The H.O.P.E. Calculator: Calculating the Ideal Hospital

The International Mathematical Modeling Challenge

Team US-7507 | April 2018

Executive Summary

Despite historical precedent, searching for a hospital is no longer as simple as locating the nearest one. Revolutionary medical breakthroughs, such as gene splicing and robotic surgery, are increasingly affecting daily lives, making the search for the "best" hospital more complex than ever. With this in mind, we build a compelling system that assists patients in choosing their "best" hospital by evaluating not just the performance of a hospital but the patient's preferences as well.

Recognizing that "best" can be both objective and subjective, we develop three metrics in which patients input personal information and a hospital to evaluate and receive three meaningful scores. Under the lens of an objectively "best" hospital, we adopt a two-pronged structure and create two metrics that mitigate a patient's likelihood of death. The first metric, the Risk Score, uses a Canadian rendition of a Hospital Standardized Mortality Score. This score evaluates the ratio of a hospital's number of deaths to a multivariate logistic regression-estimated number of in-hospital deaths in order to find the probability of death within each diagnosis group. It then employs a regression model to evaluate the severity of a disease based on sex and age. The second metric, the Success Score, evaluates an ideal four-parameter Burr Distribution among hospitals in the United States to determine whether the readmission rate places in a reasonable percentile. The last metric, Humanizing Hospitals, takes a subjective yet empirical approach. Evaluating a patient's preferences in the categories of *Comfort*, *Staff and Services*, and *Quality of Facilities*, the model adjusts the "threshold" of a Cauchy cumulative distribution function to determine whether or not a hospital is compatible with a patient. The modified-Cauchy procedure is extended to evaluate the individual statistics of a hospital, including cleanliness and surgical accidents. Additionally, we present the option of considering travel distance and personal budget if a patient has specific limits. To put it all together, we create an accessible and practical interface for our H.O.P.E. (Hospital of Preference Evaluation) calculator.

Table of Contents

1. Introduction

1.1 Background

In our world's rapidly changing biomedical technologies and revolutionizing healthcare systems, the decision to strategically choose an ideal hospital in needed times is becoming a pressing issue. The demand for more experimental research in the field is high, especially since the knowledge behind decision-making strategies is low[1]. Currently, systems that rank hospitals use criteria ranging from mortality rates, to patient happiness, to the quality of staff[2]. In this paper, we dive beyond the subjective ranking systems, gear the concept of choosing a "best" hospital towards the ability to fit a patient's' preference[3], and take steps towards assisting patients of our changing world.

1.2 Restatement of the Problem

Our ultimate goal is to assist patients in choosing their "best" hospital. We are tasked with building a metric off mortality and its factor. Then, we strive to add to our metric, considering factors beyond solely mortality. Lastly, we are asked to write a user-friendly memo for laymen to understand our mathematics.

1.3 Defining "Best"

The purpose of our model considers "best" as subjective to each patient. Therefore, our robust H.O.P.E. calculator uses both an objective and subjective approach. Our objective "best" examines whether hospitals allow patients to be healthier after their visits. This is the main drive behind the **Mortality Metrics** section. Our subjective "best" considers the experience behind a patient's visit and whether a hospital matches a patient's preferences. This is the main drive behind **Humanizing Hospitals** section.

---

[1] Fischer, Sophia. Understanding Patients' Decision-Making Strategies in Hospital Choice: Literature Review and a Call for Experimental Research. www.tandfonline.com/doi/full/10.1080/23311908.2015.1116758.
[2] Hospital Compare. Medicare, www.medicare.gov/hospitalcompare/About/Hospital-overall-ratings.html.
[3] "Where Are America's Best Hospitals?" U.S. News & World Report, U.S. News & World Report, health.usnews.com/best-hospitals.

2. Mortality Metrics

Simply considering mortality is an ineffective way to judge a hospital's worth—especially when there's a plethora of factors going into mortality rates from types of illnesses, to intake of patients, to patient demographics. In this section, we dive into the factors contributing to mortality and develop two metrics to measure a hospital's quality based on these factors—all with the goal of assisting patients with choosing a safe and risk-free hospital.

2.1 Specific Assumptions and Justifications

Assumption 1: Data regarding Hospital Standardized Mortality Rate is available at any hospital.
Justification 1: Due to lack of depth in existing data, assuming that the data is available allows our model to be as specific and robust as possible without the confinements of finding excessive data. The data we use is reasonably available.

Assumption 2: Individual data concerning the correlation between age and a specific disease and sex and a specific disease is available.
Justification 2: This data exists. However, having to find it would distract from the actual math within the model. It is easily accessible and can be incorporated into the model individually with any externally given data.

Assumption 3: Distribution of data across 19,000 United States hospitals is representative of the distribution of data across all hospitals.
Justification 3: This covers hospitals in the United States, and although it's limited to one country, it covers a wide range of hospitals—from poorer ones to more wealthy one.

Assumption 4: Diagnosis Groups[4] among all hospitals are standardized.
Justification 4: Many hospitals, and even different countries, have different labeling for diagnosis groups. Knowing the specific labeling is not relevant, but would expedite calculations.

Assumption 5: People over the age of 95 do not die solely because of a disease.
Justification 5: Stretching the age to 95 is already high, but data was available.

[4] **Diagnosis Group**: Category in which a hospital procedure belongs to

2.2 Approaching the Metric

Studies have found mortality, including standardized mortality measurements, to be a weak indicator of hospitals[5]. Therefore, our initial approach is to circumvent the faults of these indicators as much as possible and evaluate the various factors surrounding mortality.

We initially recognize that comparing and running through a database of hospitals may both be inefficient and unreasonable, so our strategy is to develop a framework surrounding mortality to output a score for a certain hospital.

We take a two-prong approach and break the metric down into two categories: *measuring risk* at a hospital and *measuring success* at a hospital—both specific to a patient's disease and a hospital. Measuring risk considers the the likeliness of death from a visit to a hospital, relating a patient's disease, age, and sex, as well as a hospital's death rate (which we elaborate upon later). Measuring success considers a patient's likeliness to be readmitted. Ultimately, we strive for two metrics that determine scores that establish success and risk within a hospital.

2.3 Methodology and Model Development

We break this section down into two parts: measuring risk and measuring success.

2.3.1 Measuring Risk—Calculating Hospital Standardized Mortality Ratio

After considering various models, we believe a rendition of the Standardized Mortality Ratio, the Hospital Standardized Mortality Ratio (HSMR), is an effective start towards measuring effectiveness within a hospital. This addresses not just mortality, but also analyzes age, sex, length-of-stay in a hospital, admission category[6], comorbidity, and whether a patient was transferred in from another institution. Our manipulation of the HSMR is inspired by the Canadian Institute for Health Information (CIHI)[7]. Ultimately, HSMR is the following:

[5] Shojania, Kaveh G., and Alan J. Forster. CMAJ : Canadian Medical Association Journal, Canadian Medical Association, 15 July 2008, www.ncbi.nlm.nih.gov/pmc/articles/PMC2443229/.
[6] **Admission Category**: Whether an admission to a hospital is urgent/emergent or elective.
[7] "Hospital Standardized Mortality Ratio." Hospital Standardized Mortality Ratio, Nov. 2016, www.cihi.ca/sites/default/files/document/hsmr_tech_notes_en.pdf.

$$\textbf{(1)}\ HSMR = \frac{Number\ of\ actual\ deaths\ in\ a\ diagnosis\ group\ while\ in{-}hospital}{Number\ of\ expected\ deaths\ in{-}hospital}$$

The first step in using HSMR is calculating the expected number of deaths at a hospital. This begins with evaluating, at each individual hospital[8], the probability of in-hospital death for a diagnosis group with its multivariate logistic regression of the following independent variables: age in years, sex, length-of-stay, admission category, comorbidity, and transfers.

