Copyright © 2014, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Acknowledgement

Laurent El Ghaoui and Don Wroblewski, my advisors.
University of California, Berkeley

Engin296 Capstone Project
Team: Kevin Johnson, Rie Kawauchi, Emilie de Longueau Saint Michel, Thibault Duchemin, Rendel Rieckmann

Health Solutions Assessment

Author: Kevin Johnson, BS

Supervisor: Laurent El Ghaoui, Ph.D.
Don Wroblewski, Ph.D.

May 14, 2014
Abstract

A great number of people in India lack access to good health care. I worked on two connected projects which each attempt to address this problem. The former project, called the Indian Health Kiosk project, involved building health “kiosks” in India which sought to provide low-cost health care to people in rural India. This project was terminated after six months due to unforeseen circumstances, leaving my team to create a new project.

Our second project attempted to systematically evaluate the effectiveness of projects which attempt to provide health care in India. For this project, called Health Solutions Assessment, we designed a simulation of India which can be used to model health care solutions. The simulation generates a population representing India, applies the effects of a solution to that population, and analyzes the results. It can be tuned and evaluated in its own right using data about previous healthcare solutions to predict the impact of new healthcare solutions. This analysis can be used to identify whether proposed solutions are worth investing in, and this information can be shared publicly to provide a platform for potential investors to learn about and discuss various proposed solutions. Due to time limitations, the simulation is not yet complete, but provides a prototype that demonstrates the potential value of the idea.

1 Introduction

As a developing nation with an immense population without substantial wealth, India faces many health struggles. It lags behind the rest of the world in many ways; there is a death of facilities to treat people [1], and a shortage of doctors [2]. Its life expectancy is a full five years behind the global average of 70 years [1]. However, recently, the healthcare industry has seen substantial developments; spending has increased by approximately 20% per year [3]. It is an area rich with opportunity for entrepreneurs and companies – both those who seek to do good in the world by aiding the large number poor and vulnerable people, and those who seek to make a profit by tapping into the immense potential market that India offers thanks to the size of its population. There have been many attempts to address these problems and provide greater access to health care to people who lack it, particularly in rural areas. The success achieved by these health solutions varies wildly.

We originally worked on a capstone project which would try to address these inequalities. This was called the Indian Health Kiosk project, and was done in collaboration with Professor Dhrubes Biswas from
the Indian Institute of Technology in Kharagpur. Our objective was to design a system of “kiosks” which would provide low-cost healthcare to the many people in rural India who lack access to good healthcare. However, this project was terminated after six months due to circumstances beyond our control, and we were unable to use any direct work from the project.

However, we were able to salvage some value from that attempt. While working on that initial attempt, we learned that it is difficult to judge whether a given health solution will succeed. As a result, we began work on a new project which would address that problem, called Health Solutions Assessment. The goal of this project is to build a tool to evaluate the effectiveness of a given health solution before it is implemented.

To do that, we create a model of India which could react appropriately to arbitrary solution designs. The model consists of a simulation which creates individual people based on the number of people in India and attempts to make them react realistically to new solutions. The model can be validated to ensure it is accurate by comparing its analysis to a more traditional analysis done on a given solution.

We do not present a simulation which is a finished product on its own. Time constraints imposed by changing projects midway through the year prohibited us from seeing the project to its logical conclusion. Instead, the simulation as it stands now is intended to be a proof of concept. The goal is to demonstrate that the fundamental idea of simulating India in this way is sound and worthy of further development.

2 Literature Review

A primary source of guidance in the development of the Health Solutions Assessment project came from interviews with experts in the field at berkeley. We interviewed Ms. Heather Lofthouse from the Blum Center, Ms. Karen Oppenheimer from World Health Partners, Ms. Clair Brown from the Institute for Research on Labor and Employment, and Ms. Caricia Catalani from InSTEDD. The feedback we received from these interviews was remarkably consistent: they thought that the simulation idea was clever and potentially valuable, but difficult to implement. They also confirmed that it was an original idea.

2.1 Existing Health Care Solutions

2.1.1 General Hospitals

Hospitals in India generally provide high quality service. In fact, India has become a growing destination for “medical tourism”, a phenomenon where people from other countries will travel to India to take advantage
of the high quality services offered there [4]. However, hospitals are inaccessible to many Indians, both due to the monetary costs of the hospitals and the monetary and time costs of transportation. Traveling to a hospital in India can be a multi-day affair, which is a long time to be away from work. That alone can prevent people who live based on the money they make from taking the time to get care they need.

