WHITE PAPER

Healthcare
Predictive Analytics

Predictive Analytics In Healthcare

A data-driven approach to transforming care delivery

Author

Andrew Bartley
Senior Health and Life Sciences Solution Architect, Intel Corporation

"The powerhouse organizations of the Internet era, which include Google and Amazon... have business models that hinge on predictive models based on machine learning [1]."

The global healthcare industry is transforming, driven by increasing costs and an aging population. By 2025 the global population is estimated to be 1 billion, with 500 million of those people over the age of 50. Multiple sources including the World Health Organization and UN predict that in that same year over 70 percent of illnesses will be chronic conditions. Place that against a backdrop of increasing global healthcare spend that is expected to reach USD 18.3 trillion by 2030 and change is a clear imperative[2].

These factors are forcing a global healthcare reform to move from volume-based payment models to outcome-, or value-based models. This shift requires providers to significantly change the way they operate. They must focus on the delivery of personalized medicine that is tailored for each patient, while also deepening their understanding of population health more broadly to better identify and respond to patterns. Both objectives can only be met by applying advanced analytics to a healthcare provider's data.

Table of Contents

Three global trends in healthcare have the potential to impact this transformation.

1. Digital transformation means more data

The digitization of healthcare is resulting in the creation of massive new data sets. One of the most significant examples of this is the Electronic Medical Record (EMR). Over 85 percent[4] of healthcare organizations in the US have now adopted an EMR system[5]. Health claims data, radiology images and lab results are other common data sets. In the future 'omics (i.e. genomics) data is expected to grow significantly as precision medicine gets increasingly integrated into standard clinical workflows.

> "Digitization has allowed us to have a more meaningful conversation around how to use predictive analytics to improve patient outcomes[3]."

Outside of the clinical setting patients generate a significant amount of quasi-health data through the use of wearable health and fitness trackers and health applications. The proliferation of health-related Internet of Things (IoT) devices, as well as apps and other health-monitoring technologies also generate a huge amount of data every day.

Healthcare providers have an unprecedented opportunity to incorporate these disparate data sources and structures into new solutions. For example, to understand population health trends, it is important to use external resources such as social media or publicly available government data, in addition to a hospital's own records. However, 94 percent of hospitals[6] are not currently capturing enough information to carry out effective population health analytics.

2. External pressures to meet the Quadruple Aims

Regulators, staff, patients and peers within the industry all have expectations of the healthcare providers that they come into contact with. The Quadruple Aims[7] that each organization must commit to addressing are:

- **Improving access** to resources and services for patients and clinicians, enabling them to find what they need, when they need it, securely and quickly
- **Reducing operating costs** and ensuring that limited financial resources can be channeled towards the most value-adding priorities while optimizing the bottom line
- **Improving treatment outcomes** by helping physicians to make diagnoses faster, deliver the most personalized treatment possible for the patient, and reducing average length of stay
- **Optimizing provider satisfaction** by ensuring that busy staff, working in a stressful environment, are able to focus on delivering patient care and achieving positive outcomes rather than grappling with complex tools and workflows

3. Government reform and policies drive change

Government policy can influence healthcare providers' data strategy in various ways. In the US, for example, policies such as the CMS EHR Incentive Program[8] and the HiTech Act[9] have increased investment in healthcare since 2011. This has accelerated the adoption of digital solutions, which in turn has led to a boom in the volumes of data a typical healthcare organization holds. Six years on, many are reaching the point at which they have a large enough pool of current and historic data to make really meaningful analytics possible. Similar trends can be seen or expected in Europe, where government initiatives in the UK, France and Sweden[10] are also spurring investment in this area.

Meanwhile, around the world, governments are placing more pressure on providers to demonstrate and track compliance with data protection and other regulations, often across lengthy processes involving multiple people and departments. Rising pressure to curb spending and the need for improved outcomes has pushed responsibility for managing the risk of poor outcomes and high costs from the payer to the provider.

Advanced analytics is closer than you think

The ability to meet many of these challenges now forms the basis of key performance indicators (KPIs) for many healthcare providers. Failure to respond can result in competitive disadvantage, a negative impact on bottom line, reputational damage, unhappy patients, or loss of staff.

