Changing Healthcare

How can we Harness Predictive Analytics for Patients, Providers and Payers?

Ian Duncan FSA FIA FCIA FCA CSPA MAAA
University of California, Santa Barbara

12th November 2018
Introduction
Ian Duncan, FSA, FIA, FCIA, FCA, CSPA, MAAA

• Fellow, Institute of Actuaries (London) 1982.
• Health actuary since 1982; consultant since 1989.
• Founder and former president of PM company, Solucia Consulting (now SCIO Health Analytics) 1998.
• Founder and president, Santa Barbara Actuaries Inc.
• Professor of Actuarial Statistics, University of California Santa Barbara;
• Adjunct Researcher, William Sansum Diabetes Research Institute, Santa Barbara
• Author of several books and peer-reviewed studies on healthcare management and predictive modeling.
• SOA Board of Directors 2012-5; Massachusetts Health Insurance Exchange board 2007-14.
Agenda

1. Managed Care
2. Technology-driven Change
3. Actuaries and payment reform
4. Example: Accountable Care Organizations
5. The last mile problem
6. “Big Data”
7. Bloopers
8. Q&A
Technology-driven Change
Think back over the amazing advances in our lifetimes in technology, medicine, communications – in most industries in fact.
Technology-driven changes in our lifetimes

THEN

NOW

A worker from the 1940s would be bewildered by today’s technology.
Data has successfully transformed businesses
Data has successfully transformed businesses

What does disruption and transformation of these industries have in common?

**ANSWER:** Through data and technology, activity is transferred to consumers, reducing cost and increasing efficiency and consumer satisfaction.
What has not changed as much?

THEN

NOW

The 1940s worker would feel at home with today’s medical delivery.
Many reasons why we haven’t transformed healthcare

• The best models aren’t very good yet.
• Medicine is complicated; so are medical data; data is generally big but incomplete.
• Professional judgement over-rides analytics.
• It takes a long time for an intervention to produce returns.
• The “Last Mile” Problem.
• The self-management/persistency problem.
• While we have many outcomes studies, its hard to know what works.
• It is very difficult to implement and manage an outcome-effective, cost-effective program of sufficient scale to impact cost.
• Innovation is life-extending rather than cost-reducing.
What hasn’t changed in medicine?

• While medicine has become more technologically-intensive, it is also a very human resource-intensive business.
  • You still have face-to-face (one-on-one) visits to the doctor.
  • Aside from their technology, hospitals are still organized as they were by Florence Nightingale.
• Other industries have used technology to increase consumer power and choice, and to decrease cost. Why not in medicine?
What hasn’t changed in medicine?

• In part (I submit) because of the way medicine is financed. Either:
  • By large insurance companies, in insurance systems (such as the U.S.) or
  • By national health systems.

• In both cases, decisions are made *for* the consumer, not *by* the consumer.
Current Predictive Models aren’t very good

• “Most current readmission risk prediction models perform poorly...Efforts to improve their performance are needed.”

...but they are better than humans

- This study assessed the predictions made by:
  - Physicians
  - Case managers
  - Nurses

- “...none of the AUC values were statistically different from chance”

Predictive Modeling in Business helps consumers make better choices, consistent with their long-term goals (cheaper, better products; tailored choices; faster access). The Progressive (and other auto insurers) experience is instructive:

- Progressive (through the use of predictive modeling) identified customers who were classified as high-risk (and assigned to a high-risk pool) but who were predicted to be low-cost.
- Their marketing focused intensely on these customers.
- The company successfully penetrated a profitable market niche.

BUT: They also focused on changing risky behavior in their existing customers.
Data has led to an explosion of devices....

It is significant that you can change your behavior in auto insurance (and benefit from a “fair price”) but not in health insurance.
Data has led to an explosion of devices....

Since 2015, the National Football League has placed chips under shoulder pads to track exactly where players move. Fans can’t get this data, but teams can. Mr. Hall tells me it can provide “instant analysis to coaches.”
And too much reliance on technology...

Fastest-growing app category: over 84,000 new health apps. published in 2017¹. However, many examples of Failed Apps²:

- Dermatology app. diagnosing cancerous moles – withdrawn as too inaccurate;
- In 2011, pharma giant Pfizer recalled a rheumatology calculator app after the company found that its swollen-joint measurements—calculated using self-reported data—were off by as much as half.

Most health apps are classified as “informational” or “entertainment” to escape FDA oversight.

1. Survey by consulting firm Research2guidance; https://research2guidance.com/
Too much reliance on adherence and self-care...

