients when Applicants Sorted by Rate Class

The applicants were also introduced in generations of 100 applicants each. The applicants were sorted according to their highest condition risk score. First, 100 applicants were introduced to the model and a least squares fit was done. Next, the first 100 applicants and the next 100 applicants, for a total of 200 applicants, were put through the model to find the coefficients. The following table shows how the coefficients change as more applicants were introduced.

Sorted by Greatest Condition Risk Score

	b1	b2	b3	b4	b0
100	1.0958	0.9271	-0.3990	0.3809	149.2883
200	0.7213	0.5004	-0.3625	-0.0829	162.9102
300	0.7404	0.3271	1.2519	-0.5587	154.8330
400	0.5903	0.1910	0.7426	-0.4275	173.2616
500	0.4917	0.2311	0.7079	-0.2761	183.3892
600	0.6744	0.1879	0.7335	-0.3279	167.6604
700	0.5391	0.2800	0.5179	-0.4166	175.5059
800	0.5004	0.3304	0.3312	-0.5831	175.3732
900	0.5666	0.2807	0.4123	-0.5005	171.8468
1000	0.6736	0.2341	0.4301	-0.2936	165.9208
1100	0.8501	0.0638	0.4458	-0.3251	156.0641
1172	1.0031	0.1631	0.2897	-0.3343	134.7957

Figure 4.11 – Coefficients when Applicants Sorted by Greatest Condition Risk Score

The least squares fit was also done on generations of 100 applicants each and the applicants were sorted by the sum of their condition risk scores. The following table shows how the coefficients change as more applicants were introduced.

By the Sum of the Condition Risk Scores

	b1	b2	b3	b4	b0
100	1.0951	0.9561	-0.3797	0.3412	148.0142
200	0.7213	0.5004	-0.3625	-0.0829	162.9102
300	0.7404	0.3271	1.2519	-0.5587	154.8330
400	0.5903	0.1910	0.7426	-0.4275	173.2616
500	0.4918	0.2310	0.7215	-0.2865	183.0980
600	0.6478	0.2090	0.6054	-0.2564	171.4074
700	0.5391	0.2800	0.5179	-0.4166	175.5059
800	0.5005	0.3304	0.3310	-0.5838	175.3599
900	0.5666	0.2807	0.4123	-0.5005	171.8468
1000	0.6736	0.2341	0.4301	-0.2936	165.9208
1100	0.8501	0.0638	0.4458	-0.3251	156.0641
1172	1.0031	0.1631	0.2897	-0.3343	134.7957

Figure 4.12 – Coefficients when Applicants Sorted by Sum of Condition Risk Scores

The final run, which uses all applicants, resulted in coefficients that were identical to the method of applicant introduction described above. However the intercepts were not as close to zero as we had anticipated. An applicant with four medical conditions, all of which are zero point conditions and are not comorbid, would still receive a risk score of 134 because of the intercept.

Since running the applicants through the model in different orders did not change the coefficients, a new approach was tried that might lower the intercept and bring more stability to the coefficients. This led to a better way of determining an applicant's "true risk score". Previously, only the underwriting decision and rate class were used to determine an applicant's target risk score. All applicants with the same underwriting decision and rate class were given the median risk score associated with that risk classification. This meant that all 880 three condition Approved applicants with a rate class of 2 were given a target risk score of 250. This created a hardship when performing a least squares fit on the data like our model needs to choose coefficients.

To help deal with this problem, the idea of sub buckets was proposed. Each risk bucket would be split into a predetermined number of sub-buckets. Four sub buckets provided better results. Again the example below uses four condition applicants.

	No Sub Buckets	2 Sub Buckets	4 Sub Buckets
b0	32.9977	15.7031	-1.0379
b1	1.056	1.132	1.2286
b2	-0.0544	-0.1201	-0.122
b3	0.0161	0.1123	0.0803
b4	0.0228	0.0906	0.0625

Figure 4.13 – Coefficients using Sub Buckets

If an applicant's condition risk scores are summed, a sense about which applicants are riskier than others can be obtained and that knowledge can be used to determine sub bucket assignment. Then, one fourth of each risk class is placed in each sub bucket. For example:

Applicant ID	U/W Decision	Risk Class	Cond1	Cond2	Cond3	Cond4	Cond_Sum	Risk Sub Bin
1	A	2	56	32	50	25	163	1
2	A	2	0	100	82	50	232	1
3	A	2	278	199	0	25	502	2
4	A	2	25	199	235	50	509	2
5	A	2	251	251	0	199	701	3
6	A	2	278	209	251	0	738	3
7	A	2	210	260	251	251	972	4
8	A	2	199	314	314	250	1077	4

Figure 4.14 – Example of Sub Bin Risk Class Placement

An underwriter would say that applicant 8 is much riskier than applicant 1, and therefore deserves a higher true risk score.

4.5 Model Inputs

The model has four inputs prepared in Microsoft Excel before being uploaded to Matlab with the built in *xlsread( )* function. The first input, *InitPop*, is the population of applicants that will initialize the model and estimate coefficients. *InitPop* is an $m \times n$ matrix, prepared in Excel, where m is the number of applicants in the population and n is the number of conditions those applicants have plus one. The first n minus one columns have the condition risk points that correlate to the applicants conditions, and the final column has the target risk scores for the applicants. The following is a sample *InitPop* input for applicants with four medical conditions.

	Condition 1	Condition 2	Condition 3	Target Risk Score
Applicant 1	255	117	83	262.5
Applicant 2	169	14	37	125
Applicant 3	235	178	222	287.5
Applicant 4	420	387	12	437.5

Figure 4.15 – Example *InitPop*

The second input is *TestPopPoints*. *TestPopPoints* is very similar to *InitPop*. Each of the applicants needs to have the same number of conditions, but the number of applicants in each population can vary. This is an $m \times n-1$ matrix, as it should include only the condition risk points and not the target risk score for each applicant. However, the true risk scores do get placed in the model by way of input three, *TestPopScores*. *TestPopScores* is a vector with the

same number of rows as *TestPopPoints* that contains the true risk scores for the applicants in

TestPopPoints. Row y of both *TestPopPoints* and *TestPopScores* should be referring to the

same applicant. The following is a sample of the testing population inputs.

	Condition 1	Condition 2	Condition 3	True Risk Bin
Applicant 5	256	100	82	262.5
Applicant 6	100	25	32	125
Applicant 7	235	189	223	287.5
Applicant 8	446	375	100	437.5

Figure 4.16 – Example *TestPopPoints* and *TestPopScores*

Finally, the last input for the model is *TestPopConds*. This input has, for each applicant

in the test population, the list of conditions the applicant has coded by their John Hancock

identifiers. *TestPopConds* will have the same dimensions as *TestPopPoints*, and furthermore

each entry in *TestPopConds* should correlate with the same entry in *TestPopPoints*.