In approaching this calculation, define the following variables:

β_0 = Regression intercept of the specific disease
β_A = Regression Coefficient of Age
β_S = Regression Coefficient of Sex
β_{LOS} = Regression Coefficient of Length-of-Stay
β_{AC} = Regression Coefficient of Admission Category
β_C = Regression Coefficient of Comorbidity
β_T = Regression Coefficient of Transfers

Ultimately, the probability of expected death from a diagnosis group within the hospital is calculated with a running multivariate logistic regression model for each variable in relation to the number of deaths and finding as such:

$$\textbf{(2)}\ P(D_i) = \frac{e^{(\beta_0+(A_P*\beta_A)+\beta_S+\beta_{LOS}+\beta_{AC}+\beta_C+\beta_T)}}{1+e^{(\beta_0+(A_P*\beta_A)+\beta_S+\beta_{LOS}+\beta_{AC}+\beta_C+\beta_T)}}$$

Where,
$P(D_i)$ = the probability of dying in-hospital from a specific disease within a diagnosis group
D_i = A diagnosis group
A_P = Age of patient.

Next, to find the number of expected deaths within a hospital, take the summation of all $P(D_i)$ within the hospital, given in this form:

$$\textbf{(3)}\ Expected\ Deaths = \sum(P(D_i))$$

[8] Referring to Specific Assumption 1

Finally, the HSMR is calculated with the following model:

$$\textbf{(1)}\ HSMR = \frac{Number\ of\ actual\ deaths\ in\ a\ diagnosis\ group\ while\ in-hospital}{Number\ of\ expected\ deaths\ in-hospital}$$

2.3.2 Measuring Risk—Risk due to Age and Sex

As patients search for a hospital, the two inherent key factors contributing to the risk factor of their disease is age[9] and sex[10]. These are contingent on the patient and not the hospital. Therefore, we consider these factors as a weight towards the ramifications of how risky the HMSR may be for a patient i.e. if a patient is a certain age and sex, they may be more prone to hospital-related risks as opposed to patients of another age and/or sex.

Using data on the relationship between age and sex[11] towards the risk of a disease, we establish a graph of age vs. number of deaths for a disease—one for each sex. Then, we establish a polynomial regression line to the 6th order (the most common and fitting trend for data sets) in order to have an age function[12], such as:

$$\textbf{(4)}\ D_M(A) = Predicted\ number\ of\ deaths\ for\ males\ at\ an\ age$$

$$\textbf{(5)}\ D_F(A) = Predicted\ number\ of\ deaths\ for\ females\ at\ an\ age$$

Where,

A=age in years.

Now, to find a severity score of a patient with a disease at a certain age, we evaluate the following function:

---

[9] "Ageing as a Risk Factor for Disease." Current Biology, Cell Press, 10 Sept. 2012, www.sciencedirect.com/science/article/pii/S0960982212008159.

[10] "Sex and Gender." National Institutes of Health, U.S. Department of Health and Human Services, 19 July 2017, newsinhealth.nih.gov/2016/05/sex-gender.

[11] Referring to Assumption 2

[12] We demonstrate this in **section 2.5**

$$\textbf{(6)}\ S(X,A) = \frac{D_X(A)}{\int_0^{95} D_X(b)\, db}$$

Where,

S(X, A) = function determining the severity rating of a disease specific to a patient's age and sex

X = Patient's sex

A = Patient's age in years

b = a dummy variable

To note, the integral expression in the denominator outputs the total number of deaths for a sex from the age of 0 to 95 (95 is used because it is a consistent age for data, although any age could be the upper bound for the integral).

2.3.3 Combining HSMR and Age and Sex Risk

The score from HSMR and the Age and Sex Risk is the following:

$$\textbf{(7)}\ Score_{Risk} = HSMR * S(X,A)$$

2.3.4 Measuring Success—Readmission Rates

In approaching success, we see the chances and rates of readmission as a key factor in whether a procedure or hospital visit is successful. In developing this portion, we use data from the *Centers for Medicare and Medicaid Services*[13] [14] of over 19,000 hospitals across the United States and their readmission rates. We are given an expected readmission rate, R_E, a predicted (or actual) readmission rate, R_P, and an excess readmission ratio, R_R. As such:

$$\textbf{(8)}\ R_R = \frac{R_P}{R_E}$$

Using these ratios across the 19,000 hospitals, we run this data through Easyfit[15], a statistical program, and find that a Four-Parameter Burr distribution is an ideal fit to this data. Through *Easyfit's* analysis, a Four-Parameter Burr distribution ranks 1st under a Kolmogorov-Smirnov test and a Chi-Square test and ranks 2nd under an Anderson-Darling test compared to other distributions—making it the ideal distribution (See **Appendix E** for details).

[13] "Datasets | Data.Medicare.gov." Data.Medicare.Gov, data.medicare.gov/data/hospital-compare.

[14] Based on their Readmission Reduction Data

[15] Technologies, MathWave. "EasyFit :: Distribution Fitting Made Easy." EasyFit - Distribution Fitting Software - Benefits, www.mathwave.com/easyfit-distribution-fitting.html.

Appendix A illustrates the Probability Density function (PDF) and graph. The more useful function is the Cumulative Distribution Function (CDF), which is the integral of the Probability Density function, given in the form:

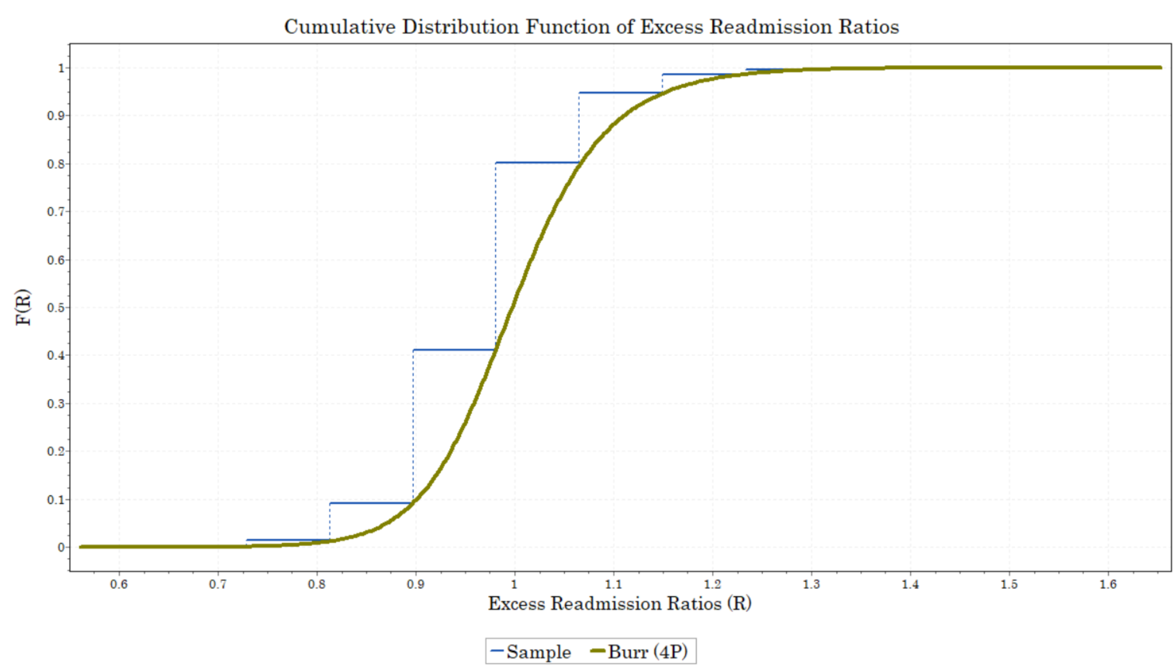
$$\textbf{(9)}\ F(x) = 1 - \left(1 + \left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{-k}$$

Parameters are given as follows:

γ = -1.5167
α = 59.099
β = 2.4962
k= 0.74744

Therefore, the function and graph for the CDF of Excess Readmission Ratios is:

$$\textbf{(10)}\ F(R_R) = 1 - \left(1 + \left(\frac{R_R+1.5167}{2.4962}\right)^{59.099}\right)^{-0.74744}$$



Ultimately, this function outputs the percentile at which a certain Excess Readmission Ratio is located.