Real-world results: Penn Medicine

Penn Medicine operates a network of healthcare facilities in Pennsylvania and southern New Jersey. Working with Intel, it has created a collaborative data science platform, based on the open source Trusted Analytics Platform* to help clinicians make faster, smarter decisions based on large-scale clinical data and big data.

In its first trials of the platform, called Penn Signals*, the organization developed algorithms to help predict and prevent two of the most common and costly issues for hospitals: sepsis and heart failure. It was able to correctly identify about 85 percent of sepsis cases (up from 50 percent), and to make these identifications as much as 30 hours before the onset of septic shock (as opposed to just two hours prior, using traditional identification methods). This meant clinicians could deliver treatment sooner, speeding time to recovery for the patient and saving resources for the hospital.

Meanwhile, Penn Medicine estimates that using traditional methods of diagnosis, between 20 and 30 percent of heart failure patients had not been properly identified. With the predictive model it was able to identify these patients and give them the education to self-manage their condition successfully. As a result, readmission rates dropped.[11]

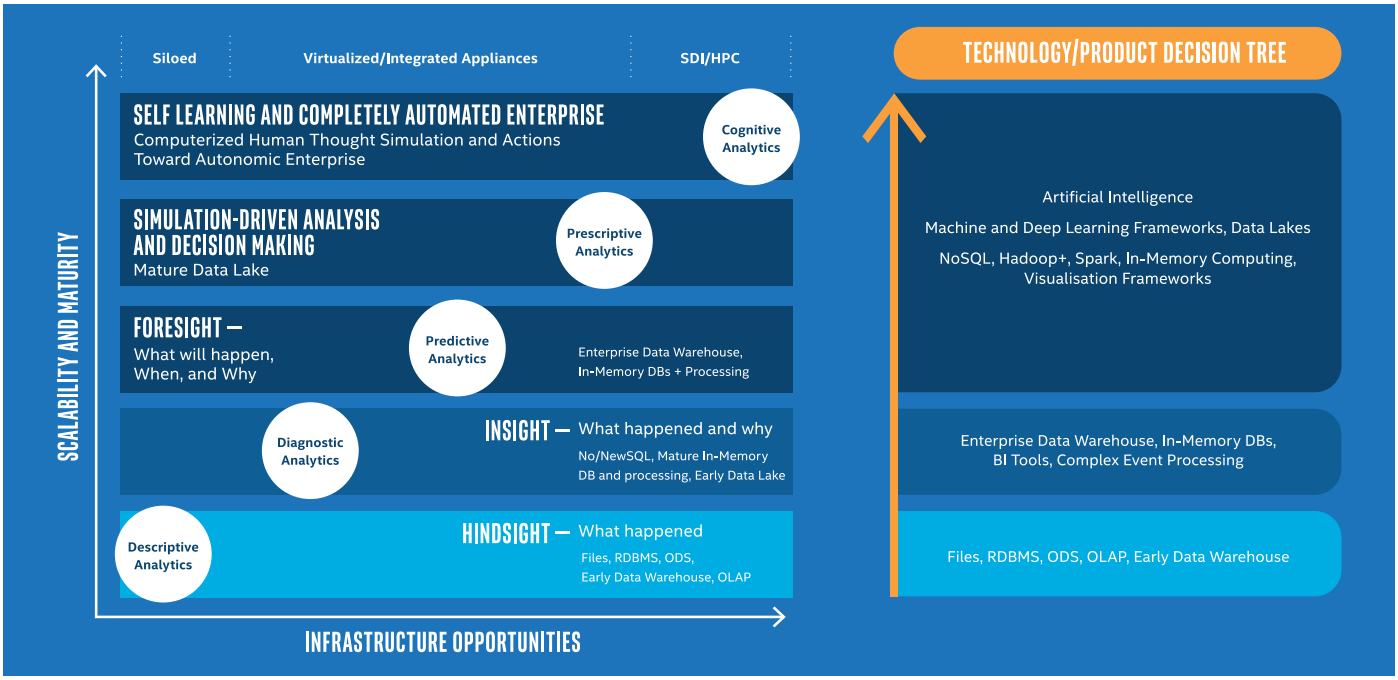


Figure 1: Stages of analytics maturity and associated technologies

Analytics has already proven helpful here. A recent Intel-commissioned report[12] from the International Institute for Analytics found that the highest performers in analytics in healthcare are using it to help improve patient engagement, population health, quality of care and business operations – areas that map closely to the Quadruple Aims.

