The Risk Adjustment Digest 2018

The Technology Issue

Contents

Executive Note

2018 promises to be a milestone year for our industry. We continue to see steady adoption of value-based contracts, across different sectors of American society. Medicare Advantage enrollment grew by 8% to 19 million members in 2017. Managed Medicaid also grew, by 2% up to 42 million members in 2017. ACA enrollment remained strong despite facing significant headwinds in the past year.

As populations under risk adjusted reimbursement contracts grow, the costs related to data acquisition, analysis, and chart review to properly capture individual disease burden grows. The demand for tools to reduce the time and cost of risk adjustment activity while increasing accuracy has never been greater.

Five years ago, Apixio introduced the first technology-based solution to assist coders with HCC chart review. At that time, we encountered a lot of industry skeptics – for good reason. Reviewing hospital and clinic charts to determine if there is sufficient documentation for one of thousands of HCC relevant diseases, based upon a plan coding guideline, is a very complicated task. There are no clear rules to follow. There are many different shades of gray when interpreting the messy world of physician documentation. The industry skepticism only deepened when we described that we didn't program rules for the computer to follow when reading the chart. Instead we developed a way for computers to learn over time how to best read and decipher key chart elements related to a risk adjusting disease.

With more experience and positive results demonstrated, the industry appears to be more comfortable with machine assisted chart review. And, based upon this comfort, more health plans are making the decision to insource their risk adjustment chart review. New technology vendors coming into the market, and older vendors are trying to integrate these technologies into their offerings. The marketing buzzwords are now all over vendor websites, tradeshow booths, and product collateral.

In less than five years, we have traveled from a landscape devoid of technology supporting risk adjustment coding to an industry in which many are proclaiming that technologies such as natural language processing and machine learning for HCC coding are commodities. We are astonished.

We see that amid the excitement and activities to adopt new technology for risk adjustment, there is more confusion than ever.

This Risk Adjustment Digest is meant to offer insights and clarity so that you can better navigate this brave new world. There are many foundational questions which need to be answered. What are the pros and cons of different technological approaches? How should accuracy be measured? How is the process going to (or should) change? How will technology impact risk adjustment operations? What is the right approach to use this technology for validation and audit activities? Will machines ever replace coders? Can this technology extend into prospective activities?

The future of risk adjustment will continue to evolve. Our team at Apixio is working every day to bring new, innovative ways to solve the challenges of risk adjustment.

Darren Schulte, MD

CEO, Apixio

Five Ways You're Already Using Machine Learning:

A Day with AI

"Machine learning" can seem like a scary term, bringing to mind images of the techno-dystopias portrayed in the Matrix, Terminator, and Black Mirror. However, far from these dark narratives, machine learning has been reaching ordinary people for a long time in simple, common, and helpful ways. Here are five such ways you're (probably) already using machine learning.

1. Checking the situation on Google Maps:

Most of us, especially those that drive, have used Google Maps before. However, not many ask: How's this nifty thing done? How does Google construct maps that reflect every street, avenue, and alley with such accuracy? The answer is artificial intelligence, or AI, and machine learning to be specific.

Google Maps is fueled by accurate real-time information from thousands of images of street signs and locations, which are analyzed by Optical Character Recognition algorithms. The task is challenging because the algorithms must know to ignore extraneous information — like the lettered shirt of the person standing next to a street sign for instance — and to pick out just the relevant information. To accomplish this tough task, Google trained its machine learning algorithms using the difficult French Street Name Signs (FSNS) database. Once the algorithm could correctly recognize these complex images, it grew much more accurate.

2. Ordering a Lyft or Uber:

Whether you're team Lyft or prefer to hop in an Uber, both services, much like Google Maps, power decisions with AI. Driver assignments, driver ETA, and your ETA at your final destination – are all calculated by algorithms

that are constantly tested and refined in real time, using machine learning and the massive quantities of data from drivers and customers.

One important thing for many; rideshare companies are using machine learning to help beat the dreaded 'surge price'. Surge pricing, or time-limited price hikes, currently compensates for times when there are not enough cars on the road to supply all the passengers who want rides. Ideally, machine learning could anticipate times of high demand (say, commute times in April on the East Coast when it frequently rains) and incent cars to be on the road, in advance.

3. Using spam filters and priority tags to keep you organized:

Your inbox probably seems an unlikely place for machine learning, but AI technologies are in fact the engine behind the "Spam Filter", one of email's most important tools.

Simple rules-based filters are used for the spam filter. Think of the words and phrases "pharmacy", "you've won the lottery" or "Nigerian prince". While you may very well be friends with a Nigerian prince, if the message seems suspicious, and is coming from an unknown sender, then it will probably get flagged and kicked to spam. The filter 'learns' the content of the emails, identifying signals by gathering inferences from word relationships. The filter is able to countermaneuver spammers that try to outsmart it with updated messages.

In addition to this general work, the filter also 'machine learns' what you, the user personally consider spam. It does so by ingesting data on what you delete and mark as spam mail. This data works in conjunction with data contributed by the entire user base.

With these approaches working in tandem, some reports place Gmail's spam filters at a 99.9% success rate. Additionally, researchers tested the effectiveness of Priority Inbox on Google employees and found that those with Priority Inbox "spent 6% less time reading email overall, and 13% less time reading unimportant email."

A similar approach is used to tag emails for predetermined primary, social, and promotion inboxes, as well as automatically labeling emails as important.

4. Preventing credit fraud:

A database of consumer complaints by the Federal Trade Commission reports 1.3 million (42%) of the 3 million complaints in 2016 were fraud related, with losses of $744 million. A total of 55% of all cases were fraud or identity theft related – that's more than 1 in every 200 Americans per year. Keep in mind, 7% of American households are unbanked, and only 77% of Americans are over the age of 18.

Big problem, right? But how do financial institutions even determine if a transaction is fraudulent? Bank of America alone has an estimated 58 million customers and like most banks, its daily transaction volumes are much too high for manual review.

To sift through this mass of data and distinguish normal purchases from illegal ones, banks use machine learning. For example, FICO, well-known producer of credit ratings, uses neural networks to predict fraudulent transactions. Named for and meant to simulate the neural networks in our own brains, these systems analyze examples of labelled data, developing their own characteristic markers, and learning how to identify unlabelled inputs. Factors considered include: the customer's recent frequency of transactions, transaction size, and the kind of retailer involved. (It's good for the bank too— MIT researchers found in 2010 that machine learning applied to customer transactions, could reduce bank losses from delinquency between 6 and 25%).

5. Posting on Facebook

Facebook uses machine learning to ensure that you see the most relevant news from your friends. Many

of us have hundreds of Facebook friends, not all of whom are equally close to us. Sociologists have actually determined that it's implausible to have more than 150 close friends — it's called the Dunbar number. Machine learning helps Facebook determine which one of your connections is truly a close friend whose life you want to hear about, and who is your cousin's ex-boyfriend who showed up at Thanksgiving once.

Facebook goes through all potential posts you could see, and assigns each one of them a score. The score is determined by a couple things; your previous "like" and "comment" activity; whether the post is an interaction between two people or between a person and a corporate page; whether the post is from one of your friends or a friend of a friend; whether the post has seen a lot of time-consuming activity (many people commenting) or is merely "liked".

The scores of all your potential posts are ranked, and that ranking becomes your Newsfeed.

Just the tip of the iceberg

All these examples are just the tip of the iceberg, AI is being used in so many different fields – anything from fashion, where companies like Stitch Fix use machine learning, customer feedback and stylist expertise to deliver clothes to your doorstep by subscription, and collect over $730 million in annual sales; to sports coaching,where it's applied to thousands of historical strategies and plays to make predictions and recommendations based on this training data and user inputs.

Machine learning helps you throughout your day— all without you realizing or thinking about it.