Yet that pales in comparison to the cost of a hospital visit. According to [5], the typical cost for a routine office visit is $16, or about 1000 rupees. Though that is cheap by U.S. standards, that is prohibitively expensive for many in India. As a result, many people go entirely without medical care.

Hospitals can served the urban and moderately wealthy parts of India very well, and they provide general purpose care that is difficult to match. However, their price and location are inadequate to meeting the demands imposed by the large numbers of poor, rural people.

2.1.2 Traditional Medicine

In the absence of care from modern medicine, it should not be surprising that many people in India would turn to alternatives. In fact, it is estimated that as much as 80% of India uses Ayurveda, an ancient traditional medicine system native to India [6]. The health benefits of such medicine are highly questionable – and, in fact, the heavy metals used in such medicines may be dangerous [7]. However, these potential risks have not substantially impeded the usage of these remedies.

2.1.3 Aravind Eye Care System

The Aravind Eye Care System is a system of hospitals, clinics, vision centers, and camps in India which exclusively provide eye care [8]. It was founded in 1976 on the principle of performing large numbers of operations rapidly using an “assembly line” system to maximize surgeon efficiency [9]. Thanks to its extraordinary efficiency, about half of the 300,000 surgeries it performs each year are free [10].

2.1.4 Smile Foundation

The Smile Foundation is a non-governmental organization, or NGO, devoted to helping children in India. Its purview is broad, and much of it is irrelevant to our project. However, it also runs a program called “Smile on Wheels”. This program consists of “mobile hospitals:” large vans with medical equipment and trained medical staff which travel to locations in India where access to health services is limited or
Figure 1: Smile On Wheels Delivery Model [12]

nonexistent [11]. The general strategy is shown in Figure 1. This is not solely the purview of rural areas; the vans also provide needed services in slums.

The Smile on Wheels program relies on low-cost delivery mechanisms to provide healthcare to places that would normally lack it. However, Smile on Wheels is not a sustainable solution. It may visit a place and help people temporarily, but it cannot provide lasting services and long-term care. However, it can provide short-term palliative care and quickly identify and aid the worst problems in an area.

2.2 Center for Health Market Innovation

The Center for Health Market Innovation maintains a sizable database of health solutions and programs around the world [13]. The website provides details of each program, including its goals, structure, scale, technology used, and financial information. Though it has nearly 1,200 health solutions and programs, it does not provide analytical tools to analyze the effectiveness of the solutions. The lack of direct analysis means it is not in competition with our simulation, but is instead complementary to it. Even if they later implement an analysis component, they are still fundamentally solving a different problem: they are analyzing solutions which already exist in the real world based on data which exists, and our project attempts to analyze solutions before they are implemented to decide whether they are worth implementing.
2.3 Consulting Firms

The fundamental service that the simulation approach provides is the same service that consulting firms provide, so they are a potential source of competition, even if the method by which they do their analysis is completely different. One key firm which provides consultation services to health care solution providers is the Advisory Board Company [14]. The most relevant service they provide is the “Population Health Advisor”, where the company helps health solution providers refine their goals and strategies, advises them whether or not to proceed with their project, and evaluates the return on clients’ health solution programs.

To the extent that our simulation also evaluates the effectiveness of a proposed solution, we have some competition with consulting firms like the Advisory Board Company. However, our approach is nothing more than a raw analysis of data. It is more scalable, but less holistic – more quantitative, but without the wisdom of subjective judgement from experts. We serve different purposes, and in a hypothetical scenario where the simulation was perfected as a distinct entity, we would have different clients. A more likely scenario is that consulting firms would try to use this simulation to supplement their existing analysis.

3 Methodology

In order to evaluate the effectiveness of health care solutions, we decided to build a model which would simulate the population of India and allow arbitrary solutions to be tested on that population. There are four main parts involved in building and using this model: two which must only be done once, and two that must be done separately for each solution to be considered. The first step is the gathering of data necessary to understand the real population we are simulating. The second is the process of converting that data, which is generally aggregate data for a state or district within India, into a useful form in the model. These two need only be done once, when the initial population is constructed. The third step is to apply the effects of the given solution to the simulated population, generating a new population where the appropriate people have been treated using whatever treatment method is appropriate. Finally, the two populations can be analyzed and compared to determined the effectiveness of the solution.

