The Age of the Algorithm is Over

More accurate forecasts of risk and cost for health plans, payers, providers, ACOs and Medical Home programs

Healthcare Business Intelligence Summit
September 20, 2012

East Bay Labs

Rising Health Care Costs

A large market that needs Business Intelligence

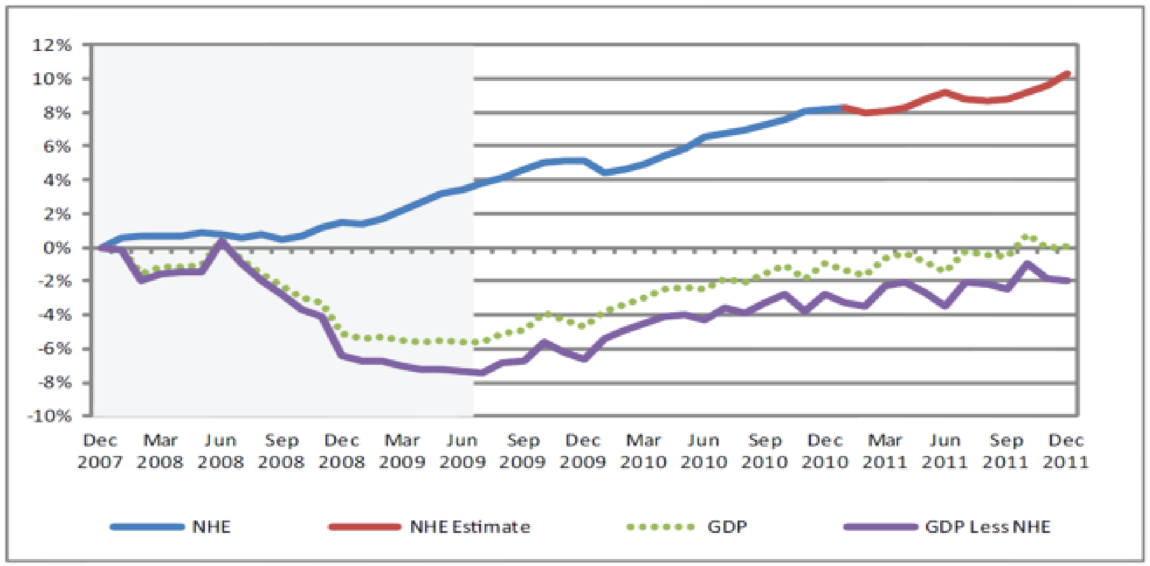
National Health Expenditures (NHE) and Gross Domestic Product (GDP)*

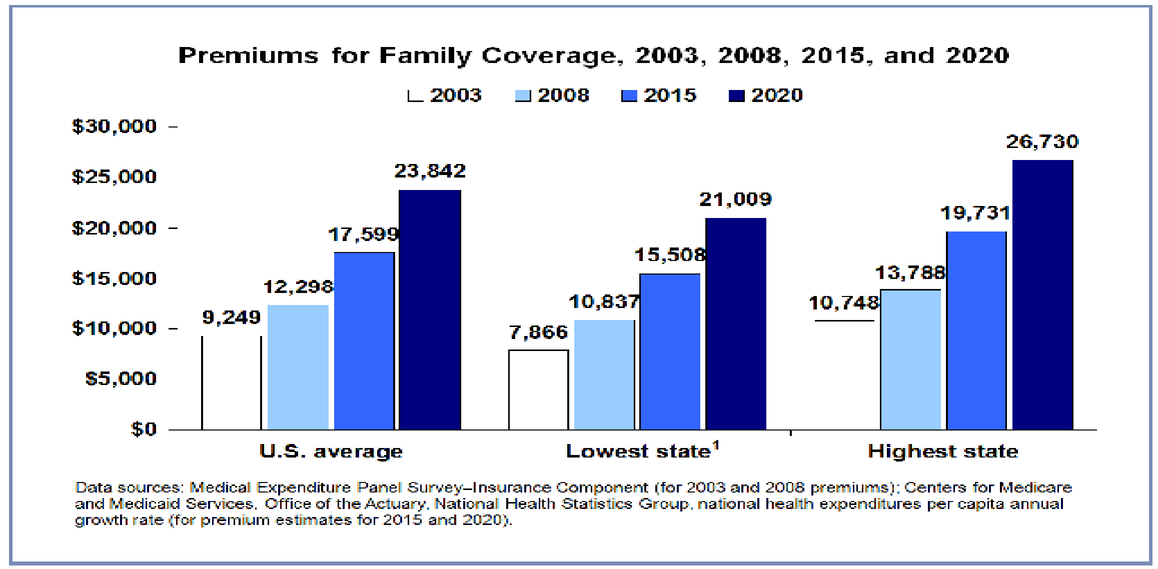
	Dec. 2009	Dec. 2010	Nov. 2011	Dec. 2011
GDP	14.04	14.82	15.27	15.24
NHE	2.54	2.66	2.75	2.76
NHE Share of GDP	18.1%	17.9%	18.0%	18.1%
NHE Share of PGDP	16.6%	16.8%	16.7%	16.8%
Growth from Prior 12 Months				
NHE	4.4%	4.6%	3.6%	3.8%
GDP	1.2%	5.6%	4.1%	2.8%
NHE minus GDP	3.2%	-1.0%	-0.5%	1.0%
NHE minus PGDP	1.8%	1.2%	-0.6%	0.0%