INTRODUCTION TO Machine Learning

At Apixio, amongst other things, we use machine learning to predict medical conditions that patients most likely have based on clinical documentation and medical claims data. After ingesting, cleaning and normalizing this data, we train the models from which we make these predictions. In the data science world when we use the term "model" what we mean is that we are trying to create a function† **f** that when given input **x** produces one or more outputs that classify the input **x**. We use a number of computational techniques to create these functions, all of which involve three important components: data points, features and weights.

†Note: this is not your typical f(x) = y function, in which a function acts on x to produce y. Instead, this is a situation where x and y are both independent variables and we are trying to discover a function that ties them together in a predictable way. In mathematical terms, this is an Xi, Xj plot.

Basic Components of Training Data

For the following explanations of data points, features and weights, let's say we are creating a classifier that predicts whether a patient has a broken leg or not.

A **data point** comes from the domain and range of the function we are trying to create to do our predictions. In this example the patient medical records are the domain and the prediction values, the information about whether the patients have a broken leg, is the range. When training a model we will have a data point for each patient containing some information on the patient (a.k.a. **features**, see below) and a binary value for "broken leg".

The difference between having the binary true/false value vs. an empty value is the difference between what we call **labeled vs.**

unlabeled data. A "labeled" data point is one where we know the answer to whether the patient has the condition "broken leg" or not, and for an "unlabeled" data point, we do not know the status of our patient's legs. Unlabeled data can be used in machine learning setting, but labeled data are always better.

Features are pieces of information (traits) we extract from the raw data that are important in the prediction task. For example, the fact that "someone recently visited an orthopedist" might be an important piece of information for predicting whether the patient has a broken leg. It does not certainly follow that the patient has a broken leg, but it makes it more probable. We get features from the structured (tabular) data and from written clinical documents (typically words). Our models may end up having hundreds or thousands of features though good models can be created from just a handful of the right features too.

Each feature has a corresponding **weight** (think - importance) assigned to it which is unknown in the beginning of training. The weights and the features serve as inputs to whatever function we choose to model our outcome with. In many cases our function is of the form f(**W**x+b), where **W** is the weights matrix, **x** is the feature vector, and **f** is a non-linear transformation.

The process of finding **W** is called the learning process or "training"." For many of our predictors we can use **logistic regression**, where the logistic function will give us a probability that the patient has the condition. During training we use a **cost function** which is an error function that compares the predicted values with the labels, and outputs the amount of aggregated error over the whole dataset. The objective of "training" then is to find a **W** that minimizes the aggregate error. If the error is small, then those weights are considered to be "good" weights and we can use them to predict whether our patients have broken legs. There are other factors, like regularization terms, that go into a cost determining "goodness" of **W**. But minimizing the cost function for aggregate error is a good starting point.

Logistic Regression is a classification technique that uses a straight line to classifiy between one of two outcomes (binary outcomes).

Fitting a Model

So we've covered the basics of what goes into creating a predictive model . Now let's look at what our functions do to data to classify something. In **FIgure 1** you can see an example of a trained linear classifier on a simple Xi,Xj plane which separates the data into two categories of blue and green. Our function here represented by the straight line, does a perfect job of separating the green and blue classes. If the training set is truly representative then this would be a perfect classifier. In real life your data is not going to line up for you like this, but the example is to show what we are expecting our function to do, separate our data points in space into separate classes.

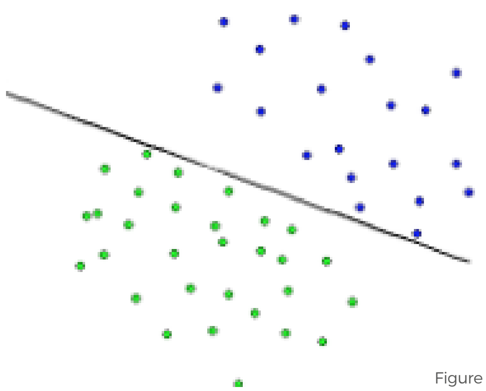


Figure 1

In **FIgure 2** we show a classifier that is a bit more realistic because the data doesn't line up neatly. The function here is not linear and clearly not perfect. If the Red class is the positive class you can see that there are Green false positives on the Red side of the line and some Red false negatives on the Green side. But by and large the function does a good job of separating the classes. Even in this simple case we see that we may not be able to separate the two classes and this represents the error in our classifier and is very common in machine learning.

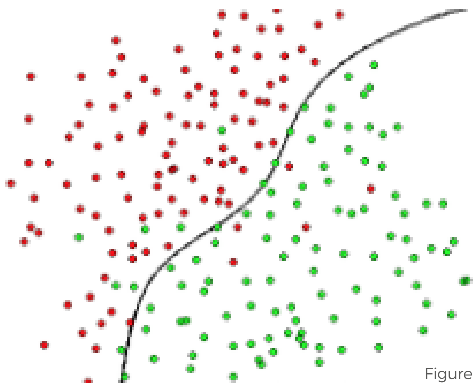


Figure 2

It is the process of training that creates a function that can separate our data into the classes we desire for our classifier.

There are different approaches and techniques to actually train a model using optimization algorithms like **gradient descent**, but we'll save these for another article. Let's assume that we have data points, features and an appropriate cost function and that we have trained and minimized the error; can we be sure the weights we get from training will result in good predictions when we apply them in prediction tasks on new data? In other words, if a new patient comes, can we extract the features, apply our logistic function with our best weights and rely on the results?

The answer is: maybe "yes" and maybe "no". We won't know unless we've held out some test data points, but if the answer is "no" then there is a culprit we always go to early on, "overfitting".

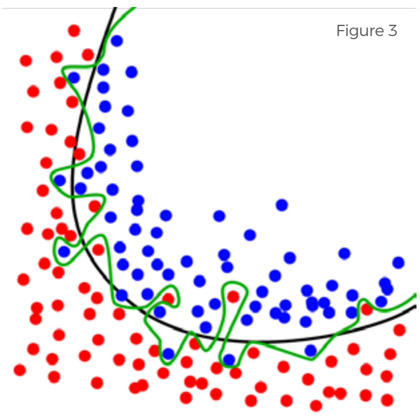


Figure 3

During the process of minimizing our cost function, we try to come up with the best weights to "fit" the training data. Two things can go wrong for us during this process. One is that we might have too few data points for the number of features, also referred to as **dimensions in the model**. This is always bad. The second common cause of overfitting is typically noise in our data set. In either case instead of achieving a smooth curve, the black line in the illustration, in **Figure 3** on the left, we end up with something that looks like the green line which tries to account for every datapoint. This kind of function that looks like a congressional districts after heavy gerrymandering is going to be much more susceptible to error when new unseen data points are run through our predictor. The issue is that our model has learned the training data too precisely, either there are too few points to balance out our function or the model has learned the noise in our training data. When we get new data points we've not seen, they are likely to not fit the model.

Solving for Overfitting

The question then is how do we fix a model that is performing poorly due to overfitting? If we're lucky, we can play with the "learning parameters". We might be able to stop or modify the learning process before it overfits or use cross validation methods (which might not be feasible if our problem is too few data points). However these often do not fix our problem. If that's the case we use one of two common approaches: **feature selection** and **regularization**. The objective of feature selection is to reduce the dimensionality of the model by removing features which are likely to be contributing to the overfitting. Removing dimensions then improves the power of the training data we do have. The idea behind regularization is to constrain the space of the functions we are learning to make them more "smooth" with the intuition being that a smooth function is less likely to overfit the data. In the previous diagram, the black curve is smooth and the green curve is not. We want "regular" functions and not "gerrymandering" (excuse the political analogy) and regularization as a way to achieve this.

So, a closing point for this article. We've just discussed the problems inherent in fitting training data to a model and the special challenge of solving for overfitting. There's a reason this is important to us at Apixio. In our business we deal with healthcare data and while one might assume that with the prevalence of health record systems these days that there's plenty of data to go around, but this simply is not the case. Data in the healthcare milieu is actually quite hard to come by, it is very noisy and there are literally thousands of conditions and procedures many of which are rare so there are not large datasets to use for training. Peter Norvig at Google coined the phrase, "the unreasonable effectiveness of data" the point being if you have a lot of it, your job gets a lot easier. In healthcare we currently do not have this luxury so our models often do not generalize well in production. Our job is to find ways to overcome these issues. Our next post will focus more closely on using regularization to solve overfitting, namely by addressing too-small training sets and noisy data.