JH 1	JH 2	JH 3
127	2	272
607	31	4338
600	689	113
100	510	607

Figure 4.17 – Example *TestPopConds*

In the sample *TestPopConds* input, Applicant 5's medical conditions are 127, 2, and 272 which

are the John Hancock condition indicators.

4.6 Model Outputs

The model features six outputs for analysis. The first output produced is *Coef*. *Coef* is a

vector with 'number of medical conditions plus one' rows. It prints out the intercept (b_0) first

and then returns the other coefficients starting with b_n, where n is the number of conditions, and

finishing with b_1, the coefficient for the first condition.

The next three outputs should be examined as a set. *PercentRight*, *PercentLow*, and

PercentHigh are just as they sound; the percentage of applicants placed correctly, the percentage

of applicants placed in a risk bucket below their true risk score, and the percentage of applicants placed in a risk bucket above their true risk score.

The final two outputs are where the adaptation of the model will come into play. *ReportHigh* and *ReportLow* give a list of John Hancock condition indicators and reports the number of times applicants with that condition where placed in a risk bucket either above or below their true risk score. Both outputs have two columns and a variable number of rows. Column one has a list of John Hancock condition indicators and column two has the number of times that an applicant with that medical condition was misplaced. Both reports are sorted so the most misplaced condition is shown in the first row.

4.7 Using the Model

First you must use a series of *xlsread( )*'s to get all the necessary variables into Matlab. For more detail on *xlsread()* please see the User's Manual. Then, with Matlab pointed to the correct directory, run:

MQPmodel(InitPop, TestPopPoints, TestPopScores, TestPopConds).

Next, the model reorders *InitPop* and *TestPopPoints* so that the highest risk points are in column one and risk points decreases to the right. Once it has reordered the conditions for each applicant it determines the values of *Coef* by performing a least squares fit on the initializing population.

The next phase of the model tests the coefficients produced in the first part using the test population provided. While doing this it counts up the number of applicants correctly placed, placed incorrectly high, or incorrectly low and divides these numbers by the number of applicants in the testing population to produce the outputs *PercentRight*, *PercentHigh*, and *PercentLow*.

The model also keeps track of any applicants that are incorrectly placed. Whenever an applicant is not placed in the correct risk bin, the model pulls all of the condition indicators for that applicant from *TestPopConds* and adds them to the running tally in either *ReportHigh* or *ReportLow* depending on how the applicant was misplaced.

Finally, all of the outputs described above are sorted and printed to the screen. *ReportHigh* and *ReportLow* can be exported to Excel and can be used to determine if changing a condition risk point for a certain condition is warranted. Once all appropriate changes have been made to medical condition risk points, re-upload the necessary inputs and run the model again.

4.8 Final Trials

Testing the model using John Hancock data was done in two phases. The first was a straight forward run of the model as described above. The second used the results of the first run to change the risk points associated with the three most misplaced conditions. The top three misplaced condition's risk points were reduced by 20. This helped determine that our model actually was adaptive, and proved the worth of an adaptive model.

These two runs of the model are compared in the table below for two, three, four and five condition applicants. The coefficients for each applicant sub group are included for each run. The table also shows the percentage of applicants placed in the correct risk class, one risk class too high, and one risk class too low as described in the About the Model section of this paper. After both runs were completed the model was then tested against the complete populations, with all applicants previously removed through filtering added back in. This percentage can be seen in the last row.

	2 Conditions		3 Conditions		4 Conditions		5 Conditions	
	Run 1	Run 2	Run 1	Run 2	Run 1	Run 2	Run 1	Run 2
b0	23.901	23.313	8.5811	6.485	-19.696	3.9432	-30.565	-23.228
b1	0.8789	0.8789	0.9492	0.9548	1.0739	0.9543	1.1515	1.1161
b2	0.0696	0.0828	0.0374	0.0449	-0.0197	0.0667	-0.1273	-0.1139
b3	-	-	0.0707	0.0828	0.0908	0.1187	0.0682	0.0648
b4	-	-	-	-	0.0007	0.0048	0.0739	0.0966
b5	-	-	-	-	-	-	0.0351	0.0271
Right	96.84%	97.05%	92.22%	92.54%	90.38%	91.48%	81.93%	86.75%
High	2.84%	2.84%	7.78%	7.46%	9.34%	8.52%	18.07%	13.25%
Low	0.33%	0.11%	0.00%	0.00%	0.27%	0.00%	0.00%	0.00%
Total Population	88.63%		84.25%		82.76%		89.00%	

Figure 4.15 – Final Model Coefficients and Placement Percentages

After the risk points of the three most misplaced conditions for each population were reduced, the percentage of misplaced applicants was diminished. The model is more accurate with two condition applicants than on other sub groups because the volume of applicants gives the model more information on which to base its parameters. The two condition population is much larger than the three condition population, and the three condition population is larger than the four condition population, and so on.

The model estimates the risk scores of the total population sub group fairly well overall. In fact, the five condition model does a better job estimating the risks of the total population than the filtered population it was initialized with. This can be explained by the fact that a large majority of applicants that were removed were considered to be "easy cases," in which the final underwriting decision was clear. The coefficients were based on the "hard cases" and are suitable to be used on the easier cases. This means that with future improvements the model will be able to predict the risk of most applicants with high accuracy.

In addition, the model was used to show how applicants were spread out across the risk bins. Below are a series of graphs for the two condition applicant population. The first graph for each population shows the distribution of applicants across all the risk class sub bins.
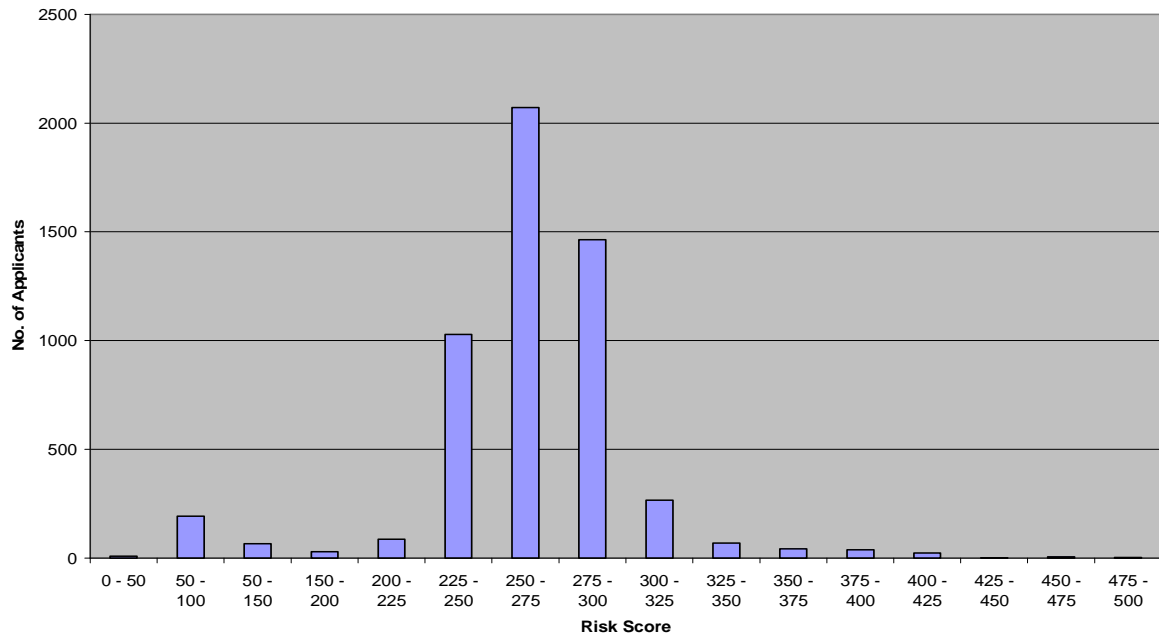
2 Condition Applicants