*Note: Spending in trillions of dollars, seasonally adjusted annual rate

Health care spending continues to rise, consuming a growing slice of the total U.S. Gross Domestic Product (GDP). As shown in these charts, the Great Recession of 2008 had virtually no impact on health care costs, which continued to rise while the economy and GDP fell sharply.

This increasing burden on the U.S. economy is not sustainable. We need to analyze data to discover ways to improve efficiency and the quality of care.





Rising Health Care Costs

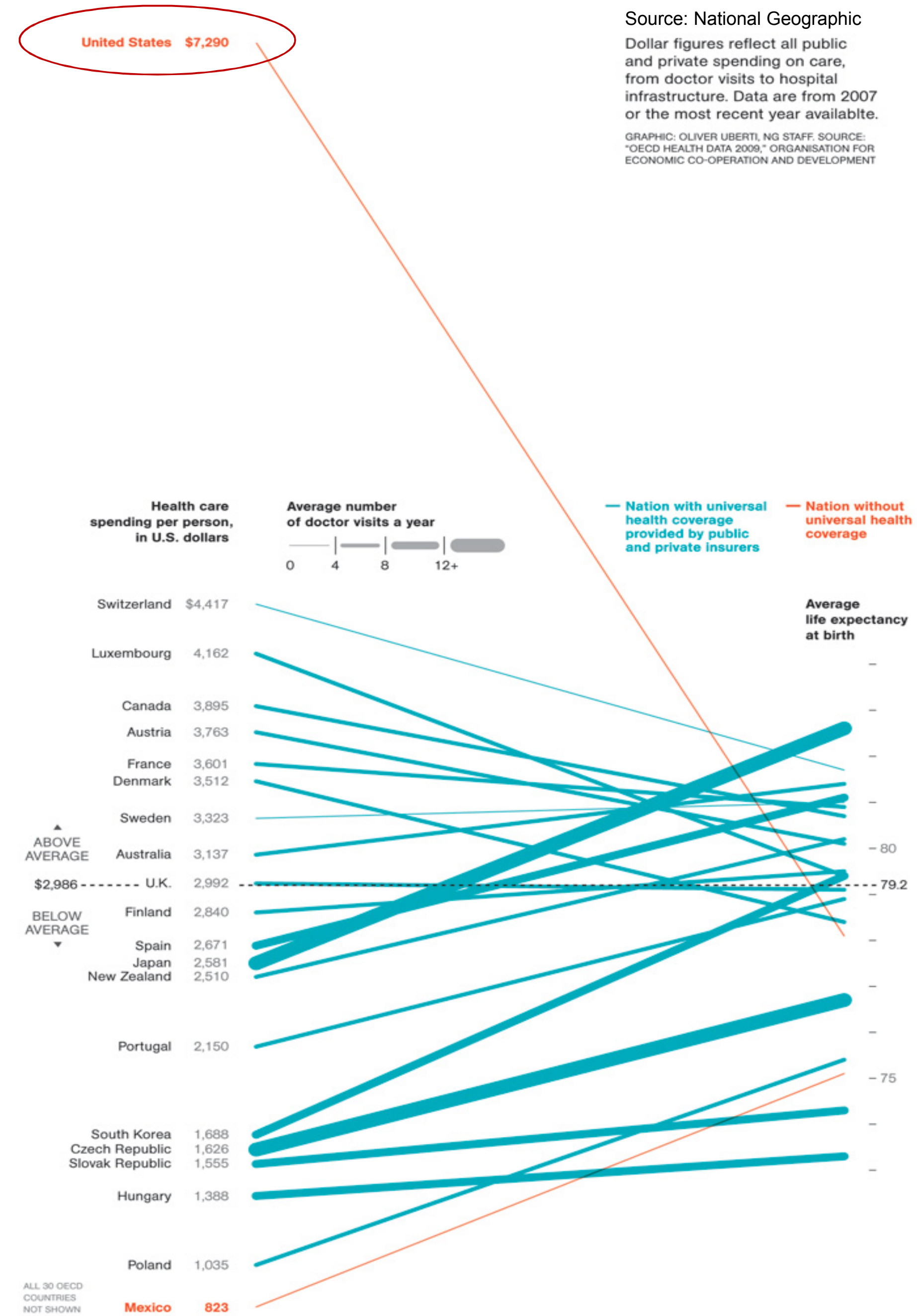
Spending too much for too little

The U.S. health care industry and policy makers know that we are spending far more on health care than every nation on earth, in return for less care and unimpressive life expectancy.

The National Geographic data analysis team created a now-famous graph in 2011 that compares health care costs per capita, the average number of doctor visits per year and life expectancy, for 20 industrialized nations. The U.S. is completely off the scale on cost, with a relatively low level of care and mediocre life expectancy.

Clearly, there are more efficient and effective ways to deliver health care services. With strong resistance in the U.S. to the universal coverage / payer model used by some nations, we need to improve efficiency in other ways.