FOUR WAYS

Artificial Intelligence Can Bend Healthcare's Cost Curve

Estimates suggest that the United States wastes more than $1 trillion each year on healthcare (one-third of all healthcare spending). This waste includes over-treatment, seeing multiple doctors unnecessarily, care delivery failure and lack of care coordination.

Unfortunately, healthcare cost-cutting measures have historically made matters worse in the United States, often cutting back access to care in a misguided race to the cost floor.

A perfect example of these cost-cutting measures is the creation of narrow network plans, which limit choice and network access in an attempt to reduce costs. According to McKinsey, 70 percent of plans on the Affordable Care Act exchanges in 2014 featured a limited network, and narrow-network plan premiums were on average 17 percent cheaper.

While less choice does typically mean lower premiums, often patients are deprived of seeing the physicians and services they need in order to receive appropriate care. There should be a better way.

How then can we cut costs while designing a more responsive, efficient and effective health care system?

Data analytics: The path forward to increasing high-value care

In the United States, more than 1.2 billion clinical documents are produced each year, and we have barely mined this information treasure trove for insights. With data analytics and artificial intelligence, we can demystify healthcare spending, design more practical provider networks, and help providers more wisely sift through care delivery options.

Here are four ways in which data analytics and AI can help bend healthcare's cost curve while improving quality.

1. Suggesting the most effective care providers

A 2014 United States Preventive Services Task Force noted that imaging tests to evaluate headaches were a common culprit contributing to millions in wasted healthcare dollars while offering little value to patients' health outcomes.

Imagine if, from the moment the patient with acute headaches began interacting with the healthcare system, their specific experience was matched with the doctor most knowledgeable about their condition who could then better filter out unnecessary tests.

Hospitals and provider networks track the health and wellness of their patient populations and have data about their patient mix based on their medical history. Leveraging data analytics, physicians can then be matched with patients based on past success in treating similar patient profiles.

This acknowledges the unique needs of individual patients, matches them to the unique capabilities of individual doctors, and reduces medical care runaround that drives up non-essential spending.

Imagine a physician asking a voice-activated medical AI (think Amazon Alexa) about the preferred, least costly surgical treatment for her patient. The artificially intelligent assistant returns the best matched surgeons, medically clears the patient for surgery, and software conducts a real-time auction among preferred surgeons. The surgical center then has options to find the best price for the patient based upon patient profile and risk.

2. Creating more detailed patient portraits

Currently, documentation for medical encounters remains effectively locked up within each hospital, clinic, surgical center, office and urgent care facility. Efforts to share and aggregate medical records for an individual are limited, but by aggregating records and using computers to decipher content, we could assemble full individual portraits of patient care.

With data from different sources, we can paint a more complete picture and personalize their treatment. Doctors would no longer have to guess at and fill gaps in documentation.

3. Identifying the best course of treatment

Machines are able to sift through millions of health records to learn what treatments work in similar patient profiles. Rather than depending on narrowly-designed studies, healthcare providers can base treatments and outcomes on data relevant to individual patients, their environments and how they live their lives.

4. More efficiently leveraging healthcare's most valuable resource: providers

With the rise of data analytics and AI, there will be a greater need for primary care physicians and less demand for some types of specialists as machines will play a greater role in analysis and diagnosis.

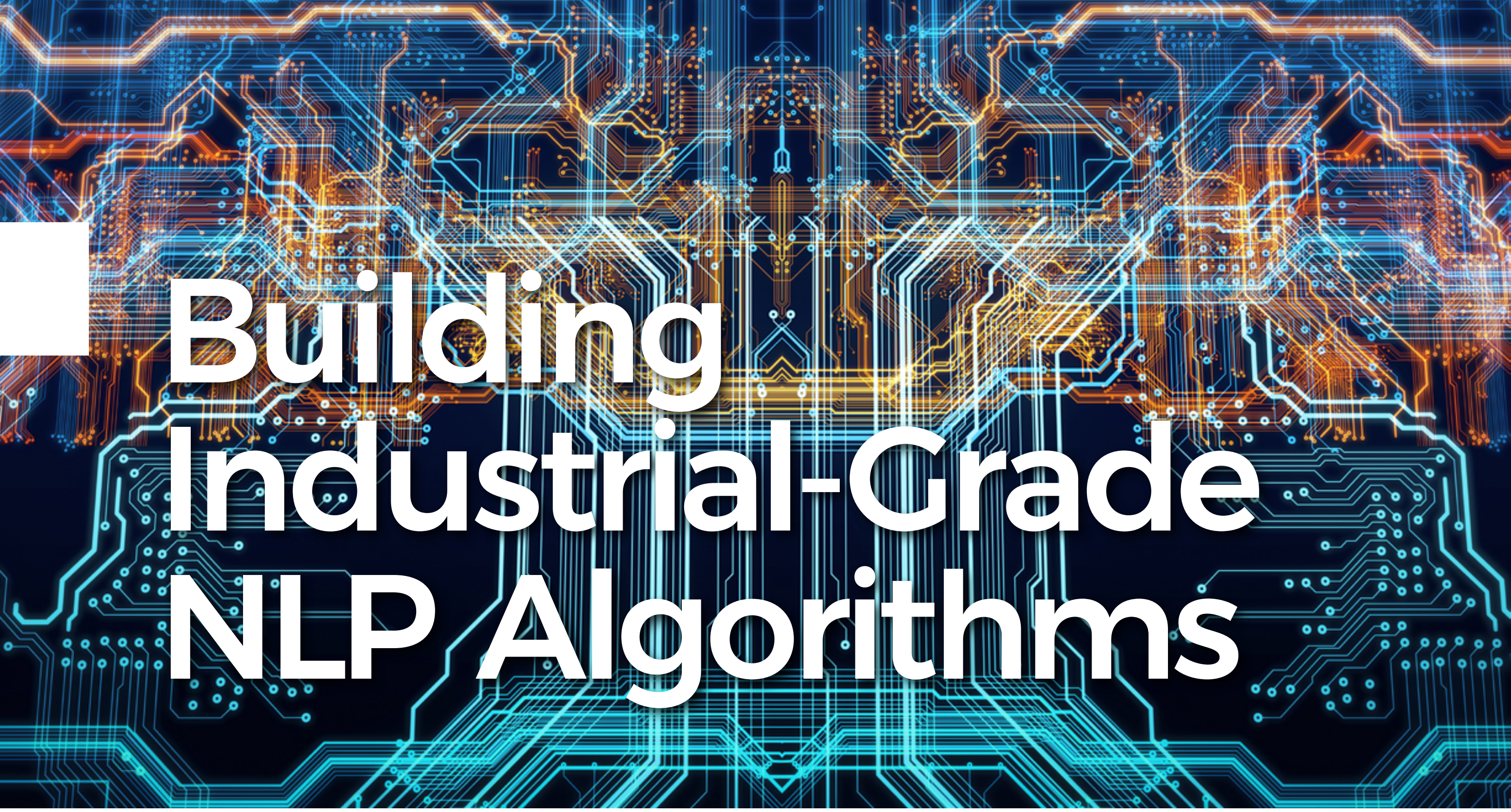
Certainly, the next generation of doctors and healthcare providers will need to understand the underlying mechanisms of health and disease and the basis for treatment, focusing on delivering what computers cannot: guidance, counseling and advocacy to help patients in the best way possible

But in this new world of healthcare, computers will be a vital aid in care delivery, automating processes and making sense of data. Meanwhile, physicians will have more time to lay hands. Given the high rate of burnout among physicians today, this should come as a welcome development.