However, for those not already proficient in analytics, knowing where to start can be tricky. Advanced analytics techniques like machine learning and artificial intelligence (AI) promise ground-breaking insights and the potential for great innovation, but they can also be complex and daunting.

The good news is that if you are doing any form of business intelligence (BI) today, then you are already on the path to advanced analytics. From there, it is simply a question of evolution and steady progress.

Analytics capabilities develop in stages (see Figure 1), each building on the one before and adding the ability to handle more complex data and draw more detailed insights in a more timely fashion.

The first two stages constitute traditional analytics methods. Descriptive analytics (or BI) reviews data about what has happened and answers questions about past actions. Diagnostic analytics takes this a step further by adding insight into why past events happened as well. With these types of analytics, you can for instance review your admissions data from last month and identify that the number of flu patients increased by 18 percent compared to the previous month.

The more advanced stages of analytics bring in capabilities that can add significant value to strategic planning, such as forecasting, real-time insight and automated decision making.

Predictive analytics analyzes historical data to predict future target events. This transition to forward-looking analytics is an important crossover for an organization from both a technology and business process perspective. Predictive analytics uses a variety of statistical and machine learning methods and are honed over time with the addition of new data. An example of predictive analytics would be to use historical data from the hospital's records along with external sources such as weather forecasts and social media to predict peaks in ER admissions for the purpose of improving staffing levels

Prescriptive analytics builds on predictive analytics by including a single or set of recommended actions based on the prediction. Prescriptive solutions can employ data techniques like simulation and machine learning as well as cognitive systems. Building on the previous example, a prescriptive solution would develop a prediction for peaks in ER admissions, and then recommend adjustments to workflow or prioritization to reduce length of stay and ensure staff can see as many patients as possible.
These recommendations can be based on discrete business processes, or in the case of cognitive systems can be based on previous activities that produced an optimal result.

Cognitive analytics uses artificial intelligence (AI) technologies such as machine learning, deep learning and reasoning or logic systems. It automates decisions using human-like analysis or provides insights and suggestions to augments human decisions. This could help create a full view of a patient's state of health, considering everything from EMR records to CT scans and the patient's fitness tracker

What is AI?

AI defines a class of techniques designed to enable machines to sense, reason, act, and adapt like humans do. There are several underlying technologies that make up AI which are shown in figure 2. Think of these as various tools that can be used to solve different types of data problems. As your AI projects become more complex, you will likely need several of these technologies as part of an end-to-end solution.

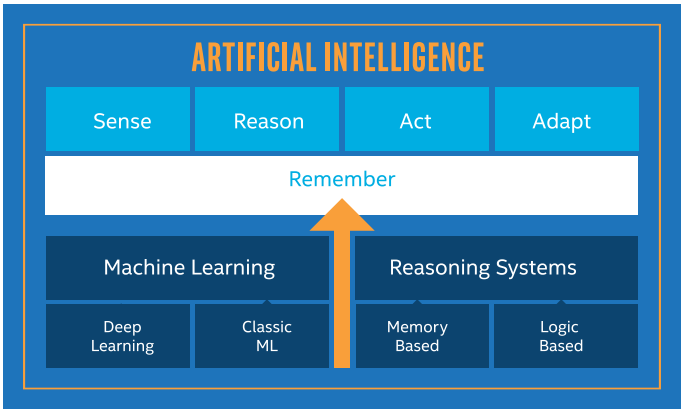


Figure 2: The technologies behind AI

feed, and then make proactive, real-time recommendations to the physician about possible treatment options or risk factors, and send alerts to the patient with suggestions on how to improve their wellbeing.

When starting out on your first advanced analytics initiative, aim to get value from predictive analytics. A focused business case with measurable ROI will help ensure proper organizational alignment. You can then move on to prescriptive and/or cognitive analytics whenever and however is appropriate for your organization. Remember though, that the goal isn't always full-scale cognitive AI – in many cases other types of analytics are more suitable. Success lies in finding the right tools for your organization, not necessarily in making it to the 'end' of the maturity journey.

Building a predictive analytics strategy: where to start

Form your predictive analytics roadmap by considering three key areas: people, process and technology.