Care and disease management is a major focus, to identify patients with the greatest risk and target care delivery. This reduces costs and improves outcomes – and the process is driven by data analysis.



Source: National Geographic

Rising Health Care Costs

Using data to identify opportunities for cost / care improvement

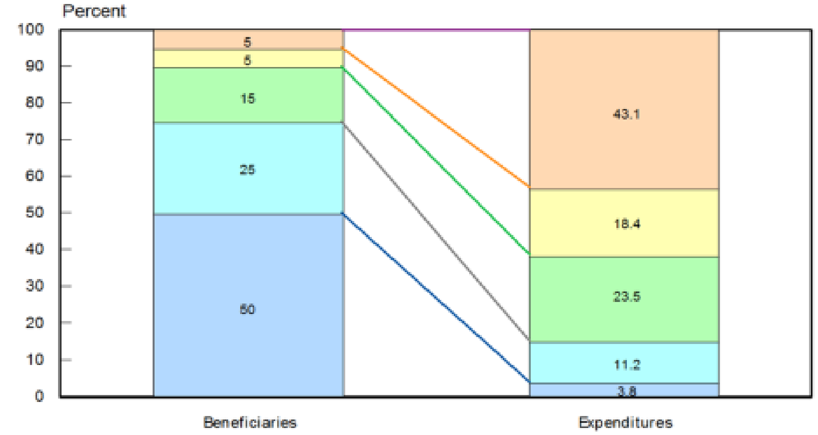
With rising costs and shrinking budgets, we need to identify causes and solutions: “Which members should we target for intervention?” and “How can we reduce readmissions?”

With Medicare, for example, 61.5% of total costs are driven by 10% of the population, and cost growth is fueled by ten chronic conditions as shown in the graphs to the right.

This highlights the key issue: payers and providers need to identify at-risk members and cost drivers, manage risk, target interventions and improve care. All of these changes require more focused and useful information – improved ‘intelligence’ for the healthcare process.

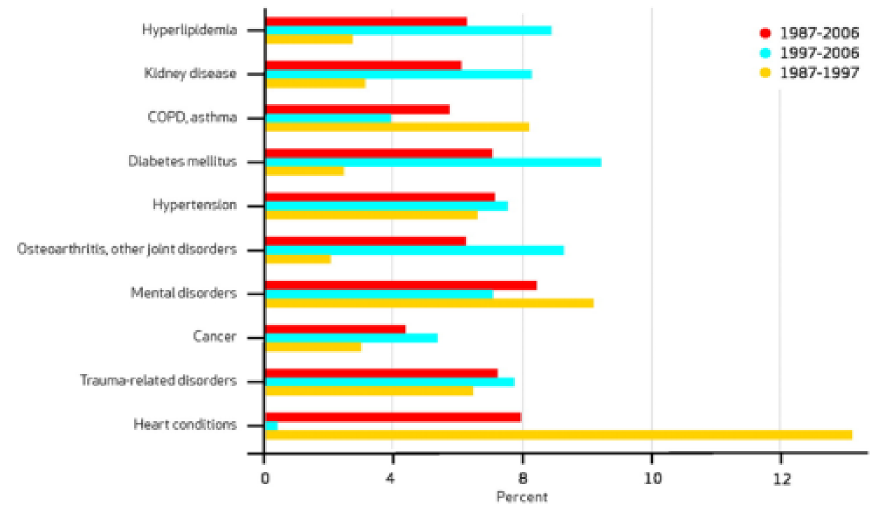
Fortunately, a revolution is sweeping the health care industry: rapid conversion to Electronic Medical Record (EMR) and Hospital Information Systems (HIS). This shift from paper to computers is creating a flood of new clinical data, which can be used to analyze results and improve outcomes.