In the next 20 to 30 years, data analytics and AI have the potential to transform the practice and consumption of healthcare from more of an art to a science and finally reach the utopia of high-quality, low-cost care.

The assumption in U.S. healthcare of tradeoffs between cost and key outcomes no longer has to be accepted. The world's most expensive healthcare system, with AI, can become its most successful.

This article originally appeared in the American City Business Journals (ACBJ).



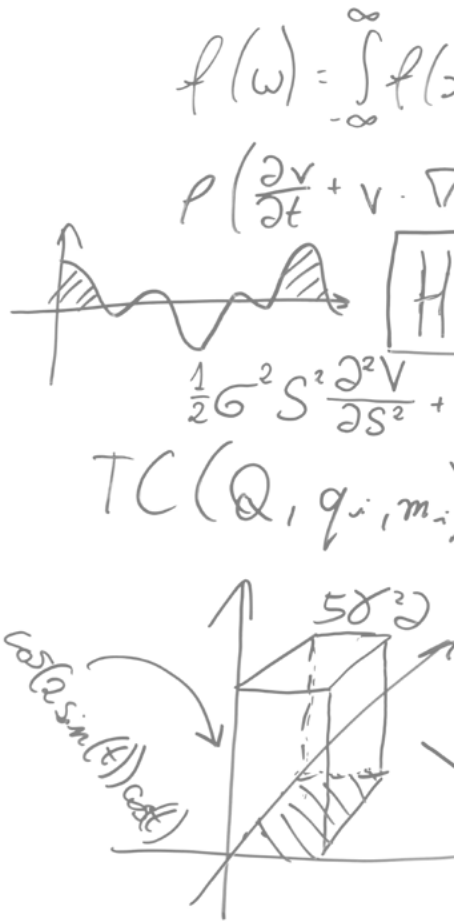
Building Industrial-Grade NLP Algorithms

Building Industrial-Grade NLP Algorithms That Impact Patients at Scale

Natural language processing (NLP) and healthcare reflect the rare case where a technology and a field evolve to the precise point where they can be useful to each other.

NLP is a field of computer science that uses machine learning and other techniques to extract meaning from the written word. It was first developed in the 1950s and has blossomed over the past 25 years with the availability of cheap, plentiful, and powerful computers. Medical documentation is the process of recording a physician's interactions with a patient during a clinical visit. With the increasing adoption of electronic health records (EHR) over the past 5 to 7 years, more clinical encounters are documented in machine-readable text rather than handwritten scribbles.

Applying NLP to machine-readable text has important use cases, including speech recognition for physician note transcription; physician note summary; structuring text into data elements; machine-assisted coding; and clinical decision support systems. Yet many of these use cases remain futuristic at best. While significant progress has been made, important obstacles remain to achieving the goal of better patient care through NLP.



Early Success Stories of NLP in Healthcare

Academics, the federal government, and vendors such as IBM Watson have reported on preliminary efforts to use NLP in healthcare. For the most part, these studies have focused upon limited patient datasets and concentrated on identification or prediction of disease within a population.

In 2013, the U.S. Department of Veterans Affairs used NLP to identify the prevalence of suicide attempts in 1.2 billion EHR documents. The algorithms developed were able to distinguish between screening for and actual mentions of suicide attempts in 10,000 patients with 80 percent accuracy.[1]

Beginning in 2014, IBM collaborated with Epic EMR, and Carilion Clinic in Virginia to identify patients suffering from heart failure based upon machine analysis of clinical notes. The idea was to extract clinical, social, and behavioral features from the medical record to predict heart failure.

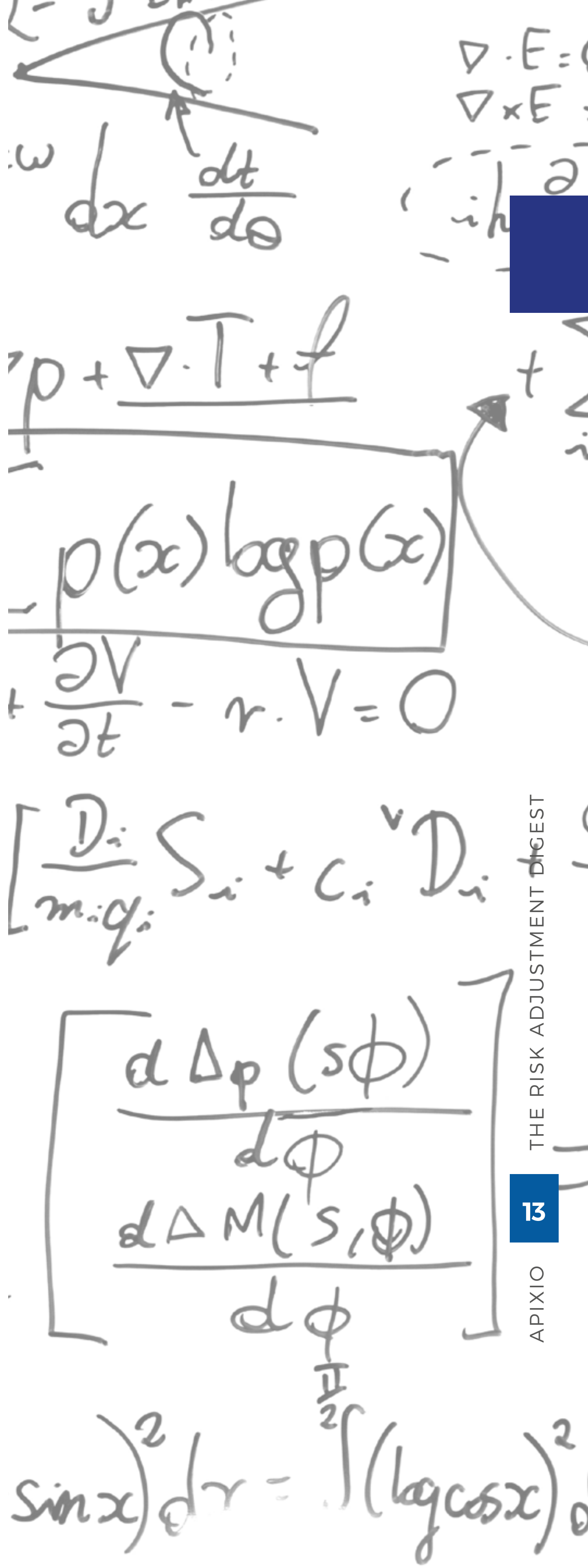
Most of this information is within the unstructured portions of the medical record as opposed to the structured fields mostly used for coding and payment, making this an ideal use case for NLP. The pilot program evaluated 21 million records in less than 2 months and identified at-risk patients with an 85 percent accuracy rate.[2]

In 2015, researchers at UCLA analyzed radiology reports, ICD-9 diagnosis codes and lab results in EHRs to detect cirrhosis (liver disease) in 4294 patients with 97% sensitivity and 77% specificity.[3]

These initiatives have moved medical NLP forward, yet they remain localized interventions. What will it take to build industrial-grade algorithms, ones that do not operate on a small population in a test situation but which impact patients at scale in the real world?

Three Steps to Moving NLP Into the Medical Mainstream

First, algorithms need to capture 95 to 98 percent of “truth” (or 95-98% of all “positive” answers in a sample set) to be useful and therefore adopted by healthcare clients. Healthcare organizations demand very high accuracy for solutions that address diagnosis, treatment, or reimbursement, as they should. This is certainly not the case for other industries in which machine learning is used, such as ad engines, social media, and e-commerce.



In fact, even the accuracy bar required for a study to be published in a peer-reviewed journal is lower than what a hospital or health plan CEO would require.

It is not good enough to achieve 85 or 90 percent of the desired insight in healthcare (see below). That means 10 or 15 percent of the truth missed, which could result in a misdiagnosis or erroneous treatment or critical lost revenue.

It is the effort to traverse the distance from 80 percent to 98 percent (see below), which will cost many times more than what it took to achieve the initial 80 percent.