Chart 4.1 – Distribution of 2 Condition Applicants across All Risk Classes
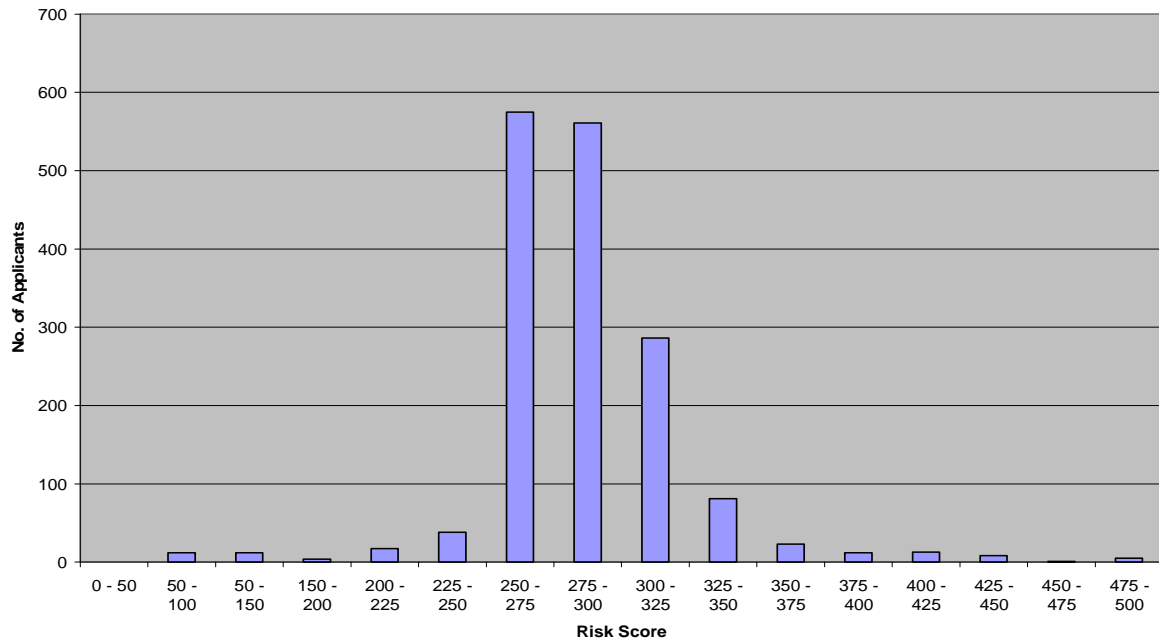
3 Condition Applicants

Chart 4.2 – Distribution of 3 Condition Applicants across All Risk Classes

4 Condition Applicants
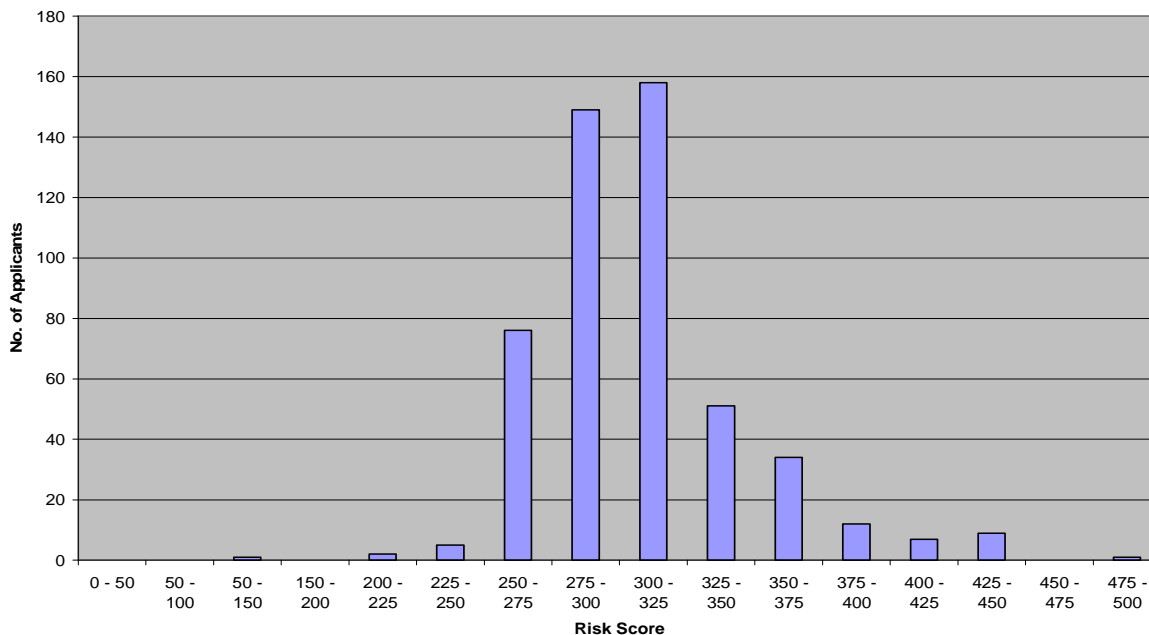
Chart 4.3 – Distribution of 4 Condition Applicants across All Risk Classes