People

Building a solid advanced analytics team can be a challenge. Competition for scarce talent is fierce. Healthcare has a unique competitive advantage inherent in the purpose of the work it does. Much like those of us working in and around healthcare, there is purpose in knowing that the work you do contributes to saving lives. Using that advantage is one way to differentiate from other organizations that may be able to offer more compensation or other perks. Make sure there is a clear tie to the work that the team completes and the outcomes that it drives.

> "Data, and the capability to extract useful knowledge from data, should be regarded as a key strategic asset[13]."

Just as important is the internal culture for this team. Practitioners of advanced analytics are problem solvers. They get energized by the opportunity to work with new data and solve new problems. Slotting them into roles that require a high amount of routine work can be frustrating, resulting in employees becoming disengaged, and increasing costly attrition. There are several good resources about how to build an analytics team – O'Reilly Media provides one that is well recognized[14]. Complement this by contacting other organizations, both healthcare and non-healthcare to identify best practices. The technology and finance industries are possible places to start.

- Which employees are currently responsible for any BI or analytics? How will their roles be impacted and how can you prepare them?
- How will you set expectations within the organization for what this team is responsible for? Where and how do handoffs between the data science and engineering teams occur?
- What new skills will you need in the team in order to make the most of your new analytics capabilities? Will you train existing staff or hire in new talent?
- What is the career path within the organization? What professional growth opportunities can you provide both from within and outside of your organization?
- What management structure will you use? Will this team fall under IT (CIO), clinical (CMIO), or does it make sense to create a new branch with reporting up to senior leadership?
- What structure will you use for assigning, tracking, and evaluating team KPIs? Should this be monthly or quarterly, and who is on the leadership team that will assign these objectives?

Process

In a busy environment like a hospital, clear processes are essential. Patients move between departments, staff change shifts and all the while potentially life-altering decisions must be made and actioned at speed. There is no margin for error or confusion. Therefore, any changes that have the potential to impact process must be planned carefully.

1. **Data management and governance**

Until relatively recently, healthcare organizations typically had a collection of best-of-breed applications to support

Real-world results: Montefiore Health System

Montefiore Health System runs a number of healthcare facilities in the Bronx, New York. It has created a Semantic Data Lake solution that pulls together data from multiple sources and in many different formats so that it can be analyzed holistically. The data ranges from clinician data, demographic, environmental, behavioral and wellness research findings and population demographics to medical imaging files and doctors' notes. Montefiore has worked with Intel, Cloudera and Hadoop to apply predictive and prescriptive analytics algorithms to its data, and to use machine learning to optimize its insights over time.

In its first implementation of the solution, Montefiore used it to identify patients at risk of death or in need of intubation within the following 48 hours, which is the window of opportunity to complete a successful intervention. It achieved accurate prediction at a rate of more than 70 percent , helping physicians deliver treatment earlier to prevent fatal episodes or respiratory failure.[15]