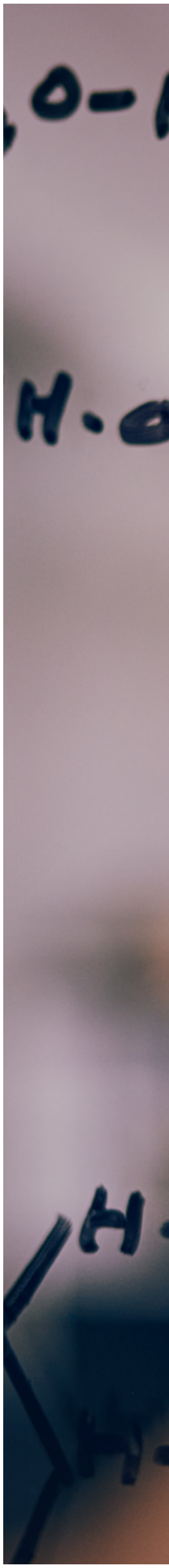
Even more important, the algorithms must be generalizable – that is, they must work well across different populations, clinical settings, document types, and data sources. In other words, they must be "scalable." This is what separates industry-grade algorithms from pilots, proof of concepts and academic studies, which is category that the large majority of NLP in healthcare currently resides.

Lastly, any attempts at using machine learning techniques must consider the ability to find truth and the cost of that effort. Consider an example of gold mining in a mountain. A company can use brute force and dig up tons of dirt to capture most of the available gold nuggets. But to sell the actual gold it must sift through lots of rock and gravel that will be tossed aside. Alternatively, the company can spend more time and money to develop a powerful metal detector that can direct the dig. Most of the gold can then be found with much less rock and gravel extracted.

Similarly, initial efforts at learning algorithms may yield most of the desired insights in a medical record (e.g. heart failure) while producing lots of irrelevant answers, which need to be sifted out at great cost and energy. This is a catch-22 situation though, as it is also expensive and time consuming to evolve an algorithm to the point where extraneous results are no longer found. A technologist must balance the cost of algorithm refinement with the cost of results QA in order to develop sustainable technology.

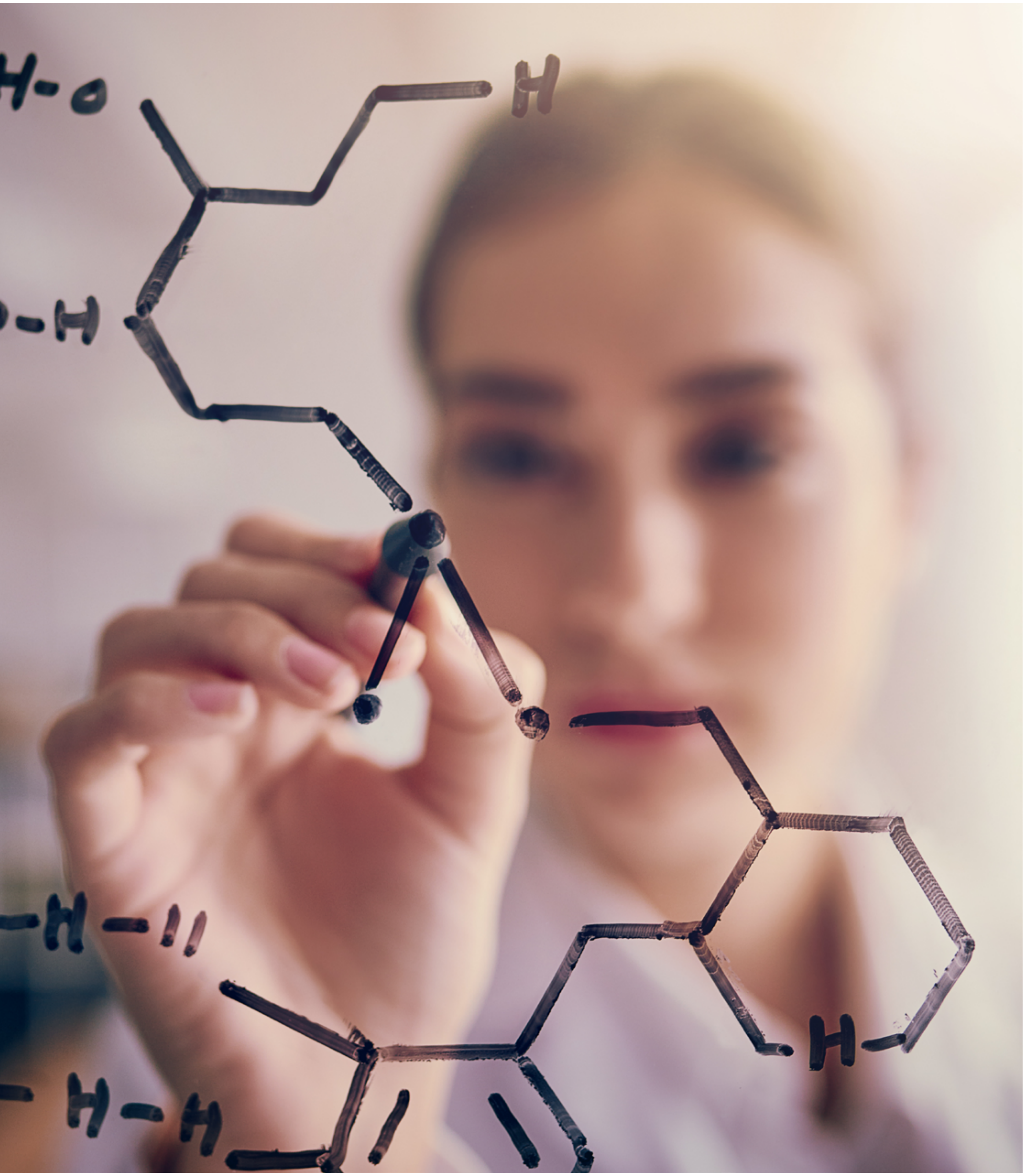
A Word of Advice on Reaching the Finish Line

Life is a game of inches: The hardest part is traversing the last 5-to-10 percent of a path towards the goal line. This is true in the world of refining analytics to solve a problem to meet user needs. And, it is particularly true when it comes to creating a solution applicable in a variety of scenarios. We are so close to delivering impactful NLP-powered applications in healthcare. These applications will provide the basis for learning from real-world care what treatment works, what doesn't, and how to improve care delivery. Of course there is more to achieving this level of performance, such as domain experience, feature engineering, and a ton of training data. But that is another article.



---

1. Hammond, Kenric & Laundry, Ryan & OLeary, T.Michael & Jones, William. (2013). Use of Text Search to Effectively Identify Lifetime Prevalence of Suicide Attempts among Veterans. Proceedings of the Annual Hawaii International Conference on System Sciences. 2676-2683. 10.1109/HICSS.2013.586.
2. IBM: Natural language, machine learning can flag heart disease. EHRIntelligence – https://ehrintelligence.com/news/ibm-natural-language-machine-learning-can-flag-heart-disease.
3. Chang EK, Yu CY, Clarke R, Hackbarth A, Sanders T, Esrailian E, Hommes DW and Runyon BA (2016) Defining a patient population with cirrhosis: an automated algorithm with natural language processing. J Clin Gastroenterol 50, 889–894.



CASE STUDY: MSP* & APIXIO

About the Plan:

- Not-for-profit multistate payer (MSP) based in the Midwest.
- 40,000 Medicare Advantage lives and 200,000+ commercial health exchange lives.
- Internal team of 9 expert coders.

About this Project:

- MSP's decision to bring all coding processes in-house sharply increased workload.
- MSP reluctant to rely on outsourced chart review for help, given increased government scrutiny.
- Apixio won business after successful pilot comparing application based coding with manual coding.