5 Condition Applicants
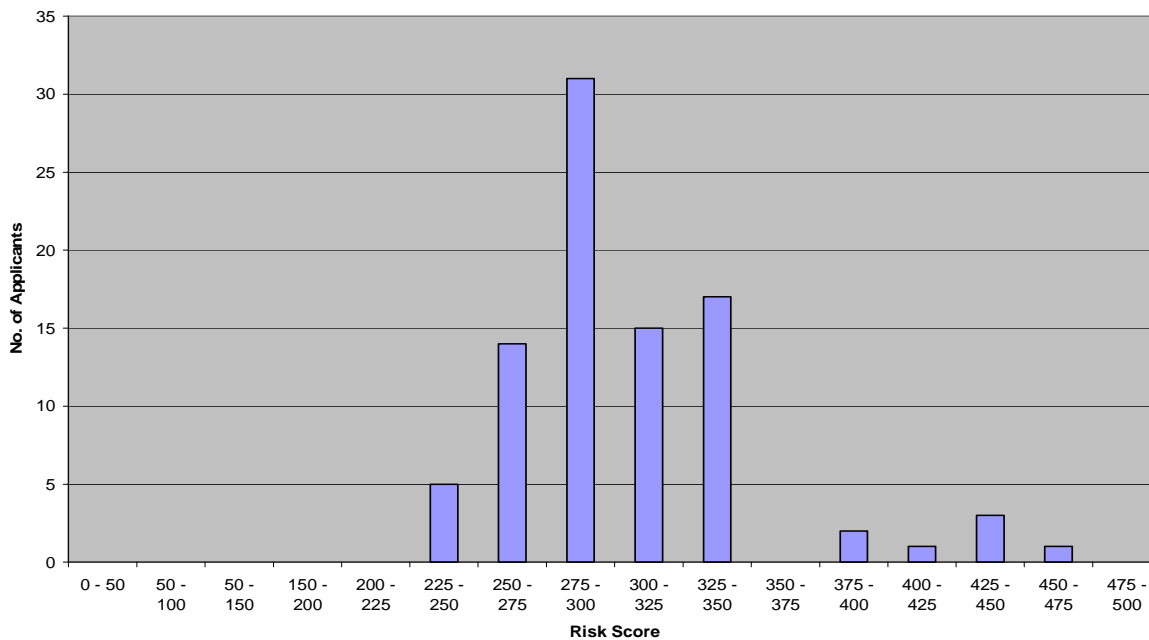
Chart 4.4 – Distribution of 5 Condition Applicants across All Risk Classes

The second set of graphs shows the distribution of applicants within the Select risk bin. Most

applicants are placed in the Select risk bin. People applying for LTCI from John Hancock will

most likely be placed in the Select risk bin. This means that the vast majority of John Hancock

LTC insureds are paying the Select premium.
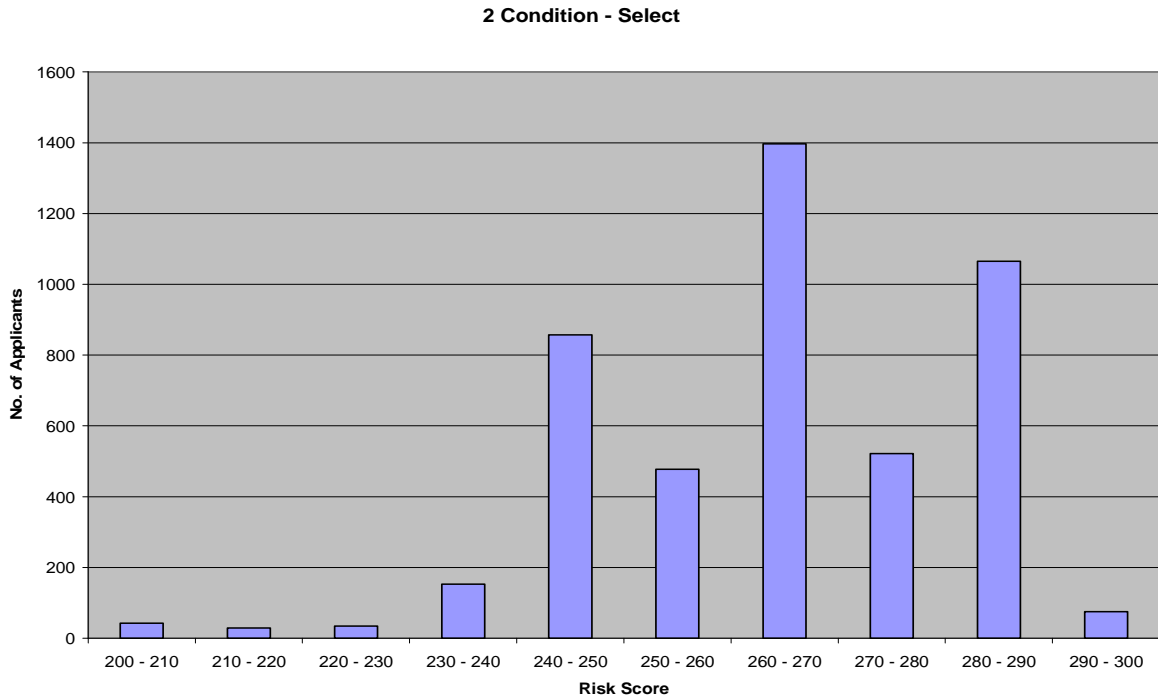
2 Condition - Select

Chart 4.5 – Distribution of 2 Condition Applicants within Select Risk Class

3 Condition - Select
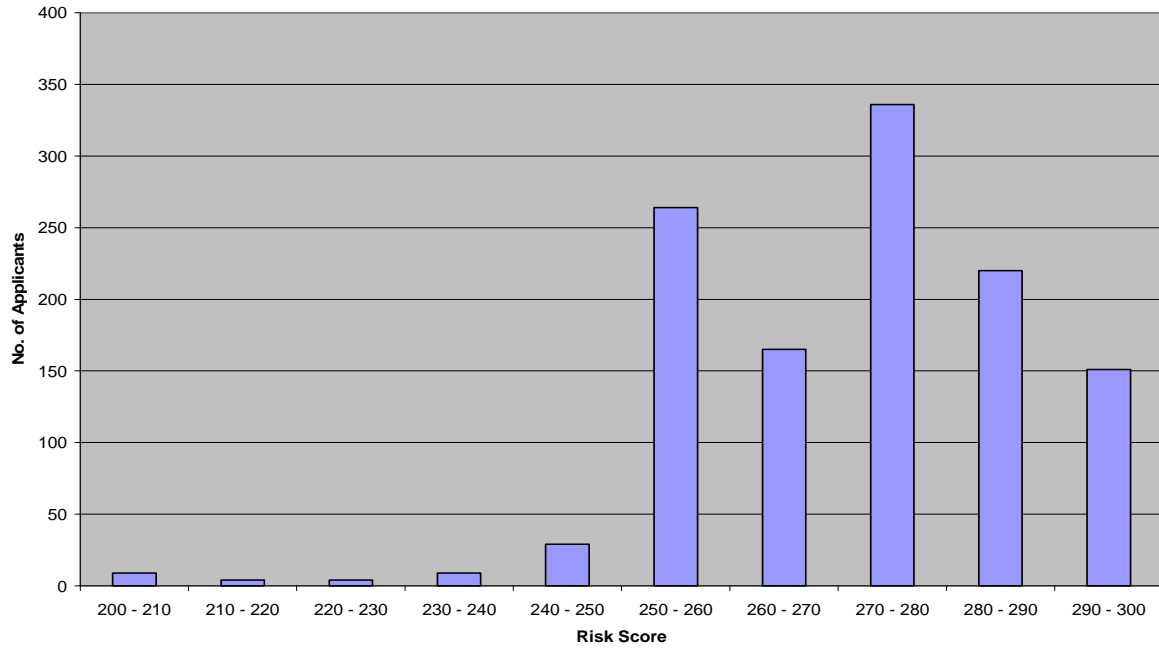
Chart 4.6 – Distribution of 3 Condition Applicants within Select Risk Class