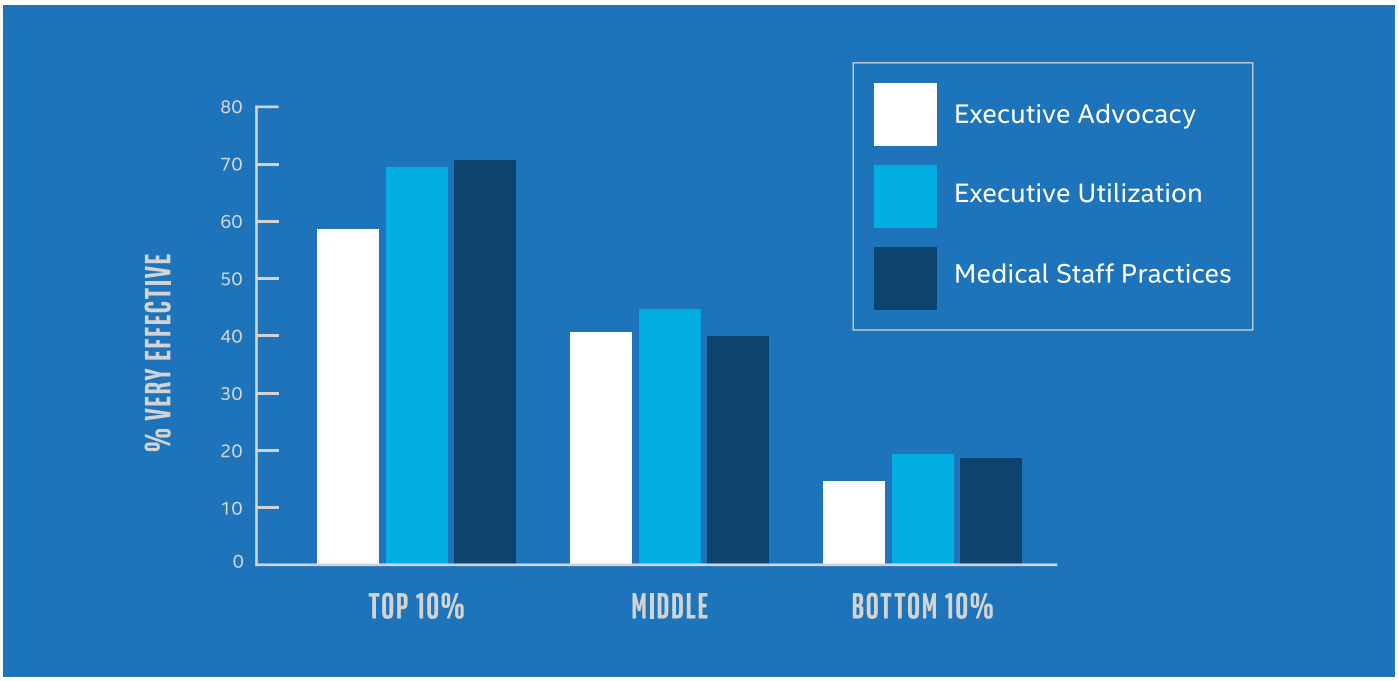


Figure 3: Levels of executive engagement are much higher in the top 10 percent of analytics in healthcare initiatives

different use cases and workflows. Now, most have larger EMR systems that pull much of this into one place. While the vendors of these systems tend to take responsibility for data governance and management within their own parameters, the evolution of analytics capabilities means healthcare providers now also want to bring in external data types which may not be so stringently managed. Properly collating and managing these disparate systems can be a challenge in itself. Add to that the fact that sentiment-based data such as social media posts – while extremely valuable – is much more ambiguous and open to interpretation than an EMR database, and the challenge becomes more complex. It's important that you work with your IT department to set rules around how certain data should (and should not) be interpreted by an analytics system.

Data collection and delivery must also be carefully controlled, with defined rules and processes to ensure that as your data grows, it uses standardized characteristics to maintain consistency and quality. If analyzing your data management capabilities for the first time, it is recommended to use a recognized framework like the Data Management Maturity (DMM) Model from the CMMI Institute[16]. Control your data fidelity and quality by putting in the time up front to define decision models and data sets upon which the system can learn and build.

Finally, ensure your methodologies for conducting analytics and applying insights are clear and accessible to your user base. By providing standard approaches, you can empower users, minimize risk and maximize the value you glean from your analytics investments.

2. Organizational alignment

Building a lasting advanced analytics capability requires tight alignment to clinical and business stakeholders. Depending on the scope of the advanced analytics program, significant investment may also need to be made in infrastructure to collect, store and process large data sets. Ensuring that this investment is available will require demonstrable ROI.

Commitment to invest must come from the top, and a study conducted by the International Institute for Analytics[18] shows that the top 10 percent of performers in healthcare analytics have a significantly higher level of executive support and engagement than other organizations (see figure 3). However, expectation of a positive ROI is not enough. Early selection of analytics projects is critical to building a solid foundation of support. Analytics teams without proper guidance run the risk of selection projects that are academically stimulating, but not well positioned to drive business value. One of the best ways to ensure that a project is on the right track is framing it within a defined business need. This need should have a clear definition for success, ideally quantified by some financial measure, and a clear path to implementation.

> "It's necessary to have a multidisciplinary team, with clinical, analytics, data science, information technology, and behavior change skill sets available from start to finish[17]."

Projects scope and strategy must also be considered. As Prashant Natarajan states in his book, Demystifying Big Data and Machine Learning For Healthcare, "many use cases, while meeting the requirement of ROI and strategic impact, have such large scope and complexity that success may be put at risk through over-reach". Meeting these requirements requires a significant amount of up front work. However the long-term payoff, being able to show measurable benefit to the organization, is well worth it.