The score from the success factor is:

$$\textbf{(11)}\ Score_{Success} = F(R_R)$$

2.3.5 Analyzing Risk and Success

Finally, we are left with two metrics—one that determine a hospital's percentile of success based on readmission rates and one that determines a score measuring a patient's risk going into a hospital at his or her age.

Starting with the Success Score, the percentile of a hospital gives patients a ranking of where the hospital places among other hospitals in terms of readmission. This is critical for two reasons. First, readmission has been found to be a consistent measure of a poor performing hospital if the rates are higher[16]. Second and more importantly, readmission rates is an object measure of how accountable a hospital is, ensuring that hospitals will be more effective and successful with their procedures[17].

With the Risk Score, we can determine how risky a hospital may be in relation to a hospital's historical precedent with deaths and a patient's status going into the hospital. Ideally, a patient wants to see a risk score of less than 1.

2.4 General Form

Metric 1A: Risk Score

$$Score_{Risk}(D_A, D_E, A, X) = \frac{D_A}{D_E} * \frac{D_X(A)}{\int_0^{95} D_X(b)\,db}$$

Where,
D_A= Number of deaths within a diagnosis group in a specific hospital in a year
D_E= Calculated expected number of deaths in a specific hospital in a year
A=Patient's Age in years
X=Patient's Sex
$D_X(A)$= A function (based on polynomial regression) of age that determines the number of how many people die at that age for the patient's specific disease based on their sex (each sex has a different function)

Metric 1B: Success Score

$$Score_{Success}(R_R) = 1 - \left(1 + \left(\frac{R_R+1.5167}{2.4962}\right)^{59.099}\right)^{-0.74744}$$

Where R_R = Readmission rate/expected readmission rate

16 "Hospital-Readmission Risk - Isolating Hospital Effects from Patient Effects | NEJM." New England Journal of Medicine, www.nejm.org/doi/full/10.1056/NEJMsa1702321.

17 "An Ounce of Evidence | Health Policy." An Ounce of Evidence Health Policy, blogs.sph.harvard.edu/ashish-jha/2013/02/14/the-30-day-readmission-rate-not-a-quality-measure-but-an-accountability-measure/.

2.5 Model Demonstration

Using data on the relationship between ages and sexes[18] towards the risk of a disease, we then find a regression for such data sets. For the case of model development, we use lung cancer as an example, although this methodology can be applied to any set of data relating age and sex to the death or severity of a disease[19]. Using Cancer Research UK[20] [21], we use data on lung cancer and the age and sex at which people perish from the disease (As shown in Appendix **B**). Then, from plotting the data, we find the 6th order polynomial which best fits the data in order to have a function of age and sex and the number of deaths to lung cancer. The graphs of the lines can be found in Appendix **C**.

Demonstration on a 64 year male:

$$S(M,A) = \frac{2*10^{-7}(A^6)-6*10^{-5}(A^5)+.0059A^4-.2811A^3+6.0615A^2-51.819A+111.46}{\int_0^{95} 2*10^{-7}(b^6)-6*10^{-5}(b^5)+.0059b^4-.2811b^3+6.0615b^2-51.819b+111.46\,db}$$

$$S(M,64) = \frac{2*10^{-7}(64^6)-6*10^{-5}(64^5)+.0059(64^4)-.2811x^3+6.0615(64^2)-51.819(64)+111.46}{\int_0^{95} 2*10^{-7}(b^6)-6*10^{-5}(b^5)+.0059*b^4-.2811*b^3+6.0615*b^2-51.819*b+111.46\,db} \quad S(M,64) =$$

$$\frac{493.28}{19449} = .023 = 2.3\%$$

The result of 2.3% is a very tangible result. This percent shows the risk of death depending on your age, sex, and disease. The lower the percent, the less likely one has a terminal case of the certain disease. An ideal S(X,A) would be zero, but this is highly unlikely as the score represents the percent of those with similar factors that have died. This is in no way a certain guarantee, as all cases are specific to each individual, but the percent provides a frame of where an individual stands within the demographic. This value would then be multiplied by an HSMR to output a score. The following table gives examples:

HSMR:	0.25	0.5	0.75	1	1.25	1.5
Risk Score:	0.00575	0.0115	0.01725	0.023	0.02875	0.0345

[18] Referring to Assumption 2
[19] Referring to Assumption **2**
[20] "Lung Cancer Survival Statistics." Cancer Research UK, 27 Sept. 2017, www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/survival.
[21] "Average Number of Deaths per Year and Age-Specific Mortality Rates per 100,000 Population, UK." www.cancerresearchuk.org/sites/default/files/cstream-node/deaths_crude_lung_M14.pdf.

3. Humanizing Hospitals

Mortality is an essential aspect of hospital performance. However, the value of several other hospital characteristics is ultimately a matter of preference.
In this part of the problem, we design a percentile-based model that evaluates the performance of a hospital based on patient preferences within the topics of comfort, staff and service, facilities, distance, and price. The result is a “percent match calculator” that generates the compatibility between any given hospital and patient.

3.1 Specific Assumptions and Justifications

Assumption 1: The distribution of hospital star rankings (government-led evaluations of overall hospital performance) is identical to that of any particular medical statistic.
Justification 1: The distribution of one overarching ranking distribution is likely to be representative of other medical criteria. Additionally, this is because the hospital star rankings evaluate and aggregate all medical criteria.

Assumption 2: Data needed from individual hospitals is accessible to the public.
Justification 2: The Center for Medicare and Medicaid services annually publishes data on a wide range of hospital characteristics. This data includes a large majority of US hospitals, but not all. Because of this, we find it reasonable that the small portion of the data that is not published can be assessed through communicating with the hospital itself.

3.2 Approaching the Metric

Since the late 1960s, hospital evaluation has been largely built upon the framework of Avedis Donabedian[22]. His approach, titled *The Donabedian Model*, consists of the consideration of three criteria: *structure*, *process*, and *outcome*. *Structure* contains the pre-existing infrastructure and staff of a hospital. *Process* describes the characteristics of the given care. *Outcome* refers to the overall result of the hospital’s actions.

These categories guide our selection of preference topics. Since mortality is heavily related to outcomes, we narrow our preference analysis to structure and process. Under these two categories, we select three topics that we felt adequately encompassed hospital characteristics: **comfort, staff and service,** and **facilities**.

[22] Frenk, Julio. “Avedis Donabedian.” www.who.int/bulletin/archives/78(12)1475.pdf.

These categories, described and elaborated upon below, shape our patient preferences model. Patients will give an importance rating (integers 1 to 10, inclusive) for each of these three categories in order to calculate compatibility with a hospital. Patients ideally should not give topics the same importance rating.