The challenge: existing health care data warehouse systems and predictive modeling products do not meet this need.



Source: CBO based on data from CMS

Contribution Of Key Medical Conditions To Inflation-Adjusted Health Care Spending Growth Among Medicare Beneficiaries, 1987–1997, 1997–2006, And 1987–2006.



Thorpe K E et al. Health Aff 2010;29:718-724

Changes in Healthcare Data Sources

From paper files to EMR, HIS and HIE systems

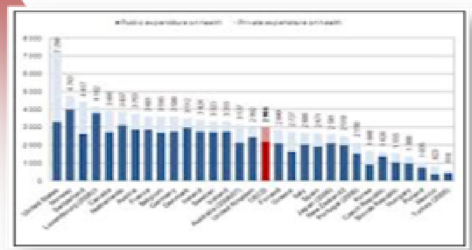
ARRA

Legislation drives change

← 1980 | 1980 - 2000 | 2000 - 2010 | 2010 →

Trapped in paper

Healthcare data was trapped in racks of file folders – a legacy that lingers today.



Shift to computers

Health care gradually moved to computerized records. First generation data warehouse systems were deployed.

Web delivery

Data warehouse systems become more advanced, with analysis and reports delivered via the Web to a wider group of users.

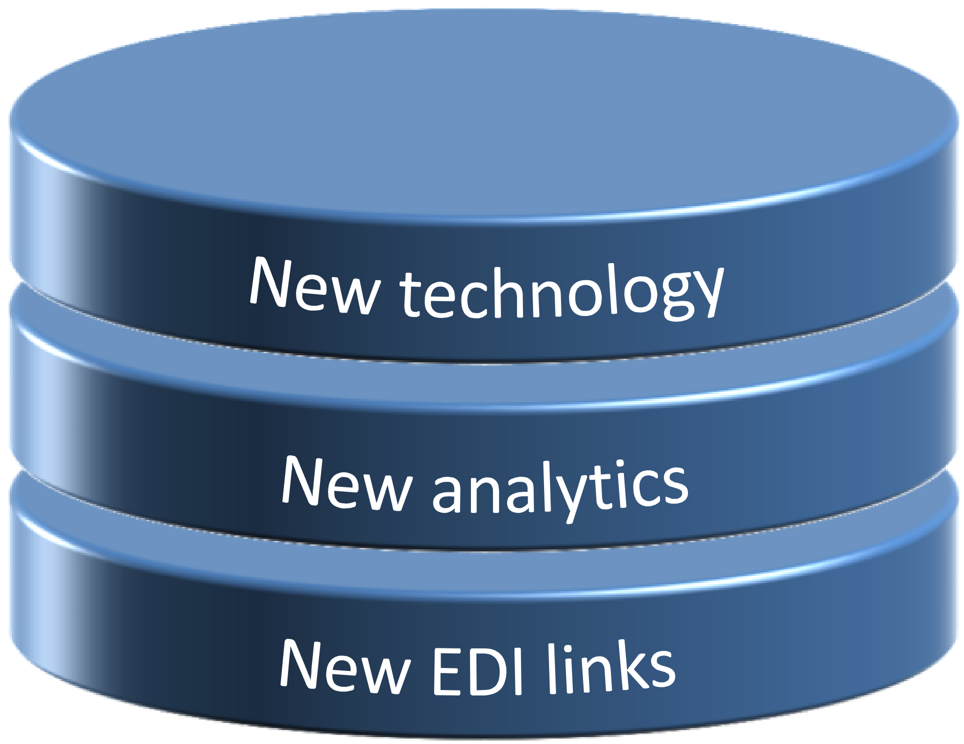
Flood of clinical data

EMR and HIS systems created a flood of new clinical data, forcing a radical change in data warehouse design, analytics and value.

East Bay Labs

Changes in Healthcare Data Sources

A different approach is required – to capture real-time data



Healthcare claims data

The primary source of health care data and reports for more than a decade, claims data is the foundation of current vendor products. This model is outdated.