This simulation itself must be tested as well to ensure that its analysis is accurate. The simulation can be validated by testing it on a solution for which we have good data. If the simulation can accurately model the effects of solutions that already exist, it can be trusted to be accurate for solutions with unknown effects.
We used the Aravind Eye Care System as a test for our model. As discussed above, Aravind operates a number of hospitals in India which exclusively treat eye health problems. It is valuable as a test because it is a highly successful and a well-studied solution, making it relatively easy to compare the actual effects of the solution with the effects projected by our simulation.

I wrote the simulation in Python, using SQLAlchemy [15] to interact with the data, which was stored in an SQLite [16] database. I ran a GitHub repository [17] to allow all members of the team to work on the code, and incorporated and adapted code they wrote using separate forks and pull requests to make sure that I could fix errors as they arose.

3.1 Gathering Population Data

In order to model the population of India, we needed statistics to understand what we were modeling. My teammates and I researched statistics for India about age, gender, income, expenditure, and life expectancy, drawing on a variety of sources [18] [19] [20] [21] [22]. I extracted the data from the original documents. The most frequently referenced data was consolidated them into an SQLite database.

In many cases, data that we needed could not be readily found, so our working estimates were more approximate. Some problems with these assumptions are described in Section

3.2 Using Population Data

The population data gathered consists of aggregate statistics that describe the country as a whole. However, it is difficult to understand the effects of solutions on the population as a whole – if we could accurately predict that, there would be no need to create this model. It is much more feasible to understand the effects of a proposed solution on an individual who is affected by the solution, such a specific person who visits a hospital and receives a particular surgery. To bridge this divide, I came up with the idea of building a randomized simulation of the population of India.

3.2.1 Building a Randomized Simulation

The core idea is simple: using the data provided, generate a number of individual “people” appropriate to the population of each district. Each person has several attributes, such as age and gender, which are assigned to the person randomly using distributions based on the data collected. I worked with Rie Kawauchi to build this randomized simulation of the population.
3.2.2 Person Attributes

Each person has several attributes, described below.

- **Gender**: Whether the person is male or female.
- **Age**: How old the person is.
- **Classification**: Whether the person lives in an urban or rural area.
- **District**: The district the person lives in.
- **State**: The state the person lives in.
- **Money**: The amount of money the person has available to spend.
- **Health Problems**: The health problems the person has.
- **Perceived Health**: How healthy the person thinks they are.
- **Health Utility**: How healthy a person actually is, on a scale from 0 to 1. This is described in Section 3.5.

This list of attributes is not an exhaustive list of every attribute which would be useful to track. For example, a person’s weight can have a wide variety of effects on their health, but it is not included. Instead, this is a limited list of attributes which were immediately relevant to the Aravind solution.

Each attribute has several effects on the rest of the simulation. For example, if a person thinks they are healthy, they will not seek out medical care, regardless of the number of health problems they actually have. The precise effects of each attribute depend on the solution, as described in Section 3.3, below.

3.2.3 Person Attribute Generation

Each attribute was generated in a different way based on the data provided, as described below. Some attributes depend on other attributes, so they are generated in order of dependency.

- **Gender**: 48.83% of India’s population is female [18]. We simply assigned gender randomly to each person.

- **Age**: Age was assigned randomly using a weighted distribution based on data for India [18]. The distribution is best seen graphically, as shown on 2.
**Location**: The district, state, and urban or rural classification of a person were not randomized. Instead, the generation of people automatically generates the appropriate number of people for each location based on exact census data [18].

**Money**: Money to spend is constructed by subtracting monthly expenditures from monthly income. We consulted the Statistics department to discuss the best model to use for random variable generation for income. We agreed that we use GDP by state (GSP) in India divided by the population of the state and set it as a mean of income per capita per state. Then we group the states into two categories; high income states which have more than Rs. 600 billion GSP, and low income state which have less than Rs. 600 billion GSP, and calculate the estimated standard deviation. As we expected, the distribution of GSP has log-normal distribution (Figure 3) and thus we randomized per-capita income generation following log-normal distribution. As rural and urban information are readily available, our goal is to narrow down the population generation per rural and urban area of a target state. GSP for rural and urban are not available, and thus we assumed that the monthly average expenditure weight for rural and urban can be applied to calculate the rural and urban income.