Topic	Description	Variables (all in percent)
Comfort	Evaluates the experiences of past patients	***r*** **=** patients who would definitely recommend the hospital ***c*** **=** patients who said their room and bathroom was "always clean" ***d*** **=** hospital bed occupancy
Staff and Service	Evaluates skill and consistency of staff procedures	***s*** **=** patients who experienced "serious complications" ***b*=** patients who experienced blood clots after surgery ***t*** **=** patients who experienced "accidental cuts and tears from medical treatment" ***i*** **=** patients who experienced bloodstream infection after surgery
Facilities	Evaluates financial position	***m*** **=** hospital operating margin

3.3 Methodology and Model Development

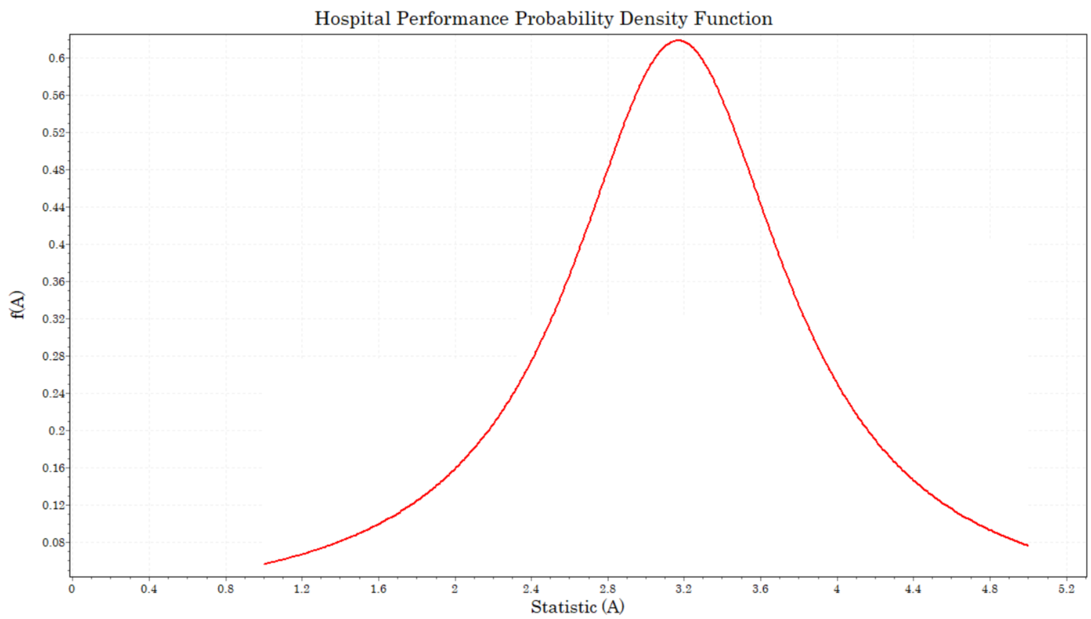
3.3.1 Establishing Distributions

To begin, we seek a general expression for a the percentile of a hospital's performance in a given statistic. The first step in this process is evaluating the distribution of hospital performance.

Using the statistical program EasyFit[23], we analyze the distribution of the *Center for Medicare and Medicaid Services* (CMS) 2017 hospital star rankings[24]. The software simulates and ranks the results of 59 different distribution tests and chooses a Cauchy distribution[25] as the greatest fit (see **appendix D** for details). This process also calculates the parameters for the probability density function, displayed and graphed on the next page.

[23] Technologies, MathWave. "EasyFit :: Distribution Fitting Made Easy." EasyFit - Distribution Fitting Software - Benefits, www.mathwave.com/easyfit-distribution-fitting.html.
[24] "2016-10-12." CMS.gov Centers for Medicare & Medicaid Services, 12 Oct. 2016, www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2016-Fact-sheets-items/2016-10-12.html.
[25] "Cauchy Distribution." From Wolfram MathWorld, mathworld.wolfram.com/CauchyDistribution.html.

$$\textbf{(1)}\quad PDF(A,A_0) = \frac{1}{0.62628\pi*[1+(\frac{A-A_0}{0.62628})^2\ ]}$$



3.3.2 Finding Percentiles

To find percentiles, we integrate Equation 1, resulting in our general cumulative density function (CDF):

$$\textbf{(2)}\quad \int PDF(A,A_0)\ dA = \int \frac{1}{0.62628\pi*[1+(\frac{A-A_0}{0.62628})^2\ ]}\ dA$$

$$\textbf{(3)}\quad \mathrm{CDF(A)} = \frac{1}{\pi}\arctan(\frac{A-A_0}{0.68628}) + \frac{1}{2}$$

As outlined in **Assumption 1**, we operate under the belief that the distribution of hospital star rankings is identical to the distribution of any one statistic. As a result, **this function acts as the fundamental scoring system with which we evaluate *each* statistic A about its center A_0**. This percentile curve is an ideal scoring system because it **1**: is bounded between 0 and 1, easing the standardization of multiple variables, **2**: remains effective after translations, and **3: heavily rewards points for surpassing A_0**. These properties are manipulated in the next section to develop personalized scoring.

3.3.3 Adjusting the Threshold

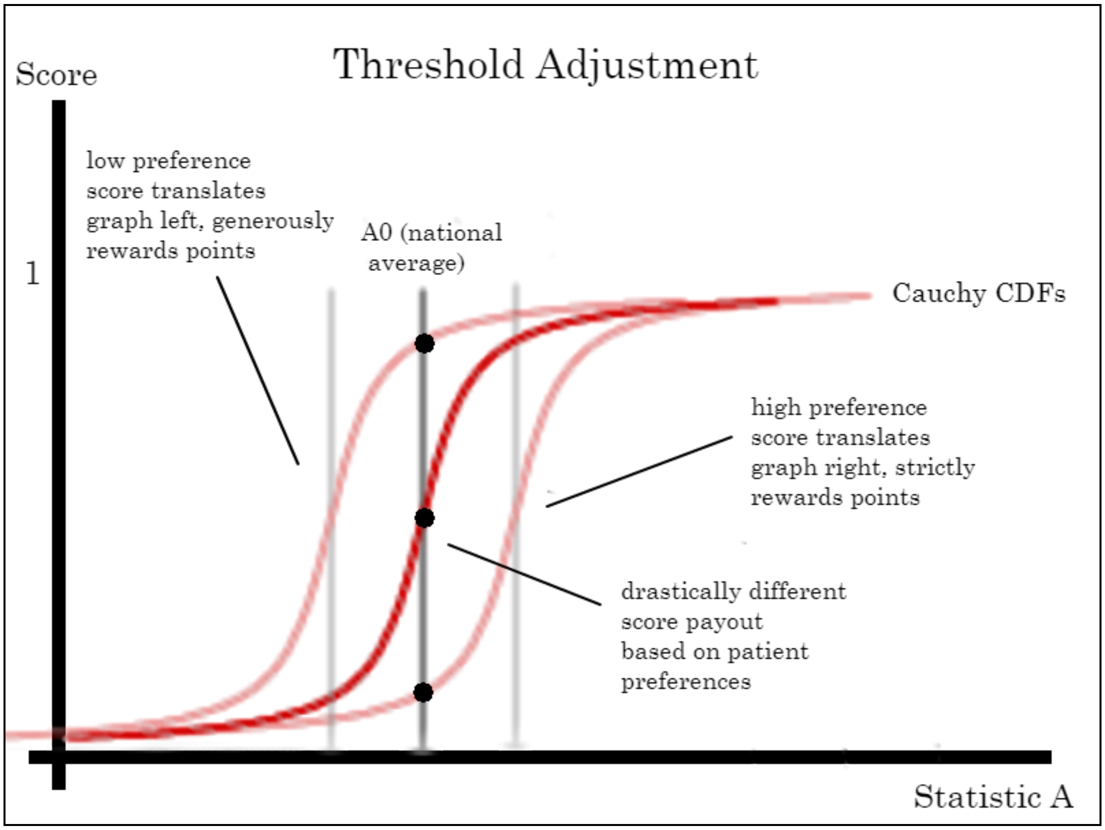
We approach the task of personalization with the goal of using a variable transformation of the Cauchy CDF. An ideal transformation would significantly alter the payout of a statistic depending on how highly a patient values its topic. Due to the drastic growth in value about A_0, we focus on its properties. More specifically, we analyze the effects of the different possible interactions between patient preference inputs (measured on a 10 point scale) and A_0.

Our preference model consists of **scaling** A_0 (the mean of a statistic), which we now refer to as a the **threshold**. Changing the threshold results in the desired meaningful differences in output for the same input.

Stated succinctly, altering a threshold results in a change in difficulty in score payout. For example, if a patient designates comfort an importance of 1, the threshold is lowered, allowing for a more generous score payout in that category. This is ideal; a patient with a low care for comfort is much more easily satisfied than one who highly values comfort.