4 Condition - Select
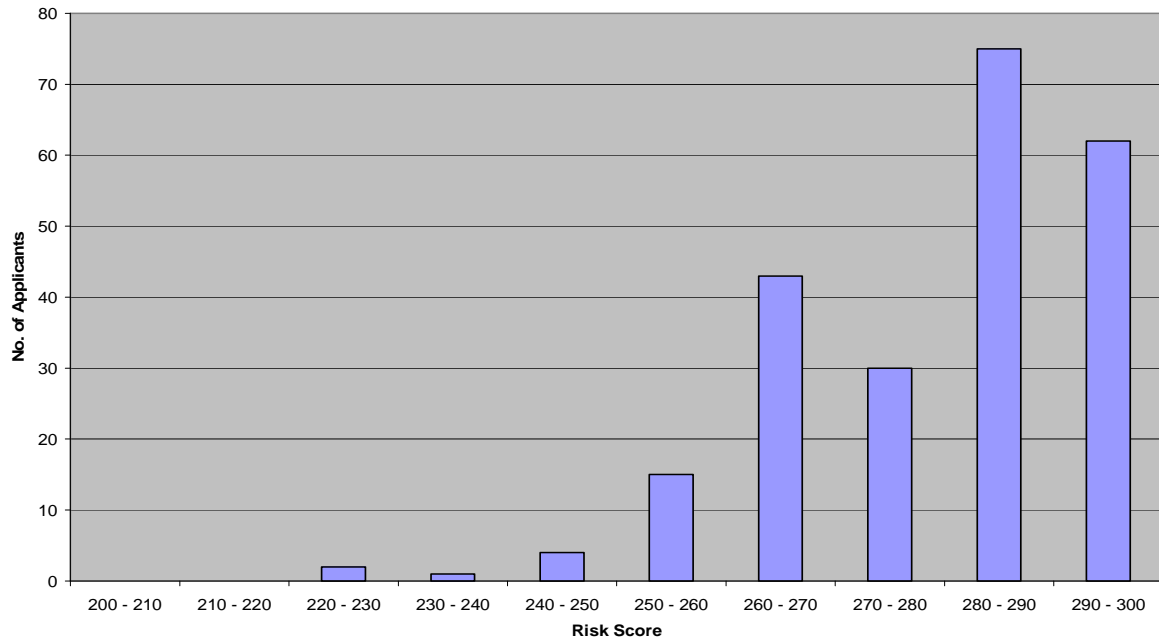
Chart 4.7 – Distribution of 4 Condition Applicants within Select Risk Class
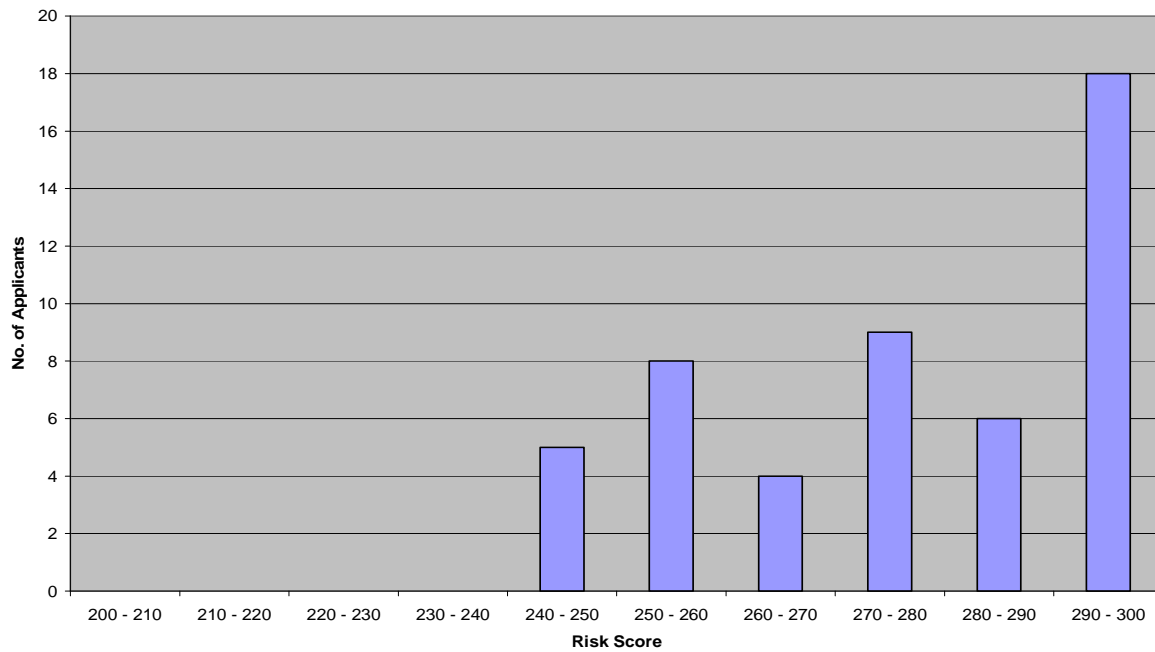
Chart 4.8 – Distribution of 5 Condition Applicants within Select Risk Class

Focusing on just the Select risk class, the range of risks for these applicants can determined more precisely. The fact that most applicants are placed on the back end of the Select risk class is important information. With this information John Hancock could lower their premiums for less risky applicants and become more competitive in the LTCI market; they could also raise premiums for higher risk applicants within a risk class.

5. Conclusions

A risk score can be assigned to Long Term Care Insurance applicants based on medical conditions, though a complete underwriting decision factors in more than just medical history. The model allows underwriters to compare applicants' risk and place these applicants into appropriate risk classes and even specific risk class sub bins.

The model assigns risk scores that can be compared for each applicant. The distribution of applicants within each risk class can be graphed. The results from the testing show that applicants in the Select risk class can be split into two further sub bins: a low risk Select and a high risk Select risk class.

The model was trained using the "hard decisions;" applicants with automatically declined medical conditions or with failing MCAS scores were removed, since those applicants are easy to underwrite and do not provide any useful information for the model. Applicants that were used to initialize and train the model provided information on how different medical conditions were underwritten, since the final underwriting decision and risk class were part of the input. The process of filtering out the easy cases created an applicant data set for future project use; the data set can be used to initialize the model to determine the coefficients.

Testing the risk scoring model using real data from John Hancock determined with what accuracy it can assign risk scores to LTCI applicants. The accuracy with which the model assigns risk scores ranges from 82.76% (for the entire population of applicants with four medical conditions) to 89% (for the entire population of applicants with five medical conditions).