The simulation is done for poor state and rich state depending on the target state GSP level (Figure 4).

The National Sample Survey Office has a great database for monthly average expenditure per capital for each state and into rural and urban classification [20]. We downloaded the data and determined that the monthly expenditure has a distribution with strong positive skew as shown in Figure 5. We used log-normal distribution to replicate the monthly expenditures and confirmed that the mean and standard deviation matches the true value (Figure 6).
Figure 3: GSP Distribution [19]

Figure 4: Simulated GSP

The simulation is done for Andhra Pradesh state MPCE (Figure 6).

Health Problems: The health problems listed in this attribute depend on on the problems that the solution being tested can address. In lieu of precise data about eye health in India, we drew on a report which stated that 10% of the population have cataracts, and 30% have poor eyesight and would need glasses [23].

We then adjusted the distribution of these problems to ensure that older people are significantly more likely to have eye problems. This is necessary because age has a very strong effect on the QALY
calculations we use to evaluate the accuracy of the simulation. If the average age of treated patients in the simulation is significantly younger than the true average age of treated patients, the simulation’s QALY will be unrealistically high. However, we lacked the data to build a sophisticated distribution of eye problems with respect to age. To solve this problem, we set a threshold below which people could not have cataract problems, and we weighted the probability of eye problems so that younger people have a decreased probability of problems and older people have an increased probability.
**Perceived Health:** The perceived health ranges from 0 to 1, and is distributed based on three stages of health defined by the Ministry of Women & Child Development in India [24].

**Health Utility:** We assigned each person an overall health utility value from a random distribution shown on Figure 7. The mean is approximately 0.8518, which we drew from research on the health utility associated with eye care procedures [23]. The starting value of the health utility is not particularly important, however; what matters for the purpose of the simulation is how much the health utility improves.

After the health utility for a person is generated, each health problem that the person has reduces their health utility. Cataracts reduce utility by 0.14 [23], and we estimate that generally poor eyesight reduces health utility by 0.7.

![Truncated Density of Eye Utility](image)

**Figure 7: Truncated Health Utility for Eye**

11
Perceived Health

Perceived health ranges randomly from 0 to 1.

3.3 Implementing the Solution

Once the population has been constructed, the next step is to make the solution treat the population appropriately. This consists of two separate substeps. The first is to filter the population and decide which subsection must be treated. The second is to decide what it means for a person to be “treated” by the solution.

With Aravind in particular, these steps are somewhat complicated by the fact that the Aravind system is composed of four separate types of facilities: hospitals, clinics, vision centers, and camps. Different facilities can treat different problems.

3.3.1 Population Filtering

Not everyone should be treated by a solution. The first form of filtering is location-based: only people who are close enough to a facility should go to it. Doing this accurately would require assigning locations to every person in the simulation and matching them to a nearby facility. The current simulation takes a simpler approach. It simply checks whether a person is in a district that has a facility or is near a district with a facility.

Most solutions would require a filtering step based on money; only people with enough money to pay for treatment should be evaluated by the solution. However, Aravind’s structure of providing primarily free or cost-reduced treatment means that this step is unnecessary; anyone too poor to receive treatment normally would get it for free from Aravind.

Finally, the model assumes that only people who want medical attention should be treated. Some people refuse care for personal reasons. This is modeled using the perceived health attribute; if a person’s perceived health is low enough, they will seek out treatment.

3.3.2 Treating the Population

Each person who has made it through the filtering process is treated by the facility they were assigned to. When a patient is treated, their problem is removed and their health utility increases by the same amount that the problem initially reduced their health utility. However, a patient can refuse treatment, which does not charge them money but prevents their health from being improved.
Aravind specifically is considered to treat two problems: cataracts and generally poor eyesight. Only hospitals can treat cataracts, but any facility can prescribe glasses for poor eyesight.

### 3.4 Analyzing the Results

After the population has been treated, that new population can be compared to the original population. When testing the validity of the simulation, this can also be compared to an equivalent calculation using real-world values. These comparisons are done using a QALY calculation, as described in Section 3.5.