Conveniently, this process also allots a similar amount of points to all hospitals above the threshold of a statistic. This results in comfort having a lessened impact on the patient's overall compatibility with a hospital. This property, as well as the ones mentioned above, work in reverse for designated importance values greater than 5.

Graphically, threshold adjustment works as follows:



But how, then, do we determine the extent to which a threshold is scaled? Fortunately, we plan on using strictly percent statistics, which is bounded by 1. We also know that our patient input can only differ from the average by at most five due to the 10 point scale. Thus, the factor ***h*** for which all percentages greater than ***n*** are floored to 100% is expressed as follows:

$$\textbf{(4)}\ n = \frac{100}{5h+1}$$

We then set **n** equal to 90 because we reason that most statistics will not exceed 90 percent. This results in $\boldsymbol{h} = \frac{1}{45}$. For ease of calculation, we round this value to a flat $\frac{1}{50}$. But what does this mean? For each change in patient preference rating (P_t , measuring distance from 5), we scale the threshold by 2%:

$$\textbf{(5)}\ A'_0(P_t, A_0) = A_0 * (\tfrac{1}{50}(P_t - 5) + 1)$$

3.3.4 The Evaluation Process

We now have all the pieces of our model ready for assembly. We use hospital statistics ***r***, ***c***, ***d***, and patient comfort preference score P_c in conjunction with national averages from *Becker's 230 Hospital Benchmarks*[26] to assemble the comfort score as follows.

$$\textbf{(6)}\ S_c(r,c,d,P_c) = S_{recommend} + S_{clean} + S_{bed}$$

We express each score as the sum of Cauchy distribution percentiles about a variable center $A_0^{'}$:

$$\textbf{(7)}\ S_c(r,c,d,P_c) = CDF(A_0^{'}r) + CDF(A_0^{'}c) + CDF(A_0^{'}d)$$

Using **Equation 2** yields:

$$\textbf{(8)}\ S_c(r,c,d,P_c) = \tfrac{1}{3} * [\tfrac{1}{\pi} \arctan \frac{r-(A_0^{'}(r,P_c))}{0.68628} + \tfrac{1}{\pi} \arctan \frac{c-(A_0^{'}(c,P_c))}{0.68628} + \tfrac{1}{\pi} \arctan \frac{s-(A_0^{'}(d,P_c))}{0.68628} + \tfrac{3}{2}]$$

Finally, substituting **Equation 5** into **Equation 8** yields:

$$\textbf{(9)}\ S_c(r,c,d,P_c) = \tfrac{1}{3} * [\tfrac{1}{\pi} \arctan \frac{r-(\frac{1}{50}*(P_{c-5})+1)(0.72)}{0.68628} + \tfrac{1}{\pi} \arctan \frac{c-(\frac{1}{50}*(P_{c-5})+1)(0.74)}{0.68628} + \tfrac{1}{\pi} \arctan \frac{s-(\frac{1}{50}*(P_{c-5})+1)(0.64)}{0.68628} + \tfrac{3}{2}]$$

In this case, the overall comfort score S_c is scaled by $\frac{1}{3}$ to reduce the output of the three modified Cauchy CDFs to that of one modified Cauchy CDF. This procedure, as well as the steps above, are then replicated for each of the three topic scores (comfort, staff and service, and facilities). The results are simply summed. Metric 2 can be viewed in its entirety in **Section 3.4**.

[26] Ellison, Ayla, and Jessica Kim Cohen. "230 Hospital Benchmarks | 2017." Becker's Hospital Review, www.beckershospitalreview.com/lists/230-hospital-benchmarks-2017.html.

3.3.5 Calculating Percent Compatibility

Based on the minimum score and score range of our model, we build a simple compatibility calculation:

$$\textbf{(10)} \quad \%_{match}(S) = \frac{Score - S_{minimum}}{\frac{1}{100}(S_{maximum} - S_{minimum})}$$

$$\textbf{(11)} \quad \%_m(S) = \frac{S - 0.912278}{0.0104699}$$

Alternatively,

$$\textbf{(12)} \quad \%_m(S) = 92.82559 * (S - 0.912278)$$

We calculate the minimum output of our model by inputting the maximum preference scores (10,9,8) and either 0% or 100% for each statistic. The same process is carried out in reverse in order to find the maximum output. Because both scenarios are implausible, most hospitals will score somewhere in the 30 to 50 percent match range. We find this to be appropriate due to the complex inner workings of hospitals; a perfect hospital does not exist. Regardless, percent match presents a prospective patient with a reliable, easy to understand method of comparison based on his or her innate and unchanging personal preferences.

3.3.6 Incorporating Price and Distance

Price and distance are immensely important factors in deciding the best hospital for an individual. In 2012 and 2013, 52 percent of patients chose the nearest hospital to them[27]. To take this into account, we propose the following proportional relationship between price and distance and prefered price and distance:

$$\textbf{(13)} \quad Score_{price,\, distance} = \frac{P_p}{price} * \frac{P_d}{distance} * S \;\; for \;\; P_p < price \; and \;\; P_d < distance.$$

This relationship results in a hospital score rapidly approaching 0 if a hospital exceeds the patient's preferred budget or distance. This score can also be inputted into **Equation 12** to generate an altered percent match. Additionally, this model ensures that both a hospital's score and percent match with a patient remain constant when within budget and target distance.

[27] Moscelli, Giuseppe, et al. Regional Science and Urban Economics, North-Holland, Sept. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC5063539/figure/f0015/.

3.4 General Form

Metric 2: Humanizing Hospitals

$$\%_m\,(P_c,\ P_s,\ P_f, r, c, d, s, b, t, i, m,\ P_P, P_d, distance,\ price) =$$

$$\left(\frac{P_p}{price} * \frac{P_d}{distance} *\right.$$

$$\left(\frac{1}{3} * \left[\frac{1}{\pi}\arctan\frac{r-(\frac{1}{50}*(P_{c-5})+1)(0.72)}{0.68628} + \frac{1}{\pi}\arctan\frac{c-(\frac{1}{50}*(P_{c-5})+1)(0.74)}{0.68628} + \frac{1}{\pi}\arctan\frac{d-(\frac{1}{50}*(P_{c-5})+1)(0.64)}{0.68628} + \frac{3}{2}\right] +\right.$$

$$\frac{1}{4} * \left[\frac{1}{\pi}\arctan\frac{s-(\frac{1}{50}*(P_{s-5})+1)(0.009)}{-0.68628} + \frac{1}{\pi}\arctan\frac{b-(\frac{1}{50}*(P_{s-5})+1)(0.0053)}{-0.68628} + \frac{1}{\pi}\arctan\frac{t-(\frac{1}{50}*(P_{s-5})+1)(0.014)}{-0.68628} + \frac{1}{\pi}\arctan\frac{i-(\frac{1}{50}*(P_{s-5})+1)(0.1)}{-0.68628} + 2\right]$$

$$\left.\left. + \left[\frac{1}{\pi}\arctan\frac{m-(\frac{1}{50}*(P_{f-5})+1)(0.034)}{0.68628}\right]\right) - 0.912268\right) * \ 92.825559$$

Note: if price and distance are not of concern, or if a hospital is within budget and the desired distance, set $\frac{P_p}{price}$ *and* $\frac{P_d}{distance}$ *to one.*