A Failed mHealth Program Offers Lessons Learned For Future Projects

An project to have FQHC patients use an mHealth app to manage their diabetes and hypertension at home collapsed after a few weeks. But researchers say they learned valuable lessons.

- 1/3 of target patients did not download app.
- 12% of patients were regular users.
- Most patients stopped using the app after a few weeks.
- Staff used office visits to train patients in use of smartphones (!).
Change Agents in Medicine

Payment Reform

Data Analytics

Behavioral Economics
Payment reform models in many countries aim to end unlimited healthcare budget growth by transferring responsibility for financial and clinical outcomes to providers and other risk-takers. Models include:

- Pay for performance
- Gainsharing
- Accountable Care Organizations
- Bundled Payments
- Patient-centered Medical Homes
- Risk-taking carve-outs
Health Actuaries and Predictive Analytics
Actuaries have practiced PM forever....

Actuaries have been using predictive methods since the profession began.

Identify patterns/segment risks  
Develop business rules  
Improve decision making

1779-1865

\[ \mu(x) = \alpha e^{\beta x} \]

1826-1891

\[ \mu_x = A + Be^{x} \]
Actuaries have practiced PM forever...

Actuaries have been using predictive methods since the profession began. Examples: Mortality Rates (survival functions = parametric predictive models).

Identify patterns/segment risks

Develop business rules

Improve decision making

Makeham

Gompertz
Example: Alternative Payment Models, Predictive Analytics and Behavioral Economics
Accountable Care Organizations

The **Medicare Shared Savings Program** was established by the Affordable Care Act. Congress created this program to better coordinate among providers to ensure quality care for Medicare Fee-For Service beneficiaries and to reduce unnecessary costs.

ACOs share 50% of any savings (projected cost minus actual cost) with Medicare. Risk adjustment is applied to the comparison population to ensure risk comparability with managed population.

1. In the UK: Accountable Care Groups
Spending at End-of-Life

5% of Medicare Beneficiaries die annually.

12% of Beneficiaries Driving 69% of the Expense

Second to last year of life represents 13% of the total Medicare FFS spend.

Last year of life represents ~25% of the total Medicare FFS spend.

Other 26%
Over-medicalized death defined as:

- Chemotherapy for cancer patients within 14 days of death
- Unplanned hospitalization within 30 days of death
- More than one emergency department (ED) visit within 30 days of death
- ICU admission within 30 days of death; or
- Life-sustaining treatment within 30 days of death
Figure 5.1.b  Cost by Place of Death and Type for Patients in Last 6 months of life
Average Hospital Cost per Day – last 6 months of life

Average Cost per day for patients who die in hospital
Medicare Hospice Payment Rates by Type of Service

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Base payment rate, 2015</th>
<th>Percent of hospice days, 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routine home care</td>
<td>Home care provided on a typical day</td>
<td>$159.34 per day</td>
<td>97.6%</td>
</tr>
<tr>
<td>Continuous home care</td>
<td>Home care provided during periods of patient crisis</td>
<td>$38.75 per hour</td>
<td>0.4</td>
</tr>
<tr>
<td>Inpatient respite care</td>
<td>Inpatient care for a short period to provide respite for primary caregiver</td>
<td>$164.81 per day</td>
<td>0.3</td>
</tr>
<tr>
<td>General inpatient care</td>
<td>Inpatient care to treat symptoms that cannot be managed in another setting</td>
<td>$708.77 per day</td>
<td>1.7</td>
</tr>
</tbody>
</table>

2018 Rates: $190.55 (Routine Home Care)

By comparison, the 2013 Average Hospital Facility per diem was $1,617.62
The Cost PMPM for members in each category varies across risk scores. The difference in costs between those that experience overmedicalized deaths versus those that experience appropriate deaths is greatest in members with risk scores > .95.
Example: Accountable Care Organizations

• The model is accurate in predicting those patients at risk of over-medicalized death.
• The economic model indicates that there are potential gains to the ACO from intervening on this population.
• The quality of care delivered in alternative settings results in less pain or anxiety and fewer side-effects to patients.

What could possibly go wrong with this win-win proposition?

*Hint*: Incentives and Behavioral Economics.

The model overlooks the *last mile problem*. 
The “Last Mile” Problem

Data/Analytics can identify issues and find opportunities. On their own, they cannot solve the “Last Mile” problem.

Overcoming the “Last Mile” problem explains why Progressive Insurance, Netflix, Amazon, the Oakland As and others have successfully implemented data analytics and healthcare has not. Consumers either enjoy lower rates (Progressive), better and cheaper players (A’s), a better consumer experience (Amazon), etc.