### 3.5 QALY: Measuring Effectiveness

The core metric used to analyze the effect of the solution is “Quality Adjusted Life Years” (QALY), a metric recommended to us by Dr. Clair Brown, a Berkeley professor [25]. We compare a QALY calculated based on Aravind’s actual results to a QALY calculated based on the simulation results and compare them; if they match, the simulation is providing accurate results.

Calculating QALY for a person is theoretically simple. Each person can be considered to have a “health utility”: a number ranging from 0 to 1 that represents the person’s overall health, with 0 representing death and 1 representing perfect health [26–28]. This value is determined by asking people about their preferences or desires for different health states [29] and about everyday life activities [30]. If a treatment improves someone’s health, their health utility increases. The amount by which the utility increases is multiplied by the estimated number of years the person has left to live. The resulting number is the QALY for that person.

Life expectancy is estimated given age using data from the World Health Organization [21]. The data is given in five-year intervals, so I interpolated the data over the intervening years to allow analysis at any age.

The simulation has the data required to compute the QALY for each person affected by the health solution. In order to get a broader measurement of the effectiveness of the treatment, the total increase in QALY can be summed across all people. This yields an evaluation of the entire solution. My teammates contributed the part of the code that actually performs the QALY calculation.
4 Discussion

I have written all aspects described above into the simulation, and it is currently functioning for the Aravind Eye Care System. Each district can be tested individually; if the district is within the purview of an Aravind facility, some people within the district will be treated. If the district is too far away from any Aravind facilities, the population is left untreated.

4.1 Simulation Accuracy

The simulation can currently generate reasonably accurate results under specific conditions. When the simulation is run on Theni, a district in the state of Tamil Nadu, it projects an QALY improvement among people with cataracts of 1.5, averaged over multiple simulations. This is not significantly different from the true average QALY among people with cataracts.

Though we could not find an existing explicitly calculated QALY for Aravind, this can be naively calculated by multiplying the average improvement in health utility by average expected years of life. Based on a health utility improvement of 0.14 for cataract surgery [23], an average age of 69.3, and a projected life expectancy among 69-year olds of 80.8 years [21], the correct average QALY is equal to $0.14 \times (80.8 - 69.3) = 1.61$.

4.2 Simulation Problems

Although the simulation is accurate by this measurement, the usefulness of this accuracy is limited. There are too many assumptions used which bear only loose resemblance to the complexity of reality. These are artifacts of the limited time available for the project. Below is non-exhaustive list of ways in which the accuracy and usefulness of the simulation could be improved.

- In the simulation, perceived health is independent from actual health. In reality, they should be related; though there exist hypochondriacs who will always look for medical attention and people who categorically refuse to seek out treatment, the vast majority of people will adjust their perceived health based on the problems they actually suffer.

- Each facility should be assigned a realistic capacity to limit the number of people it can treat.
• People should be assigned precise locations instead of the current combination of district, state, and urban/rural classification. This would allow each person to visit the specific solution facility that suits them, and is necessary to accurately model solutions which do not rely on fixed locations like Aravind does.

• Age and gender distributions should be assigned based on district-level data rather than averages for the entirety of India.

• The utility values associated with poor eyesight and other problems should be based on actual research into health utility values, rather than guesses inspired by health utility values we do have data for.

• QALY should be calculated as a sum across the entire population, allowing the breadth of the solution and the health improvement of the solution to be measured, rather than being averaged only across affected people.

• Support for additional solutions other than the Aravind Eye Care System should be built into the simulation. This can be done in a very similar fashion to Aravind, but because each solution is unique, each must be implemented manually. Features shared between solutions can be written generically so that they can be freely mixed and matched between different solutions.

• A user interface should be designed which would make testing the simulating easier than it is with the current command-line interface.

• A more robust measurement for the “true” QALY should be used.

5 Conclusion

There is a large gap between the goals of the simulation and the degree to which those goals are currently met. In its present state, the simulation is not a useful tool for analyzing health care solutions in India. There are too many assumptions made with an ambiguous basis in reality. This is a consequence of the limited time available for the project revision, its ambitious scope, and the difficulty that our group had in converting theory into working code.
However, the core idea of empirically “testing” solutions within a single simulation, ensuring consistent results regardless of the solution used, remains a powerful one. We have laid the groundwork for another group or organization to take this core concept and turn it into a project which can provide real value.

References