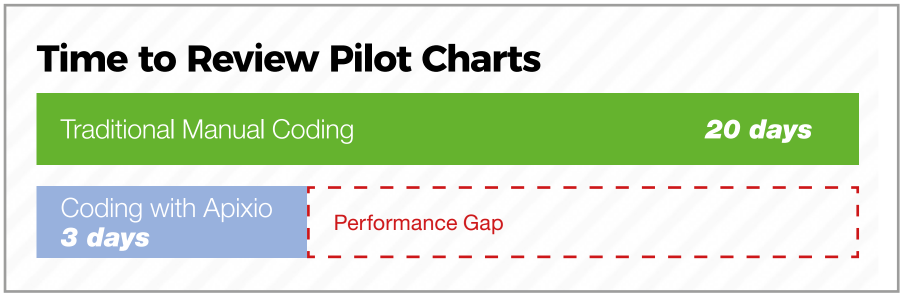
*Pseudonym

Apixio Elevates Performance Past Manual Review

Because MSP wanted to increase its coding bandwidth without taking coding outside, they were attracted to the HCC Profiler, an application that augments coder activity, enabling them to confirm the evidence behind every new, supported HCC code without the drudgery of reviewing every page of every chart.

MSP decided to do a pilot study of the HCC Profiler to test whether it could enable their coders to be faster than traditional manual review while maintaining—or increasing—their accuracy levels. Using one set of charts, they conducted retrospective chart review both with and without the HCC Profiler.

The results were conclusive. An MSP coder took twenty business days to review 1,012 charts without the HCC Profiler. She took three days to review with the HCC Profiler, and found 15 more supported HCC codes. Based upon these strong results, MSP began to work with Apixio.

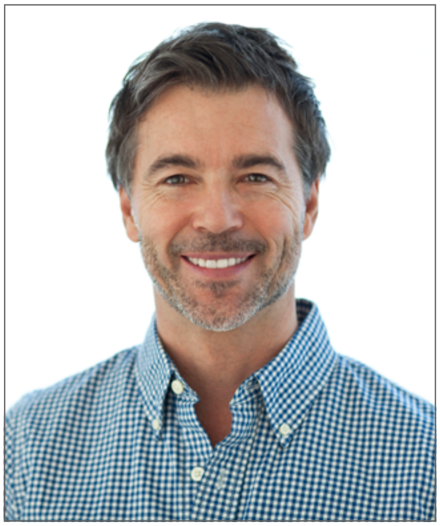


Superior Customer Service

A risk adjustment manager at MSP, says simply of Apixio's customer service, "It has been a great experience all around, I've received strong updates throughout the process and have seen good value." MSP appreciated the consideration from the contracting process — where Apixio's finance team personally discussed the many redlines and questions MSP had—through to the project itself, where Apixio's operations team responded rapidly to any inquiries.

Managerial Controls and Coding Features Impress MSP*

- Taken to encounter note with evidence instead of hunting and searching through chart
- Visibility into coder decision-making process; for education and documentation improvement
- Dashboard with individual coder's metrics such as productivity and accuracy rates
- QA can overturn "accept" or "reject" decision while providing rationale for coder
- Nearly 7 times faster than manual review.



AI IN HEALTHCARE:

Crossing the Bridge from Theory to Practice

Artificial Intelligence (AI) is poised to help make healthcare more affordable and personalized. In the hands of domain experts, AI can unlock insights from real-world clinical data to improve delivery, outcomes, and costs.

A recent CB Insights report noted that there is more venture investment for companies using AI in healthcare than any other industry; about $1.8 billion across 270 deals since 2012. Despite these investments, there is a wide chasm between the promise of AI in healthcare and the reality.

According to a McKinsey & Co. survey of C-level executives at more than 3,000 companies across 14 economic sectors, the healthcare sector has one of the slowest rates of AI adoption. One reason for this relatively low adoption is that healthcare is different. It is complicated, opaque and highly regulated. In order to use AI as a technique to solve healthcare problems, there are three principal challenges to overcome.

Challenge #1: Data Access

Data acquisition is a major bottleneck. Large amounts of varied data sets (text and codes) are required to train and test learning-based algorithms. Unlike social media or ad tech, accessing healthcare data is a more complicated matter. And we aren't just talking about challenges associated with privacy and security regulations such as HIPAA.

There are powerful disincentives dissuading hospitals and health systems from sharing medical records data. Medical records are seen as a competitive advantage for marketing, contracting, and patient acquisition and retention. Health information exchanges have proven largely unsuccessful in overcoming this because they offer no sustainable business model.

While Medicare provides access to its billing data for more than 40 million individuals with recorded diagnoses and procedure codes, this data is not nearly as rich as hospital or clinic encounter notes. Any algorithm that proves to scale across a large population will need to be trained with data not limited to a few health systems or academic centers.

The key is to acquire fully identified medical records from a variety of settings and practices for training and testing data. Those who hold the keys to this data will unlock the records once (and if) they see the value in doing so. But if AI systems don't get enough labeled data, the algorithm performance will not reach the level healthcare users require.

Healthcare demands higher performance than ad tech or social media. A false positive (or negative) that results in an ad appearing on a webpage for Cheerios has very few real-world consequences. Whereas, false positive results in healthcare will cause a physician or consumer to ignore the next healthcare recommendation, and a false negative result could have grave implications on someone's health.

Challenge #2: Getting "Wins"

AI is a fascinating technique, but, by itself, it looks like another marketing buzzword destined for the dustbin. To build trust and credibility, AI-powered technology needs to solve fundamental healthcare problems, and that hasn't happened yet.

IBM and the M.D. Anderson Cancer Center created a highly touted partnership, Oncology Expert Advisor (OEA), to harness IBM Watson and provide machine-generated advice for cancer care. After failure to get off the ground, the project was finally shut down. IBM tried to solve one of the hardest problems - cancer care decision support. It's failure to do so was a setback for AI use in healthcare. We know that there are going to be many false starts when applying new technology, but the relentless marketing hype didn't help.

There are other attempts at using AI: image recognition to assist radiologists; text recognition to assist coders for diagnosis capture; and voice recognition for patient home care assistants. Credibility will be earned once AI achieves smaller, meaningful "wins." Architects of those such solutions must apply healthcare domain knowledge from the outset. It is just as critical the solutions address problems that healthcare payers, providers, or consumers will pay real dollars to solve.

One word of advice to entrepreneurs - look outside of the clinical realm and into the administrative space to find an initial problem to solve. I know this area is not sexy, but, it offers a chance to try these techniques in an environment with a higher tolerance for failure. Smaller wins will enable entrepreneurs to gain the trust and credibility of a skeptical audience.

Challenge #3: Designing User-Friendly Applications

In healthcare, domain knowledge and experience matters. There are far too many technologists developing digital health solutions without input from industry experts. Specifically, we must understand the needs and preferences of end users.

Healthcare is riddled with technology solutions that fail to simplify users' daily tasks. For example, there are electronic health record systems with endless alerts and circuitous click patterns and applications that don't integrate with the rest of an organization's IT infrastructure. The interfaces through which data is made available to providers and administrators must be intuitive and enjoyable. The technology must serve the user, not the other way around.

AI and data analytics are at the forefront of the healthcare revolution. Or at least, they should be. When implemented properly, AI technology can can enable a "smarter" healthcare system that treats patients efficiently and well. It is precisely because of this potential that it is important to overcome AI's bottlenecks to implementation and truly turn theory into practice.

This article originally appeared in HealthcareDive.

The Wait is Over:

How We Can Achieve Widespread Interoperability Now

Let's face it, we don't have meaningful widespread data system interoperability yet, but rather than place blame on providers, vendors or the government. Let's take a closer look at the problem.

The mission of interoperability is to make data from electronic health records (EHRs), doctor notes and other relevant patient data sources available in systems and contexts that are different from the original source. The expectation is that interoperability will enable efficient care via information sharing and a reduced need for duplicate tests. This will ultimately improve health outcomes and drive down costs.

However, efforts to achieve interoperability have not yet found widespread success. Interoperability is not in and of itself a hard technical problem to solve. Other industries, notably financial and logistics, have interconnected systems that work. So why have we struggled so much with achieving interoperability in healthcare?