Finally, ensure proper cross-functional representation within the analytics program. It is critical to engage subject matter experts (SMEs) to provide context to the business problem being addressed. SMEs can help address changes in data like

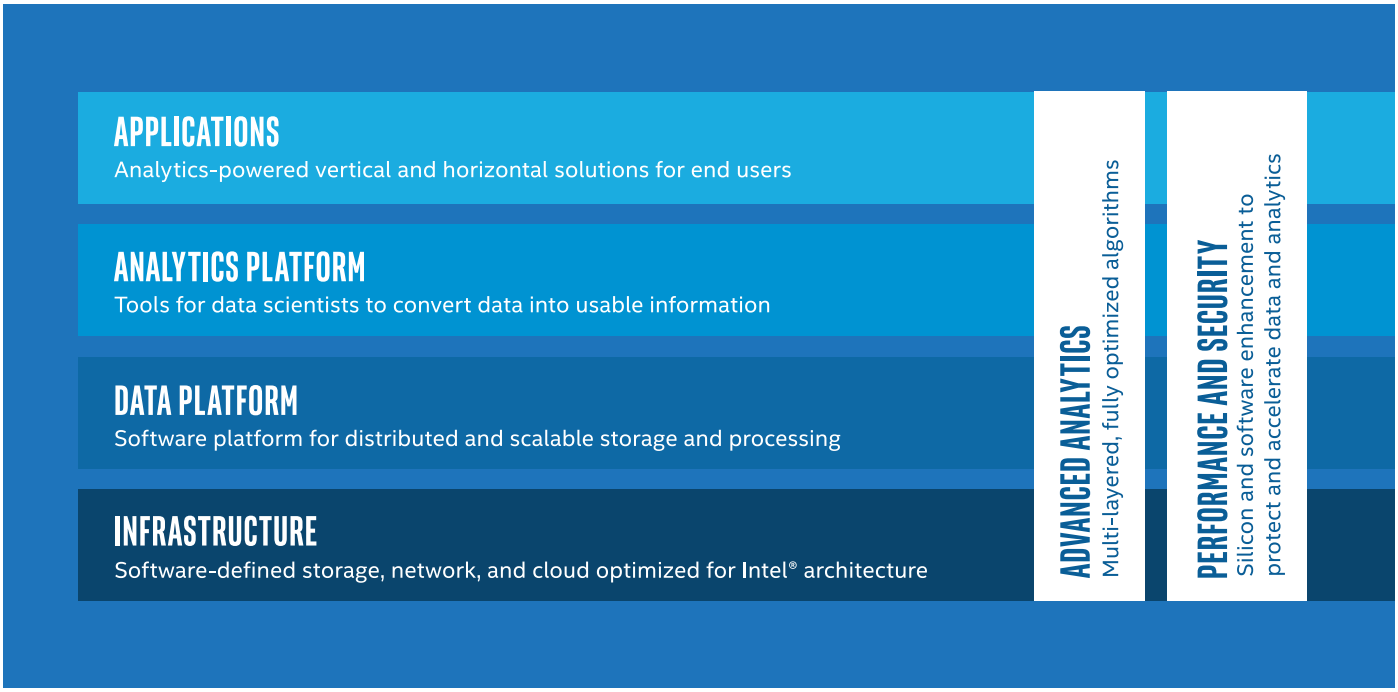


Figure 4: The analytics technology stack

when a data collection process evolves, or external factors that may impact the way that model results are used.

Some questions to consider here are:

- Who are the executives and key leaders that need to champion the predictive analytics program? What business results will speak to them?
- What are our organization's stated business goals and strategy over the short and medium term? How do proposed uses cases align to those goals?
- Who are the cross-functional stakeholders that will be impacted by the project? Who are the SMEs that should be engaged to provide context and understanding?
- When and how are successes shared? What are the key internal meetings that you can embed team members in to ensure alignment with the business?
- What will the data science group be known for? What are standard deliverables that will define the team throughout the organization?

3. Change management

Change management is one of the greatest challenges that many organizations face as they begin to implement advanced analytics. Moving from retrospective analysis to predictive analysis, requires workflows to be adaptable, sometimes in real-time, based on new data. Clearly this presents technical considerations, but just as, if not more important, are the considerations for cultural adoption of this new approach.