The accuracy of the model improves when medical condition risk points are changed to reflect the underwriting procedure at John Hancock. The model keeps track of medical conditions that are often being placed higher or lower than what the underwriting guidelines

state. The model calls attention to these most misplaced medical conditions and underwriters

can determine whether or not the risk points or underwriting procedure needs to change.

References

"American Middle Class." (March 28, 2007). *Wikipedia*. March 30, 2007.
        <http://en.wikipedia.org/w/index.php?title=American_middle_class&oldid=118558032>

Arsenault, E., & Rackliff, N. (2006). *Expert System Design for Long Term Care Unsurance Underwriting.* A Major Qualifying Project Report, Worcester Polytechnic Institute, Worcester.

"BMI – Body Mass Index." (August 26, 2006). *Centers for Disease Control and Prevention.* March 2006.
        <http://www.cdc.gov/nccdphp/dnpa/bmi/adult_BMI/about_adult_BMI.htm>

Chen, W., Danson, R., Korch, M., & Moon, M. (2006). *Adaptive Risk Score Assignment Model for Long-term Care Insurance Underwriting.* A Report for the Research Experience for Undergraduates, Worcester Polytechnic Institute, Industrial Mathematics and Statistics, Worcester.

Curry, P. (2006, November 8). "How Credit Scores Work, How a Credit Score is Calculated.*"* November 15, 2006. <http://www.bankrate.com/brm/news/credit-scoring/20031104a1.asp>

"Expert System." (2007). *Oxford English Dictionary.* March 2007.
        <http://dictionary.oed.com/cgi/entry/50080431/50080431se1?single=1&query_type=word&queryword=expert+system&first=1&max_to_show=10&hilite=50080431se1>

Fair Isaac Corporation. (2000, March 3). "Credit Scoring 101." November 15, 2006.
        <http://www.ftc.gov/bcp/creditscoring/present/index.htm>

Fair Isaac Corporation. (2001-2006). "What's In Your Score." November 15, 2006.
        <http://www.myfico.com/CreditEducation/WhatsInYourScore.aspx>

Federal Trade Commission. (1998, March). "Facts for Consumers - Equal Credit Opportunity." Novemeber 15, 2006. <http://www.ftc.gov/bcp/conline/pubs/credit/ecoa.htm>

"History of LTCI, The." (1996-2007). *LTCInsurance.com*. March 30, 2007.
        <http://www.ltcinsurance.com/what-is-long-term-care-insurance.html>

John Hancock Insurance and Financial Services. (December 5, 2005). "John Hancock Study Leads To Refined Underwriting Guidelines And Increased LTCI Approvals." March 2007. <http://www.johnhancock.com/about/news/news_dec05.jsp>

John Hancock Insurance and Financial Services. (2007). "What Should I Know About LTC Insurance?" March 30, 2007.

<http://www.johnhancocklongtermcare.com/about/abo_jhltci.jsp?selection=abo_know&c
id=johnhancock&selectstate=MA>

Johnson, Alissa. (2007). "Life, Disability, and Long-Term Care Insurance." March 30, 2007.
<http://www.ncsl.org/programs/health/genetics/lifedisLTC2004.htm>

"Long Term Care Insurance for the Elderly and Seniors." (2001-2007). *Helpguide.* March 30,
2007. <http://www.helpguide.org/elder/long_term_care_insurance.htm>

 "Market for Long-Term Care Insurance, The." (2005). *National Bureau of Economic Research.*
March 30, 2007. <http://www.nber.org/aginghealth/winter05/w10989.html>

Appendix A – User's Guide

A.1. Preparing Data Set

This section will show the user how to create a Microsoft Excel file that contains all of the necessary information for the model. It explains what format and content the model expects. Some examples will be given along the way as well as a specific explanation of how to turn raw data from John Hancock into a format for use with the model.

A.1.1 Model Inputs

There are three model inputs that need to be arranged and formatted in Excel. The first two are similar. First the initializing population is a matrix that needs to have applicants down the rows and condition risk points across the columns. The final column of the initializing population should have the applicants' target risk score. Here is an example for three condition applicants:

	Condition 1	Condition 2	Condition 3	Target Risk Score
Applicant 1	255	117	83	262.5
Applicant 2	169	14	37	125
Applicant 3	235	178	222	287.5
Applicant 4	420	387	12	437.5
etc…				

Figure A.1 – Sample Initializing Population

The next input, the testing population, is a matrix just like the initializing population. It has applicants down the rows and condition risk points across the columns with the applicant's true risk score in the final column. Here is an example for three condition applicants:

	Condition 1	Condition 2	Condition 3	True Risk Score
Applicant 5	256	100	82	262.5
Applicant 6	100	25	32	125
Applicant 7	235	189	223	287.5
Applicant 8	446	375	100	437.5
etc…				

Figure A.2 – Sample Testing Population

The final input, testing condition indicators, corresponds with the same applicants from the testing population. Each row is an applicant and must correspond to the same applicant in that row from the testing population. This means that the testing population and the testing condition indicators must have the same number of rows. The difference between these two inputs is number of columns. The testing condition indicators input has John Hancock medical condition indicators across the columns and does not need a true risk score. Here is an example for the 3 condition applicants from Figure A.2:

	Indicator 1	Indicator 2	Indicator 3
Applicant 5	127	2	272
Applicant 6	607	31	4338
Applicant 7	600	689	113
Applicant 8	100	510	607
etc…			

Figure A.3 – Sample Testing Condition Indicators

A.1.2 Modifying John Hancock Database

To create a useable data set for the model from a raw John Hancock Database first perform the applicant filtering process as described in the "Long Term Care Insurance Underwriting …" report. This will remove applicants who are automatically declined based on medical conditions or MCAS scores. It will also remove applicants with comorbid conditions. During the applicant filtering process John Hancock medical condition indicators will be translated into condition risk points for use in the model.

After all of this preparation is finished, create an excel document to hold all the applicants and their important information. For an example of what prepared data will look like, see *Clean_Apps.xls*, which has been included on this CD. Use the order of the columns in *Clean_Apps.xls*. To save some time, save the Excel workbook into the directory with the model's Matlab files.

A.2. Using the Matlab Code

The next section of this user's manual will describe exactly how to use Matlab to run data through the model and produce results. It will explain how to import data from Microsoft Excel into Matlab. It will also provide the necessary commands to run the model. Finally, this section shows how to view and export the results.

A.2.1 Importing Data

Using an Excel workbook as described in section A.1.2, the data can be easily imported into Matlab. From within Matlab the first step is to navigate to the correct directory. Use the Current Directory window in Matlab to navigate to the directory that contains the data set and model code. Now, use the built in Matlab function "*xlsread*" to import the three inputs discussed in section A.1.1.