Avoiding the "Field of Dreams" Scenario

First, let's dispel the notion that interoperability is currently nonexistent. There is some interoperability, but it's narrowly confined to managing adjudicated claims streams. Other initiatives, like Health Information Exchange (HIE), portals and the migration of patient data across systems have been far less fruitful when you consider their widespread benefits and adoption. In these cases, interoperability has been pursued more as a theoretical benefit than a solution to a specific use case with far-reaching demand.

This is a classic "field of dreams" scenario, where we're trying to build a product that lacks a defined user or widespread demand and doesn't clearly fit into existing workflows. Defining a set of standards and hoping for widespread implementation is another form of this thinking. Material demand for intersystems interoperability needs to drive standards and technologies. We have to expand our thinking beyond "one-size fits all" to solve these challenges.

There are noteworthy interoperability successes in healthcare. For example, OCHIN, an Oregon-based nonprofit, recently launched a real-time data aggregation system that serves more than 170 organizations and boasts interoperability between different implementations of the same electronic health record (EHR). Another good example is how users of Epic EHR systems can request patient charts from other institutions that use Epic. Note the use-cases. In the first case the system is a locally brewed solution created by like-minded providers and the second is available to users within a vendor's ecosystem.

How Do We Improve Interoperability?

The best way to achieve interoperability and realize the benefits is to approach the problem one use case at a time and pursue more targeted solutions.

Do we want practitioners of a certain specialty like oncology to share in a common pool of knowledge countrywide? Do we need clinical data shared with researchers? Do we need to pull data for analytics or reporting? Do we want to meaningfully migrate a chart between two EHR systems? These are the use cases that will drive the technology and solutions that are ultimately built and adopted.

We need to let go of the build-it-and-they-will-come approach where interoperability is an inherent good that will inevitably confer benefits if everyone would just play the game fairly. The benefits of this top-down approach are too diffuse and long-term for the market to cooperate in the near future.

Instead, we must start at the grassroots market level and look for the specific use cases to solve. Doing so will do more to define interoperability standards going forward than any single effort.

This article originally appeared in www.himss.org.

Health Tech On the Horizon:

Blockchain

Cryptocurrencies have exploded in popularity over the past several months, with everyone from boxer Floyd "Money" Mayweather, who declared "You can call me Floyd "Crypto" Mayweather from now on," to the Winklevoss twins (named the first bitcoin billionaires) investing in digital currencies like Bitcoin and Ethereum.

For healthcare however, the quick climb, dramatic falls, and dizzying climb again of cryptocurrencies have interesting implications, beyond their potential as valuable assets. It's actually the revealed potential of the underlying blockchain technology to impact healthcare in big ways across the next decade that has our attention.

The 'Blockchain 101'

Blockchain is a digital method of making, recording and validating events or transactions. Imagine an infinite digital receipt that creates transparency and significantly reduces fraud risk. At its core is a reliance on a decentralized approach to authorization and edits.

Who made this revolutionary tech? No clue. There's lots of conspiracy theories out there, but we won't dive into them here.

With blockchain three existing technologies are applied in a novel way, as the cryptocurrency blog Coindesk writes. First is private (secret) key cryptography— the use of unique key to encrypt and sign transactions, delivering currency or data via "public key' for recipient access. Second is a distributed network with a shared ledger – this means independent, unique copies of the ledger are held across the entire network, but simultaneously updated only when a transaction is verified by all, making it impossible to edit or rewrite. Third is an incentive to service the network's transactions, record-keeping and security: Simply, this creates an incentive for computers to lend their processing power to the blockchain endeavor.

Blockchain has rapidly transitioned from its beginnings in the virtual currency ecosystem to more traditional industries like mainstream finance, manufacturing, and healthcare in a big way.

Finance, seeing obvious value in verifiable transactions, has lead the pack with major investment by big firms in blockchain currencies or developing blockchain based payment and transaction systems. Financial transactions depend on parties having trust and good faith. Potential conflict is eliminated by blockchain which certifies the execution of the conditions for the transactions. This is one of the reasons the technology is also attractive to healthcare.

Potential and current implementation in health

Blockchain could have far-reaching effects on the healthcare industry in several ways.

The first is electronic medical record (EMR) interoperability. Critical to EMR interoperability is a trusted, authenticated environment for clinical reference and decision making that meets privacy and security concerns. This is where blockchain's ability to timestamp and keep a central point of truth – while managing simultaneous multiple user access comes into play. A new physician would simply be added to a patient's blockchain and from there access the same information as everyone else already participating. Several organizations are already pursuing initiatives along these lines.

For example, researchers at MIT's Media Lab developed a blockchain-based system called MedRec, which lets patients authorize access to their private, secure EHR-approved changes, and govern sharing between disconnected providers.

A second area of impact is claims processing and reimbursement. Blockchain can improve revenue cycle management and reduce the payment inefficiencies plaguing providers. One avenue might be 'Smart contracts' – made possible by blockchain's transparent transaction and event validation. Validations might also impact claims adjudication and billing management. This could cut the high instances of medical billing and reimbursement related fraud; a linchpin in lowering overall system costs.

Also central to industry interest in blockchain adoption is health data security. Depending on the technology's implementation, patients who are part of their EMR blockchain may be able to approve or deny any sharing or changes to their data, helping to ensure a higher level of privacy and greater consumer control.

Cyber security experts say health plans and systems can expect to see unprecedented data breaches through 2018. Blockchain can help because it enables the gathering and integrating of data from a distributed network of participants in the healthcare value chain. To break that down – a thief can't steal all your belongings if if they're scattered across many, many homes. And as discussed earlier, every blockchain event or transaction is time-stamped and then unalterable, ensuring the source and integrity of data, and establishing trust among participants so they can share data securely.

Blockchain does have its risks. For one, what does it mean to keep staggering dollar amounts of real world investments in digital wallets (a January 3, 2017 market cap of over $700B and counting) that can be hacked and stolen, disappearing overnight? In a similar vein, confidential medical records could be a valuable cybercrime target.

Possible impacts for big data and healthcare companies

One of the greatest promises of big data in healthcare is that of precision medicine. This requires access to all available data, which we discussed in an earlier article about EMR interoperability. But the increased accuracy and integrity of healthcare data due to blockchain may allow AI platforms like Apixio's to train their 'machines' better, opening doors to things like predictive medicine even faster.

In truth, widespread implementation of blockchain for healthcare is on the distant horizon. Bitcoin and Ethereum et al may be hot news, but healthcare has long demonstrated that the underpinning architecture has both long term profitability, and societal good.

How 4 Different Organizations Used **Machine Learning to Transform their Risk Adjustment**

At the 2017 RISE Nashville conference Apixio hosted a panel with four of our clients: Doug Loop of OptumCare, Russ Shust of Group Health Cooperative, Jennifer Pereur of Hill Physicians Medical Group, and Alicia Wilbur of Martin's Point Health Care.

The panel was moderated by Tam Pham, a widely respected risk adjustment expert with experience at SCAN Health Plan and the University of Pennsylvania Hospital.

A crowd of over 150 showed up to hear these four leaders discuss their thoughts on risk adjustment technology and the challenges facing the industry today. If you missed it, we've got you covered—highlights are below:

Strategies to reduce provider abrasion over chart pull

The panel discussed strategies for securing patient charts while maintaining strong relationships with physicians. Physicians notoriously chafe at pulling charts for risk adjustment, because they just don't see any benefit on their end from this process.

"We started working with Apixio, and the value that this technology has for us first and foremost is provider abrasion. We're tired of the same process year after year. Our physicians are tired. I was seeing a real drop in compliance in just getting chart retrieval. Every year we were getting a smaller portion of the charts we were targeting.

The way we use this technology is by directly accessing the EHR., We don't have to bother the office, we just pull the data., We grab all that unstructured data, we run it through the machine, get that valuable information, and there is no burden that's placed on the office. That's a real win for us."