3.5 Model Demonstration

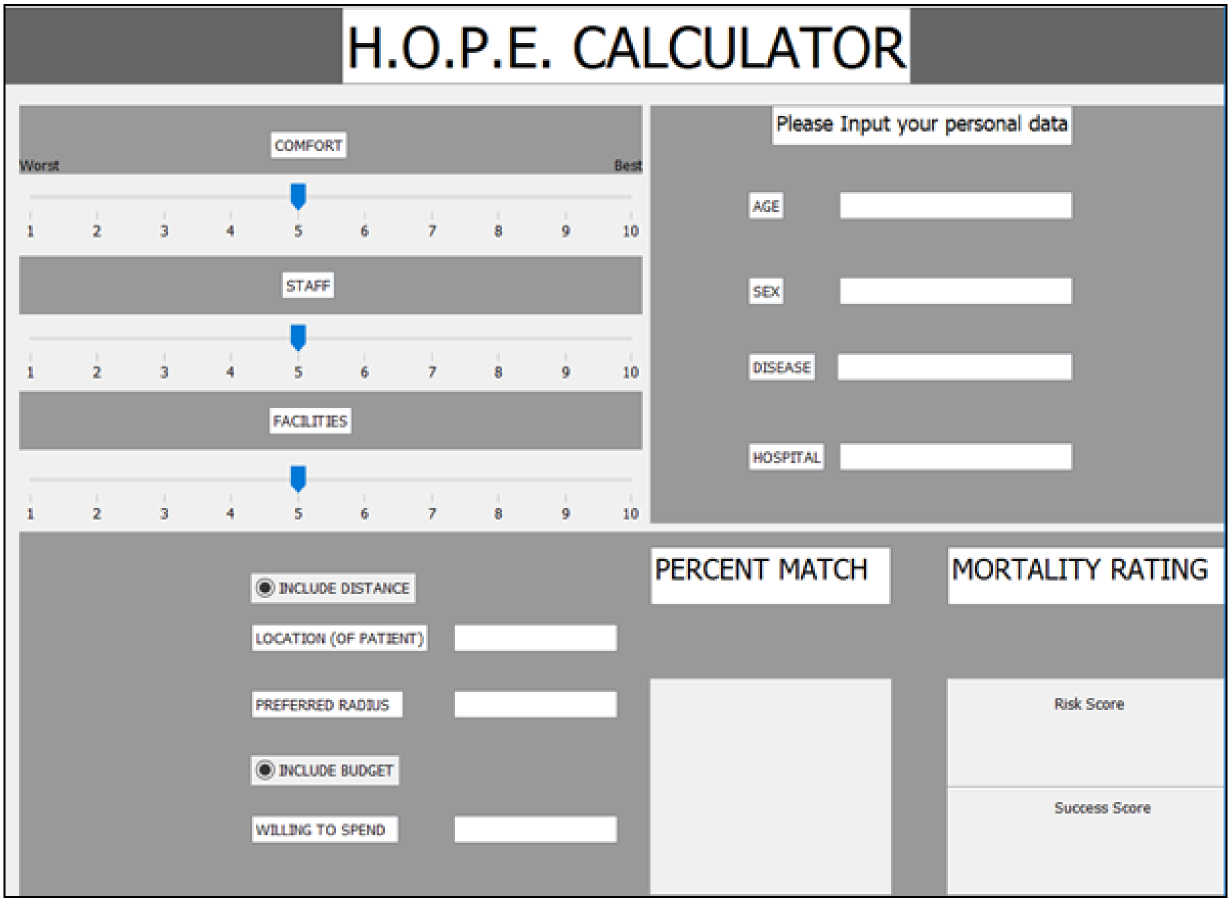
Consider the decision of the following patient. They are torn between Hospital A and Hospital B, which are both within their budget and desired location. The patient designates **comfort an importance of 9, staff and service an importance of 7, and facilities an importance of 2**.

Statistic	Hospital A	National Average	Hospital B
r	0.5	0.72	0.7
c	0.4	0.74	0.8
d	0.6	0.64	0.6
s*	0.009	0.009	0.008
b*	0.004	0.0053	0.004
t*	0.013	0.014	0.012
i*	0.1	0.1	0.1
m	0.04	0.034	0.02
% Match	**45.999582**		**53.241024**

**Note: In these categories, a lower percent is ideal.*

As expected, Hospital B matches the preferences of the patient better! Despite having half the operating margin of Hospital A, Hospital B excelled in the topics that the patient valued the most, giving it the winning edge.

4. H.O.P.E. Calculator Interface



When using this calculator, a patient uses three individual sliders to indicate how they value the topics of comfort, staff, and facilities. Additionally, patients input their age, sex, and disease. Through a checkbox, patients are then presented with the opportunity to include distance or budget (or both) in the calculator's evaluation. Finally, the patient inputs the hospital that they are evaluating. The calculator then displays the results of **Metric 1A, Metric 1B, and Metric 2.**

5. Reflection

5.1 Strengths and Weaknesses

Strengths

- Most current models do not take into account the personal preferences of the patients. Our model focuses on finding the best, personal match by allowing patient input into how factors will be scaled.
- Each person's version of our model is distinct to them.
- Our model accounts for robust diversity in patients. No matter demographic, disease, or preferences, out model finds a unique fit for the specific patient.
- Our model allows patients to directly compare different aspects of the hospital, it does not just create some generic one score.
- Our model redefines what the "best" hospital is. While most hospital ranking systems rely on classical conventions, our model mixes the conventional with personal, creating a more holistic rank.
- Our inputs and outputs are accessible and easy to comprehend. It its essential for patients to understand what the ranking means and why the hospital is ranked that way. Our system of breaking down rankings allows patients to understand how we rank hospitals, and use their own preferences to aid in their decision.

Weaknesses

- Our model depends highly on acquired data. It assumes that all the necessary data will be readily available.
- Our model does not integrate mortality and quality factors, and thus does not create one overall score. We have seperate scores, and the lack of an overarching score makes comparing hospitals more difficult.

5.2 Further Work

- We could use more detailed existing data to refine the model. This would allow for easier access to those using the model and more precise rankings.
- Ideally, the model would be linked to a database of every hospital. This would allow our model to not just score hospitals, but to provide the patient with the best possible hospital match.
- We would dive into various sorting or searching algorithms to save patients the expense of inputting hospitals, and instead, we'd be able to search for the most compatible hospital.
- After having a more in-depth model, we would work to combine our model more efficiently, in order to output one easily readable score.

6. User-Friendly Memo

Dear Prospective Users of the H.O.P.E. Calculator:

We are extremely glad that you have found our up-and-coming application designed to streamline your search in choosing the "best" hospital. Unlike many other hospital ranking systems, the H.O.P.E. Calculator considers both an objective and subjective lens to assist you in your search.

An acronym for Hospital of Personal Evaluation, H.O.P.E. requires a list of inputs from you. First, we would like to know more about you and your case: your age, sex, and the disease or sickness for which you need care. Then, we would like three values that describe how you value three hospital characteristics: comfort, staff and services, and the quality of facilities. To get the most out of the model, we'd recommend you don't have any factor with the same rating score. Lastly, if your budget or willingness to travel is a factor, we include a checkbox in case you have limits on your budget or travel distance.

While using the H.O.P.E. calculator, you may wonder how the application fundamentally works. Essentially, we have three metrics.