Organizations must have a strategy for how to implement change across the business. As your data yields more insights, how will you adapt existing clinical workflows to make the most of them? Predictive analytics projects are constantly evolving. Many follow the CRISP-DM lifecycle process which lays out six phases from Business Understanding to Deployment[19]. This continuous optimization is key to extracting those game-changing insights that will get your patients feeling better sooner or let your busy nurses finish their shifts on time. However, it does mean that your data management and clinical teams may need to change the way they work and collaborate, incorporating more time for feedback and iteration to ensure long-term success.

- Which aspects of organizational culture will be affected and how will you ensure your innovations work with, rather than against, it? James Merlino, MD's book Service Fanatics describes the significant effort that Cleveland Clinic undertook to change its internal culture to be more patient-centric. While a different goal, it shows how much work can go into influencing a hospital culture.
- Who will lead change management efforts, an internal team or external consultant? How might this change over time?
- Who are the formal and informal leaders that can bring support for change initiatives?
- How will insights be shared with hospital managers, clinical staff and consultants? What would be the most useful format, channel and cadence for them?
- How will you demonstrate the value of any changes to those they impact?

Technology

When building an analytics environment, picture it as forming four complementary layers (see figure 4), each including a variety of technologies, depending on your own organizational needs, legacy systems and preferences, and the type of analytics you need.

The infrastructure layer: This is the foundation that will enable you to acquire, store and protect your data, and to run commercial and open source analytics. A typical infrastructure layer will support a number of different

frameworks, databases and applications.

The data layer: Pre-analytics, this layer may have been comprised mostly of internal, siloed databases. As today's mix of data is more complex, distributed and unstructured, this layer may take the form of a data lake (for example based on the Cloudera Enterprise Data Hub*) that pulls together multiple data streams – from EMR records to clinician's handwritten notes to fitness tracker feeds – into one holistic environment. Many organizations will have more than one data platform, such as a relational database and big data platform based on technology like Hadoop* or NoSQL* side-by-side, to serve different analytics and reporting requirements.

The analytics layer: This is the layer that will be subject to the most change. There are a wide range of open source and commercial solutions available, with new solutions emerging regularly. Digital transformation is still new for many organizations, so there are many opportunities to innovate around data analysis. The analytics layer should support many different types of algorithms. In particular for predictive analytics, regression, classification, and clustering algorithms are three of the more common techniques used. Most analytics platforms will support these algorithm types along with many others to meet different needs.

Most organizations will have several technologies and/or solutions built into the analytics layer to support increasingly complex data analytics pipelines. Due to the fact that much healthcare data is unstructured, data analytics tools like a natural language processing (NLP) solution from Linguamatics* or Deep Learning for image detection can be incorporated to analyze specific data sources. These point solutions will generally feed back into the core platform to allow for developing models across all of the relevant data.

The applications layer: This layer supports the visualization of predictive model results. Open source or commercial software provides tailored analytics capabilities that support your specific use cases and workflows. Give careful consideration to how you will deliver the results of predictive models back to clinical stakeholders. It is highly recommended to engage a cross-functional team when deciding the best way to visualize and deliver results. Some questions to ask here are:

- What data should be delivered in a pull vs. push method?
- Should results be integrated into an existing application like the EMR or a new application?
- What data needs to be displayed to a health worker? How will you calibrate results to maximize signal vs. noise?
- How will interaction with the model results, either positive or negative, be captured to improve the model over time?

These layers enable you to take advantage of your existing data management systems as well as newly implemented technologies. Additional technologies like AI and performance or security solutions can be applied across the whole stack to accelerate insights and strengthen data protection. Solution providers with expertise in delivering analytics solutions for healthcare organizations, like ProKarma and NTT Data, can help you define the right mix of technologies for you.

Real-world results: Sharp HealthCare

Based in San Diego, California, Sharp HealthCare runs a specialized rapid response team for medical emergencies that occur within its hospital. It has developed a predictive model intended to detect patient health decline and identify patients at risk for requiring the team's help. It worked with Intel, Cloudera and ProKarma to build the solution, which analyzes a wide range of data – including blood pressure, temperature and pulse rate – from the hospital's EMR system, and trains the algorithm over time using machine learning.