Jennifer Pereur, Hill Physicians Medical Group

Two of our panelists laid out how Apixio helped them manage this dynamic in their organizations. Jennifer Pereur described how Hill Physicians moved from manual chart pull to direct EHR extraction for a portion of its providers, enabling them to secure the necessary information while maintaining high physician engagement.

Alicia Wilbur of Martin's Point Health Care in Maine had a very different situation. Her physicians didn't want Martin's Point to touch their EHR at all, and were comfortable with their existing chart chase vendor who scanned their charts into PDFs. Apixio was able to process and analyze these charts very smoothly, so Alicia didn't have to make any other asks of her physicians.

> *"With this technology and being able to read PDFs, it's really helped us with that portion of our chart review. We're at a place where we're about 50/50 with EHR reviews and with the traditional chart chase PDF reviews, and it's really added some value to that PDF review for us in particular."*
>
> Alicia Wilbur, Martin's Point Health Care

The panel also discussed ways to give back to physicians during risk adjustment, so they truly see a benefit from the process, and buy in. Russ Shust floated the idea of giving risk adjustment analyses back to physicians in real-time—telling them when they had missed asking a legacy diabetes patient about their diabetes, for instance. To Russ, creating a joint initiative around quality will help physicians feel like the risk adjustment was a two-way street.

Increasing the speed and scope of chart review

The panelists discussed their hopes to include different types of documentation in the risk adjustment process. Jennifer said that she wants to use authorization requests and care transition documentation. This, she said, is where real efficiencies will be created and how the true picture of patient health can be found— by looking at interactions across the clinical spectrum.

Martin's Point is already looking at all outpatient records for Medicare Advantage patients, and Alicia said that in 2017, they will also look at all inpatient records. Additionally, they are partnering with an outside quality auditing firm, to receive the documents that they pull as well.

Alicia and Russ made the point that technology like Apixio's is exactly what enables organizations to increase the scope of their review. Because Apixio speeds the review process, organizations are able to take on more documentation and data instead of worrying if they have the bandwidth to process it. Speaking to this fact, Alicia said: "Technology is helping us create bandwidth within our coding team to be able to look into those records."

> *"Time is limited, if you had a finite number of coders to look at your medical records, why not serve them the data they need to look at vs going after a hundred pages of a medical record. The efficiency of a tool like this, it's impressive."*
>
> Russ Shust, Group Health Cooperative

The importance of strong quality assurance programs

Every single one of the panelists emphasized the need to QA as many risk adjustment results as possible, to have full confidence in the accuracy of the process.

Russ mentioned that during GroupHealth's initial project with Apixio, they gave a group of charts which had already been looked through three times (by an auditor, another vendor, and a coder); Apixio was still able to find additional supporti ng HCC evidence. They then audited Apixio's results to confirm accuracy.

Alicia mentioned that Martin's Point used a similar process, and gave Apixio charts that had been looked through by their coders several times. They gave Apixio's findings back to their "quality assurance guru" for validation, and found that they agreed with 94% of Apixio's results.

Accuracy was of paramount importance to the panelists because of increased CMS scrutiny on risk adjustment in general. Russ mentioned that an especially important part of his risk adjustment process was validating existing codes by cross-chec king that they linked up with a qualifying claim.

Tam pointed out that all RAPS and EDPS files should be given to vendors before the chart review process begins. This will allow for efficiency in the review process to identify conditions that have not been previously captured or to validate conditions which were previously submitted.

The future is bright for risk adjustment

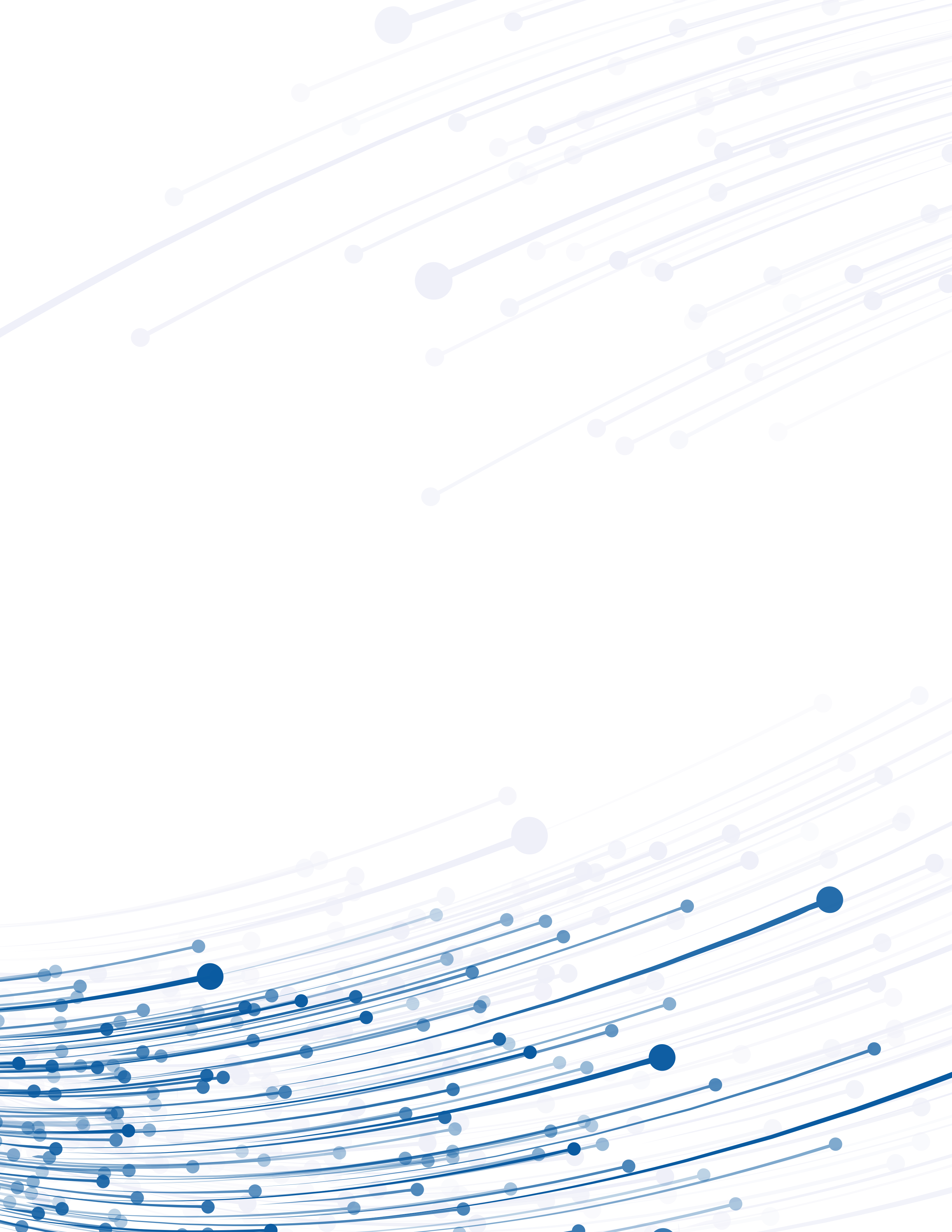
The panel is hopeful about the future of risk adjustment. Doug Loop mentioned that the CMS advance notice didn't include marked changes to Medicare Advantage, signaling that under the new administration the regulations might remain static, in the near term. And Jennifer mentioned that, if anything, it seemed like risk adjustment might move into new spheres in the coming years, potentially even commercial risk in California.

Jennifer noted that the more widespread risk adjustment became in insurance programs the easier it would be to do high-quality risk adjustment. The burden of engaging physicians would become less once risk adjustment affects more of their patients.

You can watch the entire presentation here:
https://youtu.be/gMP3Z7quit4

Notes

Notes



APIXIO©

/ f / in @Apixio / apixio.com

info@apixio.com 1825 South Grant Street #350, San Mateo, CA 94402