©2018 Santa Barbara Actuaries Inc. Confidential and Proprietary. All rights reserved.
Example: Big Data
“Big Data”

Plus:
- Unstructured
- Novelty
- Messy
- Machine learning

• External
• New Tools
• Patterns
• Algorithms

- User-created
- Enabling technology
- Etc.

Source:
Big data has come to refer to the use of predictive analytics, user behavior analytics, or other advanced data analytics methods that extract value from data, and seldom to a particular size of data set. "There is little doubt that the quantities of data now available are indeed large, but that’s not the most relevant characteristic of this new data ecosystem."\(^1\) Analysis of data sets can find new correlations to "spot business trends, prevent diseases, combat crime and so on."\(^2\)

Machine Learning: a field of computer science that gives computers the ability to learn without being explicitly programmed.

The issue of missing data: Wearable devices

Example of a “Big Data” Dataset
Detailed source data were not available (privacy). Summarized data categorized as:

<table>
<thead>
<tr>
<th>Classification of Physical Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device-recorded verified workouts</td>
</tr>
<tr>
<td>Light Workouts</td>
</tr>
<tr>
<td>Standard Workouts</td>
</tr>
<tr>
<td>Steps</td>
</tr>
<tr>
<td>Calories</td>
</tr>
<tr>
<td>Time at 60% Maximum Heart Rate</td>
</tr>
</tbody>
</table>

Approximately 300,000 participants over 4 years; continuously reported data, including clinical data available on 8,519 participants between January 1, 2013 and August 31, 2015.
### Number of workouts per week

<table>
<thead>
<tr>
<th>Year</th>
<th>Light Workouts</th>
<th>Standard Workouts</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Mean (SD)</td>
<td>Max</td>
</tr>
<tr>
<td>2013</td>
<td>0</td>
<td>0.86 (1.05)</td>
<td>6.1</td>
</tr>
<tr>
<td>2014</td>
<td>0</td>
<td>1.27 (1.29)</td>
<td>6.8</td>
</tr>
<tr>
<td>2015</td>
<td>0</td>
<td>1.09 (1.04)</td>
<td>7.0</td>
</tr>
</tbody>
</table>

Other available measures include BMI, Age, Sex, Smoking status, depression status, no. alcoholic drinks per week, blood pressure and serum cholesterol level.
## Predicted 20-month BMI Measure for Two Sample Participants

<table>
<thead>
<tr>
<th>Sample Participant</th>
<th>Baseline BMI Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>45</td>
</tr>
<tr>
<td>30-year old Female; 3 std./1 light w/out weekly</td>
<td>17.94 (5.5)</td>
</tr>
<tr>
<td></td>
<td>25.46 (1.9)</td>
</tr>
<tr>
<td></td>
<td>34.87 (-0.4)</td>
</tr>
<tr>
<td></td>
<td>44.30 (-1.6)</td>
</tr>
<tr>
<td>60-year old Male; 5 std./1 light w/out weekly</td>
<td>17.70 (4.1)</td>
</tr>
<tr>
<td></td>
<td>24.99 (0.0)</td>
</tr>
<tr>
<td></td>
<td>34.11 (-2.5)</td>
</tr>
<tr>
<td></td>
<td>43.23 (-3.9)</td>
</tr>
</tbody>
</table>
Model 1: BMI

Effect of Exercise Levels on BMI for selected participants

Female - Age 30 - 1 Light Workout

Male - Age 60 - 1 Light Workout
Conclusions

• Physical activity even at low levels can have positive impacts on measurable health metrics.

• Physical activity levels (light and standard) had the largest impacts on BMI and HDL cholesterol levels, but little to no effect on either blood pressure or LDL cholesterol levels.

• A measureable impact on health outcomes requires frequent, intense exercise.
“Big Data” - What is the member’s probability of having Diabetes?

Step 1
- Logistic regression model (all variables, based on whole data set)
- Random forests model (all variables, based on whole data set)

Step 2 Sub-model
- Logistic regression model (important variables, based on whole data set)

Step 3 Balanced Data
- Logistic regression model (important variables, based on balanced data set)
- Random forest model (all variables, based on balanced data set)
Variable Importance