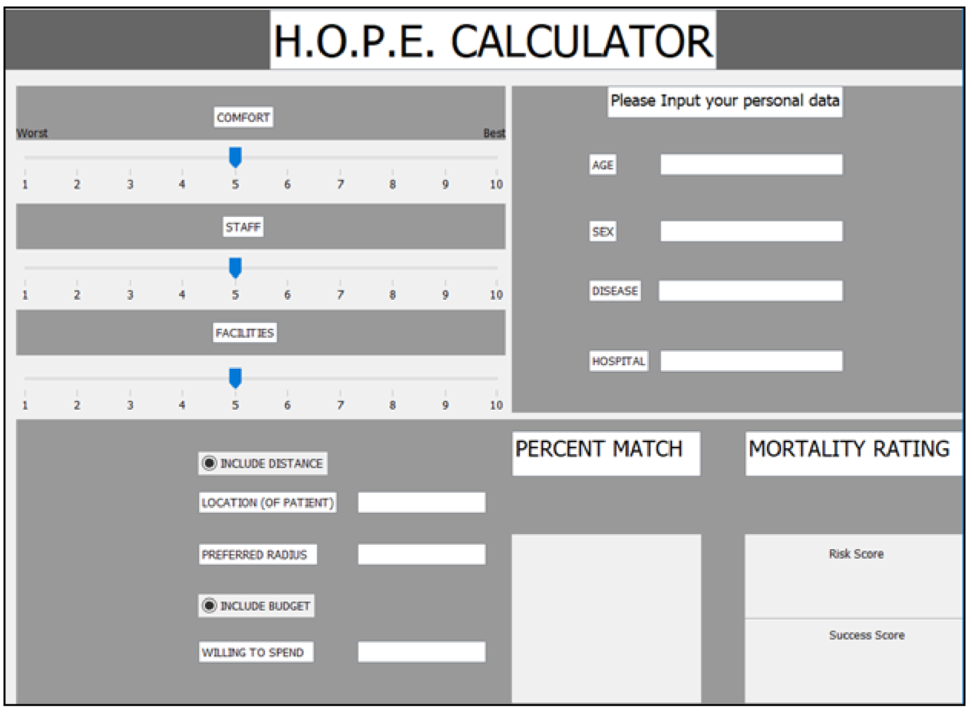
Our first and second metric considers a more objective approach.

The first metric—the risk score—calculates the rate at which deaths have occured during hospital procedure in relation to how many deaths a hospital expects. This means that if a hospital has less deaths than they expect, you can expect better treatment from your disease. The risk score then factors in your age and sex to see if you are at higher risks for faults. We then multiply the scores to see if a hospital is risky for you. We recommend you find a hospital with a score less than 1.

The second metric—the success score—considers how well a hospital performs using data on whether patients tend to be readmitted. We analyzed 19,000 hospitals across the United States to get an idea of the distribution of how hospitals tend to perform. We then take the hospital you want us to analyze and output a percentile of where the hospital places among other hospitals. We recommend you search for hospitals in lower percentiles.

The third metric examines a more subjective approach to determine the “best” hospital for you. In this Humanizing Hospitals metric, the math takes into account the rating score of each of the three categories (comfort, staff and services, and facilities), and determines if the hospital that you are evaluating would meet your preferences based on the hospital’s relation to others. For example, if you score categories low, more hospitals would be deemed satisfactory for you, while if you score categories high, less hospitals would meet your standards. In the end, a percent match is outputted to see how compatible you would be with your desired hospital, and we recommend as high of a percent match as possible.

As you use our application, you have the full flexibility to consider any hospital and compare hospitals to see if they fit your qualifications of “best”. We have attached an image of our calculator to assist you in visualizing how your could use our newly created tool. We *hope* for the *best* in your search for a hospital.



Best Regards,
Team US-7507

p.s. We are a start-up group looking for angel investors or generous donors. You may see that our interface is not yet exciting, interactive, nor aesthetically pleasing.

7. Appendices

Appendix A—Probability Density Function of Excess Readmission Ratios

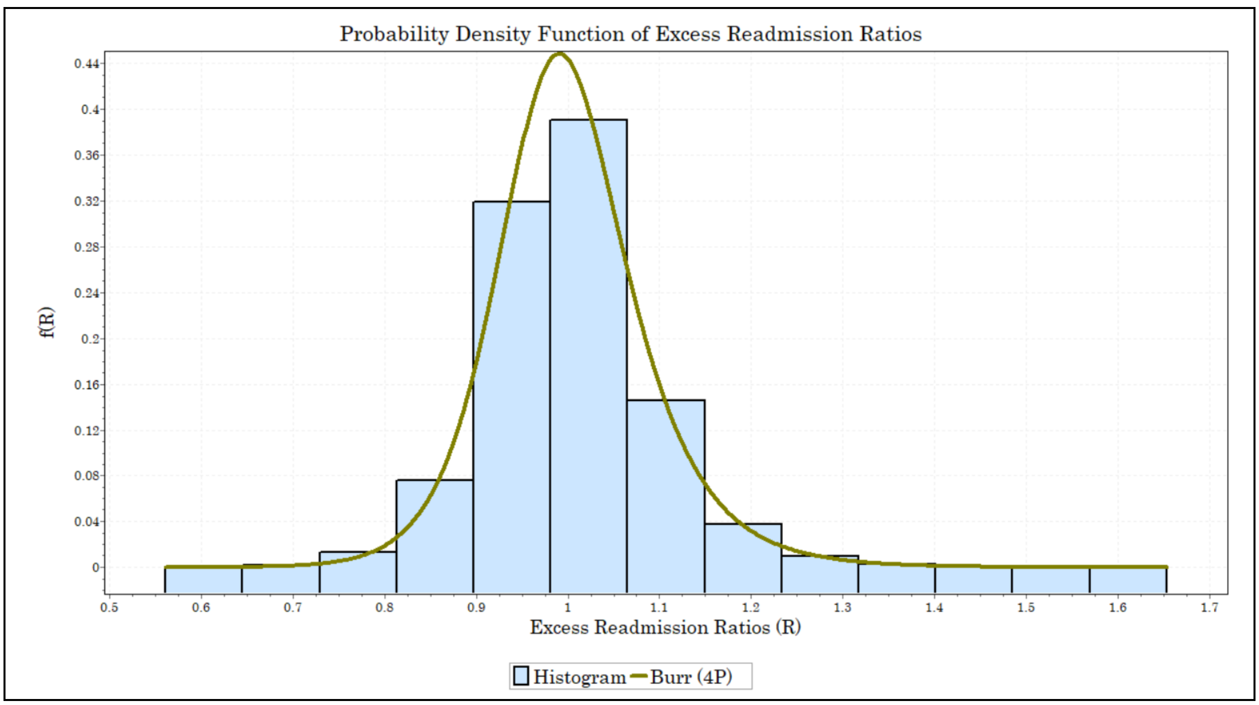
$$f(x) = \frac{\alpha k\left(\frac{\chi-\gamma}{\beta}\right)^{\alpha-1}}{\beta\left(1+\left(\frac{\chi-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}}$$

Parameters are given as follows:

γ = -1.5167
α = 59.099
β = 2.4962
k= 0.74744

This makes the Probability Density Function:

$$f(R_R) = \frac{59.099k\left(\frac{R_R+1.5167}{2.4962}\right)^{59.099-1}}{2.4962*\left(1+\left(\frac{R_R+1.5167}{2.4962}\right)^{59.099}\right)^{0.74744+1}}$$



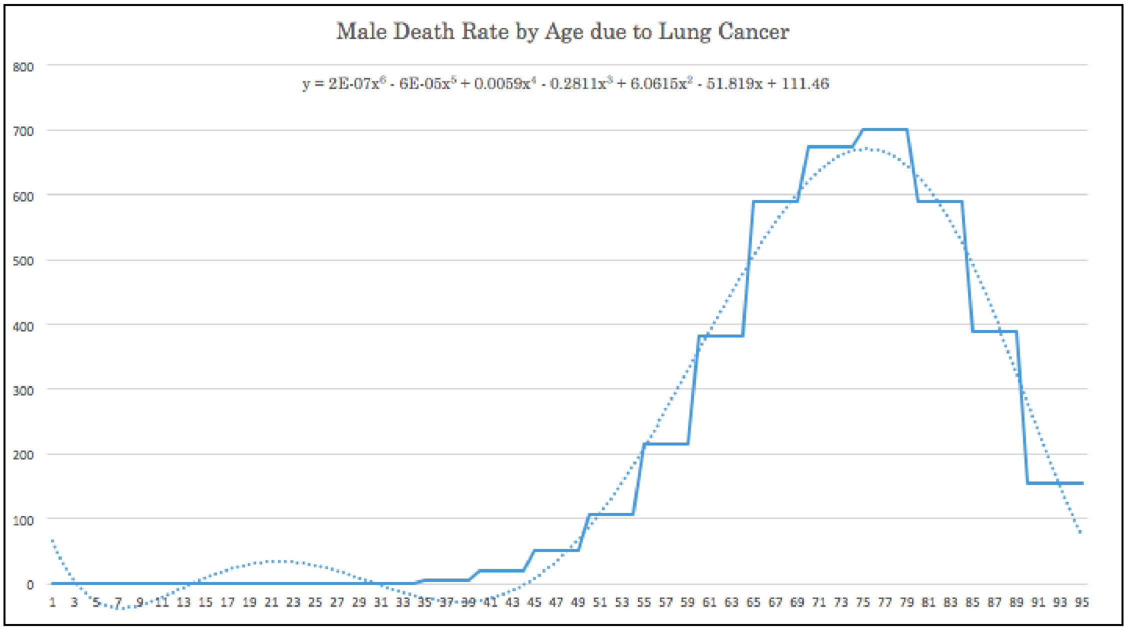
Appendix B—Lung Cancer and its Mortality Rate