In a proof of concept (PoC), Sharp tested the model against historical data and found that the model was 80 percent accurate in predicting the likelihood of a rapid response team event within the next hour, demonstrating the potential to drive real-time clinical interventions, improve outcomes and enhance resource utilization.[20]

Continue your journey

'Analytics' is not a once-and-done initiative – rather than a single project, it's an ongoing, evolving way of working that should adapt with your organization as needs and priorities change. Start with a simple project that will deliver clear value to the business in a short time frame, and use this initial success to build support from business users and longer-term commitment from senior leadership.

Once you have your foundations in place, you can add more analytics capabilities as your technology resources evolve and your data sources become richer. As your clinicians and other user groups become more familiar with the potential of analytics, they may also start to propose new use cases themselves.

In this paper we have focused on making the move from BI to predictive analytics, but there is plenty more potential for development once you've done that – for example you can look ahead to more automated possibilities that come with the next stages in the analytics journey, such as AI, deep learning and cognitive computing. Some organizations are already exploring such use cases.

The key to all this is data. The data that sits in your organization today. Make sure your data and your analytics strategy are aligned to meet tomorrow's challenges by considering these questions:

1. What are your data sources today and (how) are they connected?

2. What's your strategy for collecting and managing your data today and tomorrow?

3. What information are you not capturing that you should be?

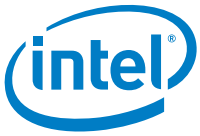
4. What will you need historical data for in two years' time and are you collecting it now (eg unstructured content like handwritten doctors' notes, images etc)?

5. What was the last question you were unable to answer? What would have enabled you to answer it?

Read on to learn more about how Intel technology can help you create your transformational analytics environment:

Report: Driving Clinical and Operational Performance Through Analytics

For more information about Intel's healthcare solutions, and to review the most up to date content visit: intel.com/healthcare



1 Professor Vasant Dhar, Stern School of Business, New York University, *Predictive Analytics: The power to predict who will click, buy, lie, or die,* Eric Siegel
2 All statistics in this paragraph are from: *Healthcare Disrupted: Next Generation Business Models and Strategies,* Jeff Felton and Anne O'Riordan
3 Vindell Washington, MD, National Coordinator for HIT in the US Department of Health and Human Services , Healthcare Executive Magazine, volume 31, Number 5, September/October 2016
4 https://www.cdc.gov/nchs/fastats/electronic-medical-records.htm
5 https://dashboard.healthit.gov/quickstats/quickstats.php
6 *Healthcare Executive Magazine,* volume 31, Number 5, September/October 2016
7 *From Triple to Quadruple Aim: Care of the Patient Requires Care of the Provider,* Thomas Bodenheimer, MD and Christine Sinsky, MD, http://www.annfammed.org/content/12/6/573.full
8 https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/index.html?redirect=/ehrincentiveprograms
9 https://www.healthit.gov/policy-researchers-implementers/health-it-legislation
10 France: http://blog.ihs.com/frances-2016-social-security-financing-act-a-figure-is-worth-a-thousand-words
UK: https://www.eiuperspectives.economist.com/sites/default/files/ValuebasedhealthcareUK.pdf
Sweden: http://ki.se/en/lime/value-based-health-care
11 https://www.intel.co.uk/content/www/uk/en/big-data/penn-medicine-trusted-analytics-platform.html
12 *Driving Clinical and Operational Performance Through Analytics,* International Institute of Analytics, August 2017
13 *Data Science For Business: What you need to know about Data Mining and Data-Analytic Thinking,* Foster Provost & Tom Fawcett
14 *Building Data Science Teams*, O'Reilly Media: http://www.oreilly.com/data/free/building-data-science-teams.csp
15 https://www.intel.co.uk/content/www/uk/en/healthcare-it/solutions/documents/montefiore-advance-patient-care-solution-brief.html
16 http://cmmiinstitute.com/data-management-maturity
17 *Getting Buy-In for Predictive Analytics in Health Care,* Meetali Kakad, MD; Ronen Rozenblum; David Westfall Bates, MD, https://hbr.org/2017/06/getting-buy-in-for-predictive-analytics-in-health-care
18 *Driving Clinical and Operational Performance Through Analytics,* International Institute for Analytics, August 2017
19 Machine Learning for Predictive Data Analytics, John Kelleher, Brian Mac Namee, Aoife D' Arcy
20 https://www.intel.com/content/www/us/en/healthcare-it/solutions/documents/using-machine-learning-and-emr-data-to-predict-patient-decline-case-study.html