Random forest

Logistic regression model

<table>
<thead>
<tr>
<th>Important variable</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Waist_circumference</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Average_of_other_program_activities</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Education</td>
<td>0.00012</td>
</tr>
<tr>
<td>Education6</td>
<td>8.83E-13</td>
</tr>
<tr>
<td>Average_alcohol_drinks_per_week</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>BMI</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Kessler_stress_score</td>
<td>3.19E-15</td>
</tr>
<tr>
<td>Tobacco_use</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>HDL</td>
<td>8.18E-14</td>
</tr>
<tr>
<td>DBP</td>
<td>1.23E-15</td>
</tr>
<tr>
<td>LDL</td>
<td>2.34E-07</td>
</tr>
<tr>
<td>TC</td>
<td>4.85E-05</td>
</tr>
<tr>
<td>Heart_disease</td>
<td>4.03E-16</td>
</tr>
<tr>
<td>Heart_disease</td>
<td>4.35E-14</td>
</tr>
<tr>
<td>chronic_lung_disease</td>
<td>6.85E-06</td>
</tr>
</tbody>
</table>
## Model Evaluation

<table>
<thead>
<tr>
<th>Model</th>
<th>Dataset Used</th>
<th>Predictors</th>
<th>Threshold</th>
<th>AUC</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>CV Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>Whole dataset</td>
<td>All variables</td>
<td>0.08</td>
<td>0.7903</td>
<td>0.7127</td>
<td><strong>0.7380</strong></td>
<td>0.7359</td>
<td></td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>Balanced dataset</td>
<td>All variables</td>
<td>0.50</td>
<td><strong>0.8000</strong></td>
<td>0.8486</td>
<td>0.5753</td>
<td><strong>0.7465</strong></td>
<td>0.7279</td>
</tr>
<tr>
<td>Random Forests</td>
<td>Whole dataset</td>
<td>All variables</td>
<td>0.08</td>
<td>0.7982</td>
<td>0.7514</td>
<td>0.7041</td>
<td>0.7078</td>
<td></td>
</tr>
<tr>
<td>Random Forests</td>
<td>Balanced dataset</td>
<td>All variables</td>
<td>0.50</td>
<td>0.7973</td>
<td>0.7678</td>
<td>0.6908</td>
<td>0.7445</td>
<td></td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>Whole dataset</td>
<td>Important variables</td>
<td>0.08</td>
<td>0.7897</td>
<td>0.7094</td>
<td><strong>0.7382</strong></td>
<td>0.7358</td>
<td></td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>Balanced dataset</td>
<td>Important variables</td>
<td>0.50</td>
<td><strong>0.7999</strong></td>
<td><strong>0.8486</strong></td>
<td>0.5757</td>
<td><strong>0.7466</strong></td>
<td>0.7262</td>
</tr>
</tbody>
</table>

**Validation of Predicted Values**

1. Derivations from confusion matrix
2. AUC

Sensitivity: True positive rate
Specificity: True negative rate
## Logistic Regression Odds-ratios

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Variable Description</th>
<th>Analysis: Logistic Regression on Balanced Dataset (Undersampling)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>In years; range is 18 - 80</td>
<td>Odds Ratio (95% Conf. Int.) p-value</td>
</tr>
<tr>
<td>Education (11)</td>
<td>College degree</td>
<td>1.0502 (1.0475, 1.0529) &lt; 0.001</td>
</tr>
<tr>
<td>Education (6)</td>
<td>Postgraduate degrees</td>
<td>0.9402 (0.8806, 1.0038) 0.0648</td>
</tr>
<tr>
<td>Daily Sedentary Hours</td>
<td>In hours</td>
<td>0.8214 (0.7515, 0.8978) 0.0648</td>
</tr>
<tr>
<td>Kessler Stress Score</td>
<td>Range is 1 - 50</td>
<td>1.0137 (1.0065, 1.0210) 0.0002</td>
</tr>
<tr>
<td>Average Alcohol Drinks per Week</td>
<td>Average drinks</td>
<td>1.0140 (1.0073, 1.0207) 0.0001</td>
</tr>
<tr>
<td>Chronic Lung Disease (2)</td>
<td>Yes versus No; does not have disease</td>
<td>0.9661 (0.9586, 0.9735) 0.0001</td>
</tr>
<tr>
<td>Heart Disease (2)</td>
<td>Yes versus No; does not have disease</td>
<td>0.5944 (0.3592, 0.9549) 0.0365</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>In kilograms per meter squared</td>
<td>0.6696 (0.5467, 0.8161) 0.0001</td>
</tr>
<tr>
<td>Waist Circumference</td>
<td>In centimeters</td>
<td>1.0343 (1.0268, 1.0417) 0.0001</td>
</tr>
<tr>
<td>Systolic Blood Pressure</td>
<td>In mmHg</td>
<td>1.0169 (1.0142, 1.0195) 0.0001</td>
</tr>
<tr>
<td>Diastolic Blood Pressure</td>
<td>In mmHg</td>
<td>0.9908 (0.9870, 0.9946) 0.0001</td>
</tr>
<tr>
<td>Total Cholesterol</td>
<td>In mmol/L</td>
<td>1.0217 (1.0186, 1.0249) 0.0001</td>
</tr>
<tr>
<td>High Density Lipoprotein</td>
<td>In mmol/L</td>
<td>0.7860 (0.6946, 0.8890) 0.0001</td>
</tr>
<tr>
<td>Low Density Lipoprotein</td>
<td>In mmol/L</td>
<td>0.8379 (0.7286, 0.9640) 0.0132</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>In mmol/L</td>
<td>0.9243 (0.8164, 1.0468) 0.214</td>
</tr>
<tr>
<td>Count of Other Program Activities</td>
<td>Count per week, self-reported</td>
<td>1.5019 (1.4119, 1.5988) 0.214</td>
</tr>
</tbody>
</table>