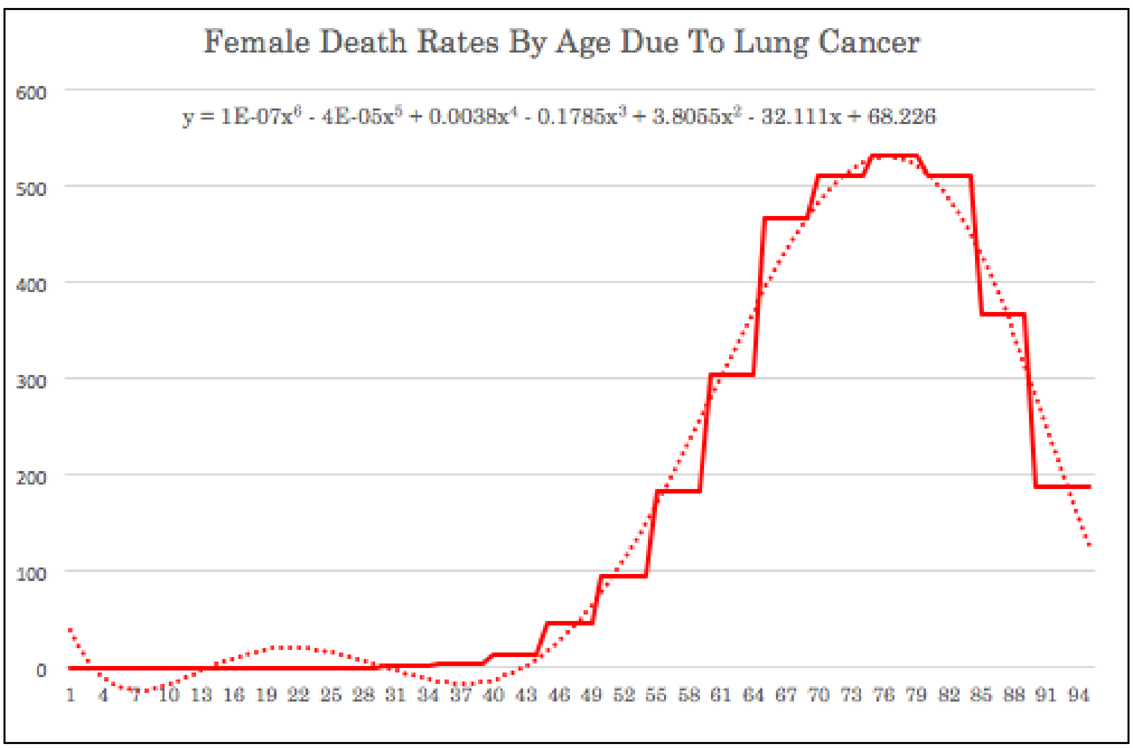
Lung Cancer (C33-C34): 2012-2014

Average Number of Deaths per Year and Age-Specific Mortality Rates per 100,000 Population, UK

Age Range	Male Deaths	Female Deaths	Male Rates	Female Rates
0 to 04	0	0	0.0	0.0
05 to 09	0	0	0.0	0.0
10 to 14	0	0	0.0	0.0
15 to 19	0	0	0.0	0.0
20 to 24	1	1	0.1	0.0
25 to 29	3	3	0.1	0.1
30 to 34	8	10	0.4	0.5
35 to 39	27	22	1.4	1.1
40 to 44	96	68	4.3	3.0
45 to 49	252	228	10.9	9.6
50 to 54	535	474	24.9	21.6
55 to 59	1079	914	58.1	48.0
60 to 64	1914	1527	109.8	84.1
65 to 69	2969	2335	176.2	131.1
70 to 74	3379	2562	279.9	190.9
75 to 79	3510	2666	367.8	234.1
80 to 84	2953	2551	453.6	283.6
85 to 89	1947	1831	564.4	309.4
90+	777	937	521.2	245.6
All Ages	19449	16129	61.6	49.5

Appendix C—Graphs of Rate of Death to Lung Cancer





Appendix D: Hospital Star Ranking Distributions

#	Distribution	Kolmogorov Smirnov		Anderson Darling	
		Statistic	Rank	Statistic	Rank
1	Beta	0.25811	41	3.659	53
2	Burr	0.18957	14	0.63077	29
3	Burr (4P)	0.1845	8	0.57976	20
4	Cauchy	0.18461	9	0.29477	2
5	Chi-Squared	0.35717	48	0.368	3
6	Chi-Squared (2P)	0.36832	49	0.42699	4
7	Dagum	0.1893	13	0.54213	10
8	Dagum (4P)	0.17736	4	0.6228	27
9	Erlang	0.3116	46	0.81889	41
10	Erlang (3P)	0.18168	6	0.58174	22
11	Error	0.17794	5	0.59121	24
12	Error Function	0.899	59	17.36	59
13	Exponential	0.39954	51	0.56611	17
14	Exponential (2P)	0.33106	47	6.1884	56
15	Fatigue Life	0.26022	42	0.69947	36
16	Fatigue Life (3P)	0.1918	18	0.57998	21
17	Frechet	0.44565	55	0.69444	35
18	Frechet (3P)	0.22931	33	0.54872	11
19	Gamma	0.21384	26	0.91843	43
20	Gamma (3P)	0.19893	22	0.55656	13
21	Gen. Gamma	0.22608	31	0.66598	34
22	Gen. Gamma (4P)	0.18287	7	0.63943	31
23	Gen. Pareto	0.30643	45	2.8414	50
24	Gumbel Max	0.23593	36	1.3494	47
25	Gumbel Min	0.22177	29	0.8892	42

Appendix E: Readmission Rate Distribution

#	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.04156	23	19.977	24	128.72	22
2	Burr	0.0094	2	0.80846	3	7.0426	2
3	Burr (4P)	0.00767	1	0.62701	2	6.5623	1
4	Cauchy	0.06655	33	47.772	33	595.27	35
5	Chi-Squared	0.61779	56	2296.5	56	62420.0	49
6	Dagum	0.0117	5	1.1255	6	9.5107	5
7	Dagum (4P)	0.01182	6	1.0701	4	9.6885	6
8	Erlang	0.04441	25	21.455	25	142.19	25
9	Erlang (3P)	0.041	21	19.091	22	122.01	18
10	Error	0.05584	30	17.934	16	154.04	27
11	Error Function	1.0	57	2.1354E+5	57	N/A	
12	Exponential	0.54145	52	1854.9	52	39680.0	47
13	Exponential (2P)	0.45241	50	1452.5	50	14602.0	43
14	Fatigue Life	0.03773	14	18.2	18	120.27	17
15	Fatigue Life (3P)	0.03783	15	18.207	19	119.55	16
16	Frechet	0.08064	36	81.556	35	N/A	
17	Frechet (3P)	0.10243	42	122.52	39	N/A	
18	Gamma	0.04277	24	21.684	26	140.84	24
19	Gamma (3P)	0.03888	18	18.781	20	122.31	19
20	Gen. Extreme Value	0.0339	9	42.512	31	N/A	