The *xlsread* function requires 3 arguments. The first is the filename of the workbook that it should import from. The second is the worksheet name, and the last is the range of cells to import from that worksheet. It looks like this:

xlsread('NameOfFile.xls' , 'Name of Worksheet' , 'Name of Range')

The three arguments need to be saved to variables. The names are arbitrary but need to be remembered. It is suggested that *InitPop*, *TestPop*, and *TestPopConds* be used. Now, import

half of your population as the initializing population and half as the testing population. (It is not necessary to do half and half. Any combination can be used.) Here is an example using the provided files:

InitPop = xlsread('CleanApps.xls', '3 Cond Apps', 'K2:N1473');

TestPop = xlsread('CleanApps.xls', '3 Cond Apps', 'K1474:N2546');

TestPopConds = xlsread('CleanApps.xls', '3 Cond Apps', 'C1474:E2546');

A.2.2 Running the Model

Running the model is very simple. Make sure the Current Directory window is pointing to the folder with the model's Matlab files. Next, type the following command (all on one line) into the Command Window using the inputs imported in section A.2.1.

[Coef, PercentRight, PercentHigh, PercentLow, ReportHigh, ReportLow] = ...

... MQPmodel(InitPop, TestPop, TestPopConds);

A.2.3 Viewing Results

There are six separate outputs of the model that can be viewed from within Matlab. To view them, simply type the name of the variable in the Matlab Command Window and press enter.

Coef	The model coefficients
PercentRight	Percentage of Testing Population applicants given the correct Underwriting Assignment
PercentHigh	Percentage of Testing Population applicants given an Underwriting Assignment too risky
PercentLow	Percentage of Testing Population applicants given the Underwriting Assignment too safe
ReportHigh	A report of the most misplaced John Hancock medical condtion indicators for applicants in PercentHigh
ReportLow	A report of the most misplaced John Hancock medical condtion indicators for applicants in PercentLow

Figure A.4 – Descriptions of Model Variables

A.2.4 Exporting Outputs

Any of the outputs viewed in section A.2.3 can be exported to Microsoft Excel by using "*xlswrite*". The *xlswrite* function only requires the name of the Excel file to save to and the variable of data to save. There are other optional inputs for *xlswrite*. Please see the Matlab help for more information. Here is an example of how to export a variable to Excel using *xlswrite*:

xlswrite('Most_Misplaced_Conditions.xls', ReportHigh);

Appendix B: Contents of CD

B.1 *Raw_Data.xls*

This Microsoft Excel workbook contains 5 worksheets. One for each of the one, two, three, four, and five medical condition populations as they were straight from John Hancock. Each tab contains applicant ID numbers, their medical conditions, their underwriting decision, and height and weight from a number of different sources.

B.2 *Condition_Mapping.xls*

This Excel workbook houses the Avon condition matrix that is used to assign condition risk points to John Hancock medical conditions. The other tab in this file, Mapper, shows the translation from John Hancock medical condition indicator to Avon condition matrix risk point.

B.3 *Clean_Apps.xls*

Clean_Apps.xls contains five tabs; one for each of the five populations. This data set is the remaining applicants after all the data filtering has been completed. *Clean_Apps.xls* is the file that should be used when training and testing the model.

B.4 *User Manual.doc*

This Microsoft Word document gives step by step instructions on how to use the model. It starts with how to create a data set and goes on to explain importing inputs. It explains how to run the model and export the outputs. *User Manual.doc* also shows examples from this analysis.

B.5 *MQPmodel.m*

This is the main Matlab file. It is used to run the model and calls *Initialize.m*, *TestModel.m* and *Reports.m*.

B.6 *Initialize.m*

Initialize.m is Matlab code that takes the initializing population and performs the least squares fit to determine the model coefficients.

B.7 *TestModel.m*

This Matlab file takes the coefficients from *Initialize.m* and tests them against the testing population. It sends the John Hancock medical condition indicators for misplaced applicants to *Reports.m*.

B.8 *Reports.m*

Reports.m takes John Hancock medical condition indicators for misplaced applicants from *TestModel.m* and compiles them into reports. Then, it sorts these reports so that the most misplaced medical condition is reported first.

B.9 *LongTermCareMQP.doc*

This is the report.

Appendix C – Sample of Raw Data from John Hancock

The example shows two condition applicants. Control number is each applicant's unique identifier. Decision made is accepted or declined for coverage. Key Health Condition Coding #1 and #2 are John Hancock medical condition indicators. Rate class is Preferred, Select, etc. TIP wt, ft, and in are the applicants weight, and height. There are other entry methods for height and weight but they are not shown in this example. For a more in depth look at raw data, open Raw_Data.xls from the CD.

Control number	Decision made	Key Health Condition Coding #1	Key Health Condition Coding #2	Rate Class	TIP wt	TIP ft	TIP in
#######	D	580	348	2			
#######	A	40342	40336	1	150	5	4
#######	A	120	84	2	113	5	6
#######	A	341	293	5	300	6	1
#######	A	339	484	2	128	5	4
#######	A	4667	638	2	192	5	11
#######	A	341	2161	2	165	5	5
#######	A	438	311	2	126	5	2
#######	A	2161	293	2	190	5	10
#######	D	4064	40335	2	190	5	7

Figure C.1 – Sample of Raw Data

Appendix D – Sample of Condition Mapping

This is an excerpt from the Condition Mapping table in Condition_Mapping.xls. Indicator is the John Hancock medical condition indicator. Condition Name and Description are the name of the medical condition and a description of when an underwriter should use this specific John Hancock medical condition indicator. Decision is the underwriting decision that should be made for each medical condition. Avon Condition and Severity are the medical condition indicators from the Avon medical condition matrix that coincide with a given John Hancock medical condition. Finally, score is the risk point associated with the medical condition.

Indicator	Condition	Description	Decision	Avon Cond	Avon Sev	Score
17	Alcoholism	Abstinence >18 months, normal blood studies, no complications	Select	8	5	209
18	Alcoholism	Abstinence <18 months, normal blood studies, no complications	Postpone 18 months	8	7	1000
19	Alcoholism	Alcohol abuse with detox admission, current abstinence >18 months	Select	8	5	209
20	Allergies	0	Preferred	10	1	0
21	Alzheimer's Disease	0	Decline	12	7	1000
22	Amaurosis Fugax	With diagnosis of diabetes	Decline	13	7	900
23	Amaurosis Fugax	2 or more episodes	Decline	13	7	900
24	Amaurosis Fugax	Asymptomatic, single episode, no cognitive or physical abnormalities, >12 months	Class1 w/90 day EP	13	4	355
25	Amputation	Due to trauma or congenital, independent, single limb, >12 months	Select - Class 1	16	5	268
26	Amputation	Due to disease	Decline	15	7	1000

Figure D.1 – Sample of Condition Mapping

Appendix E – Sample of Modified Avon Condition Matrix

The following figure is an excerpt from the Avon medical condition matrix. ID is the Avon medical condition identifier and Medical Condition is the name. Desired Severity 1 and 7 Score are, respectively, the lowest and highest risk points associated with a given condition. Relative Steepness describes fast the risk increases as severity of the condition increases. Finally, 1 through 7 are the risk points associated with the medical condition at all 7 severity levels.