©2018 Santa Barbara Actuaries Inc. Confidential and Proprietary. All rights reserved.
Cluster Analysis

- Want to investigate the data to see any interesting subpopulations.
- Clustering data of mixed type:
  - Gower distance
  - Partition around medoids (PAM)
  - Silhouette width
- 10,683 observations used (5% of data)
- t-Distributed Stochastic Neighbor Embedding (t-SNE)
Machine Learning Application: 10-Cluster Plot

1. Male with post graduate degree
2. Male college graduate; AA alcohol consumption
3. Female college dropout; high stress levels
4. Female college graduate; smoking history
5. Mostly female high school graduate; BA alcohol consumption
6. Male college dropout; AA BMI, AA waist circumference
7. Female college graduate; no smoking;
8. Female with post graduate degree; BA BMI, BA waist circumference
9. Female college dropout; high depression rate with a smoking history
10. Male college graduate; no smoking, AA waist circumference

* AA – above average; BA- below average
Whoops! You need to understand the data

Google Flu example

Why Google Flu Is A Failure

It seemed like such a good idea at the time.

People with the flu (the influenza virus, that is) will probably go online to find out how to treat it, or to search for other information about the flu. So Google reasoned, decided to track such behavior, hoping it might be able to predict flu outbreaks even faster than traditional health authorities such as the Centers for Disease Control (CDC).

Indeed, as the authors of a new article in Science explain, we got “big data flu.” David Lieven and colleagues explain that:

"Big data flu" is the often implicit assumption that big data are a substitute for, rather than a supplement to, traditional data collection and analysis.

The folks at Google figured that, with all their massive data, they could outsmart anyone.

Symptoms of Influenza

Central

Headache

Nasopharynx

FEVER PEAKS
A comparison of three different methods of measuring the proportion of the US population with an influenza-like illness.
“My daughter got this in the mail!” he said. “She’s still in high school, and you’re sending her coupons for baby clothes and cribs? Are you trying to encourage her to get pregnant?”

The manager didn’t have any idea what the man was talking about. He looked at the mailer. Sure enough, it was addressed to the man’s daughter and contained advertisements for maternity clothing, nursery furniture and pictures of smiling infants. The manager apologized and then called a few days later to apologize again.
Whoops! Prediction-based interventions are great, but...

Original Investigation

November 12, 2017

Association of the Hospital Readmissions Reduction Program Implementation With Readmission and Mortality Outcomes in Heart Failure

Ankur Gupta, MD, PhD; Larry A. Allen, MD, MHS; Deepak L. Bhatt, MD, MPH; et al

Author Affiliations

JAMA Cardiol. Published online November 12, 2017. doi:10.1001/jamacardio.2017.4265
RIP Google Health

by John Moore | June 24, 2011

Chilmark Research has not had a very good feeling about Google Health for well over a year now. Back in early May of this year we felt that Google had all but given up and had put Google Health in stasis. Today, Google made it official, Google Health has a little more than six months to live, then it will get the Kevorkian treatment with Larry Page administering the final lethal dose.
Cambridge Analytica was allowed to pull that profile data. Facebook only changed its policy in early 2015. But before the general election, the Trump campaign dropped Cambridge Analytica for the Republican National Committee data, reportedly never using any of the “psychographic” information. According to CBS News, in September 2016, it had “tested the RNC data, and it proved to be vastly more accurate.”
Health Actuaries and Big Data
Actuaries are well-positioned to *lead* in this new world:

1. Actuaries combine the three components necessary for successful change:
   - Data analytics;
   - Payment reform; and
   - Behavioral Economics.

2. Actuaries comply with ethical and professional standards.

3. Actuaries understand the business and risk context in which medicine operates.

To do so, however, will require health actuaries to upgrade their statistical and clinical knowledge. Fortunately, many actuarial organizations are offering educational opportunities in these areas.
Thank you