ID	Medical Condition	Desired Severity 1 Score	Relative Steepness	Desired Severity 7 Score	1	2	3	4	5	6	7
2	Achilles Tendonitis- Refer to Bursitis and Tendonitis	25	STEEP	500	25	30	42	68	124	243	500
3	Acoustic Neuroma	75	STEEP	750	75	83	100	137	216	385	750
4	Acromegaly	1000	FLAT	1000	1000	1000	1000	1000	1000	1000	1000
5	Actinic Keratosis	0	FLAT	0	0	0	0	0	0	0	0
6	Addison's Disease (adrenocortical insufficiency)	100	STEEP	900	100	110	130	173	267	468	900
1	AIDS	1000	FLAT	1000	1000	1000	1000	1000	1000	1000	1000
145	AIDS Dementia	1000	FLAT	1000	1000	1000	1000	1000	1000	1000	1000
7	Alcohol Abuse	100	STEEP	1000	100	111	134	182	288	514	1000

Figure E.1 – Sample of Modified Avon Matrix

Appendix F – Final Presentation

Overview

To test the previously developed risk scoring model for Long Term Care Insurance underwriting using real data from John Hancock.

WPI

Project History

- Expert System Design (2005-2006)
  - Elizabeth Arsenault, Nicholas Rackliff
- Adaptive Risk Score Assignment Model (2006)
  - Wendy Chen, Rachel Danson, Mary Korch, Melissa Moon

WPI

Long Term Care Insurance

- What it is
  - Description
  - History
- Why it is needed
  - Protects family members
  - High cost of LTC services
- How it is triggered
  - Activities of Daily Living (ADLs)
- LTCI market

WPI

Original Data

- 67,466 Applicants
- JH Medical Condition Indicators
- MCAS Scores
- Weight and height
- Underwriter decisions
- Data Obscurities

6

WPI

Filtered Applicants

- Easy Cases
  - Failing MCAS scores
  - Automatic declined medical conditions
  - Zero point medical conditions
- Comorbidities
  - BMI > 35
- Incomplete Underwriting Information
  - Final decisions
  - Unknown medical condition indicators
- Underwriting decision can not be explained by database information

7

WPI

Used Applicants

No. of Conditions	Applicants Received	Useable Applicants	% Useable
1	16999	11371	66.9%
2	14505	5398	37.2%
3	10639	1648	15.5%
4	6680	506	7.6%
5	6399	89	1.4%
Total	55222	19012	34.4%

8

WPI

Defining Risk Class Bins

- Previously developed
- Arbitrary

9

Condition Mapping

- Mapped medical conditions in JH guidelines to AVON risk point matrix
- JH Underwriting Guidelines
  - 959 medical condition indicators
- AVON risk point matrix
  - 486 conditions
  - Up to 7 severities

10

Assigning Target Risk Scores

- Example of Applicants
  - Underwriting Decision of Select

	Cond 1	Cond 2	Cond 3	Sum of Cond Scores	TRS
Applicant 1	255	117	83	455	237.5
Applicant 2	233	14	37	284	212.5
Applicant 3	235	178	222	635	287.5
Applicant 4	278	205	12	495	262.5

11

The Model

- **Medical Condition Risk Points:** P_n

- **Model Coefficients:** b_n

- **Medical Condition Risk Score:**

$$CRS = b_0 + b_1 P_1 + b_2 P_2 + b_3 P_3 + b_4 P_4 + b_5 P_5$$

- **Model Assumptions**

WPI

Model Input

- Initializing Population
  - Medical Condition Risk Points
  - Target Risk Score

	Cond 1	Cond 2	Cond 3	TRS
Applicant 1	255	117	83	262.5
Applicant 2	169	14	37	125
Applicant 3	235	178	222	287.5
Applicant 4	420	387	12	437.5

WPI

Model Input

- Testing Population
  - Medical Condition Risk Points
  - True Risk Bin
  - Medical Condition Indicators

	Cond 1	Cond 2	Cond 3	TRB	JH 1	JH 2	JH 3
App 5	256	100	82	262.5	127	2	272
App 6	100	25	32	125	607	31	4338
App 7	235	189	223	287.5	600	689	113
App 8	446	375	100	437.5	100	510	607

WPI

Sample Results

- **Coefficients**

b0	8.5811
b1	0.9492
b2	0.0374
b3	0.0707

- **Bin Assignment**

Correct	92.22%
High	7.78%
Low	0.00%

- **Reports**

JH Condition Indicator	Total Misplaced
37	45
40336	36
446	19
4 way tie	5

15

Results

	2 Conditions		3 Conditions		4 Conditions		5 Conditions	
	Run 1	Run 2	Run 1	Run 2	Run 1	Run 2	Run 1	Run 2
b0	23.9010	23.3130	8.5811	6.4850	-19.696	3.9432	-30.565	-23.228
b1	0.8189	0.8189	0.9492	0.9548	1.0139	0.9543	1.1515	1.1161
b2	0.0696	0.0828	0.0314	0.0449	-0.0197	0.0667	-0.1273	-0.1139
b3	-	-	0.0707	0.0828	0.0908	0.1187	0.0682	0.0648
b4	-	-	-	-	0.0007	0.0048	0.0739	0.0966
b5	-	-	-	-	-	-	0.0351	0.0271
Right	96.84%	97.05%	92.22%	92.54%	90.38%	91.48%	81.93%	86.75%
High	2.84%	2.84%	7.78%	7.46%	9.34%	8.52%	18.07%	13.25%
Low	0.33%	0.11%	0.00%	0.00%	0.27%	0.00%	0.00%	0.00%
Total Apps	88.63%		84.25%		82.16%		89.00%	

16

Results

Model Placement of 3 Condition Applicants

(bar chart: Count of Apps vs Risk Score Range)
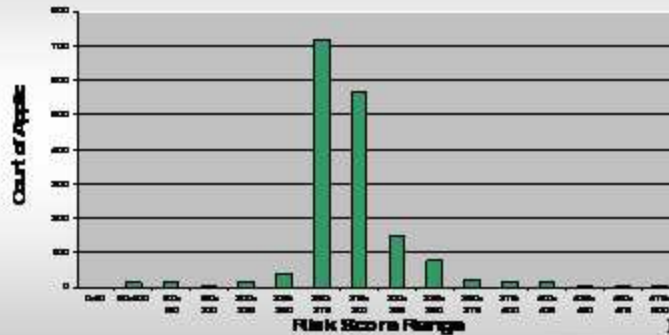
17

Future Work

- Reports
  - Automate process of updating Medical Condition Risk Points
- Model
  - Add comorbidities
  - Automate data filters
  - Deal with lifestyles and medications

18

WPI

Summary & Conclusions

- Real data is difficult to work with
- Eliminated easy decisions
- Worked with difficult decisions
  - Applicants give much information for model
- Created applicant data set for future projects
- Model results improve after changing risk points

19

WPI