Required: new data warehouse / BI systems

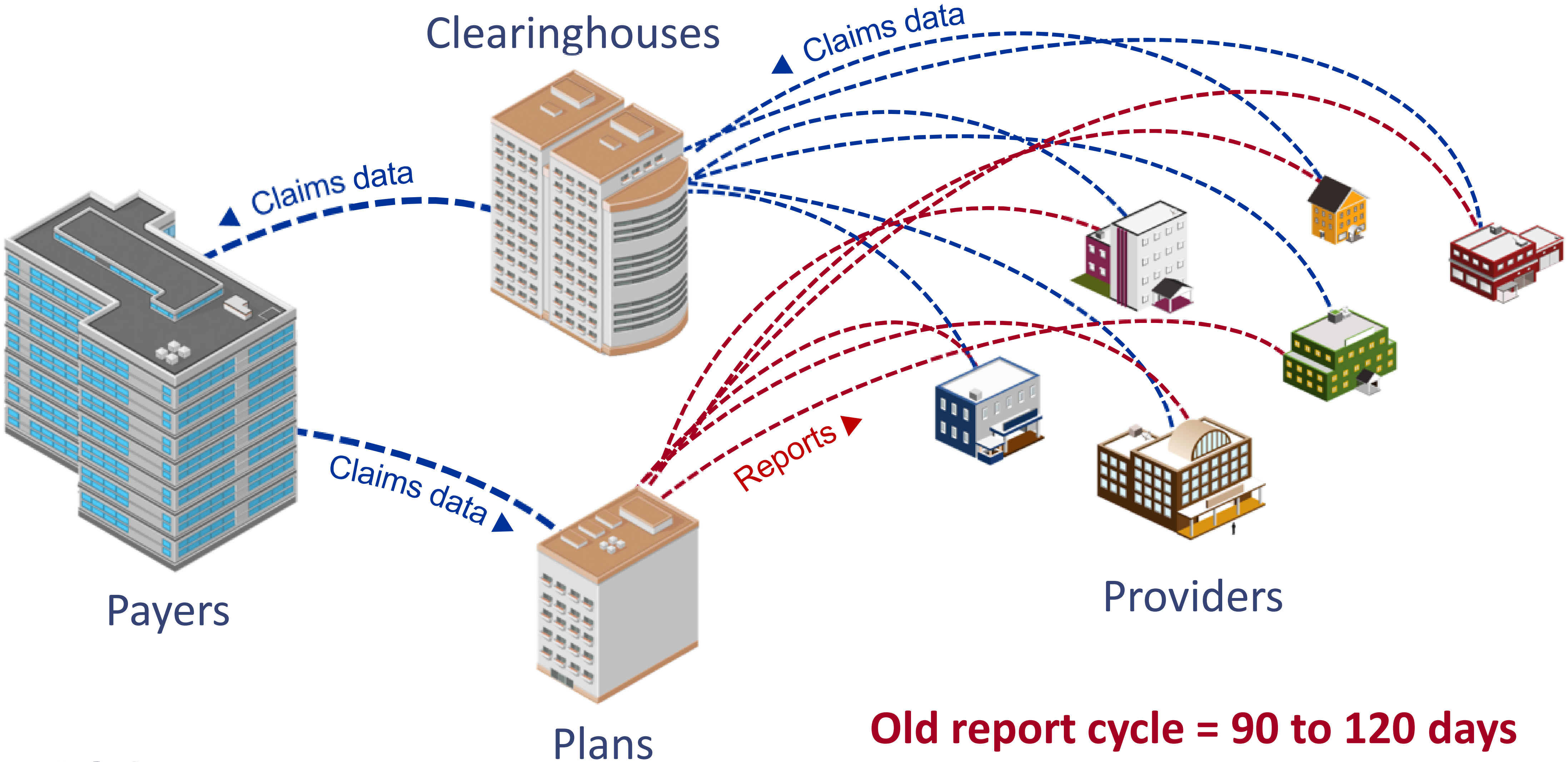
The flood of new clinical data requires a complete redesign of the classic 'healthcare data warehouse' model, data loading process and analytic reports.

New clinical data sources

Electronic Medical Record (EMR) and Hospital Information System (HIS) solutions are generating a rich set of 'Big' data that offers a clearer view of patient care and status – which drives advanced analysis and reporting.

Changes in the Healthcare Report Model

The old ‘collect and report‘ approach is outdated



Old report cycle = 90 to 120 days

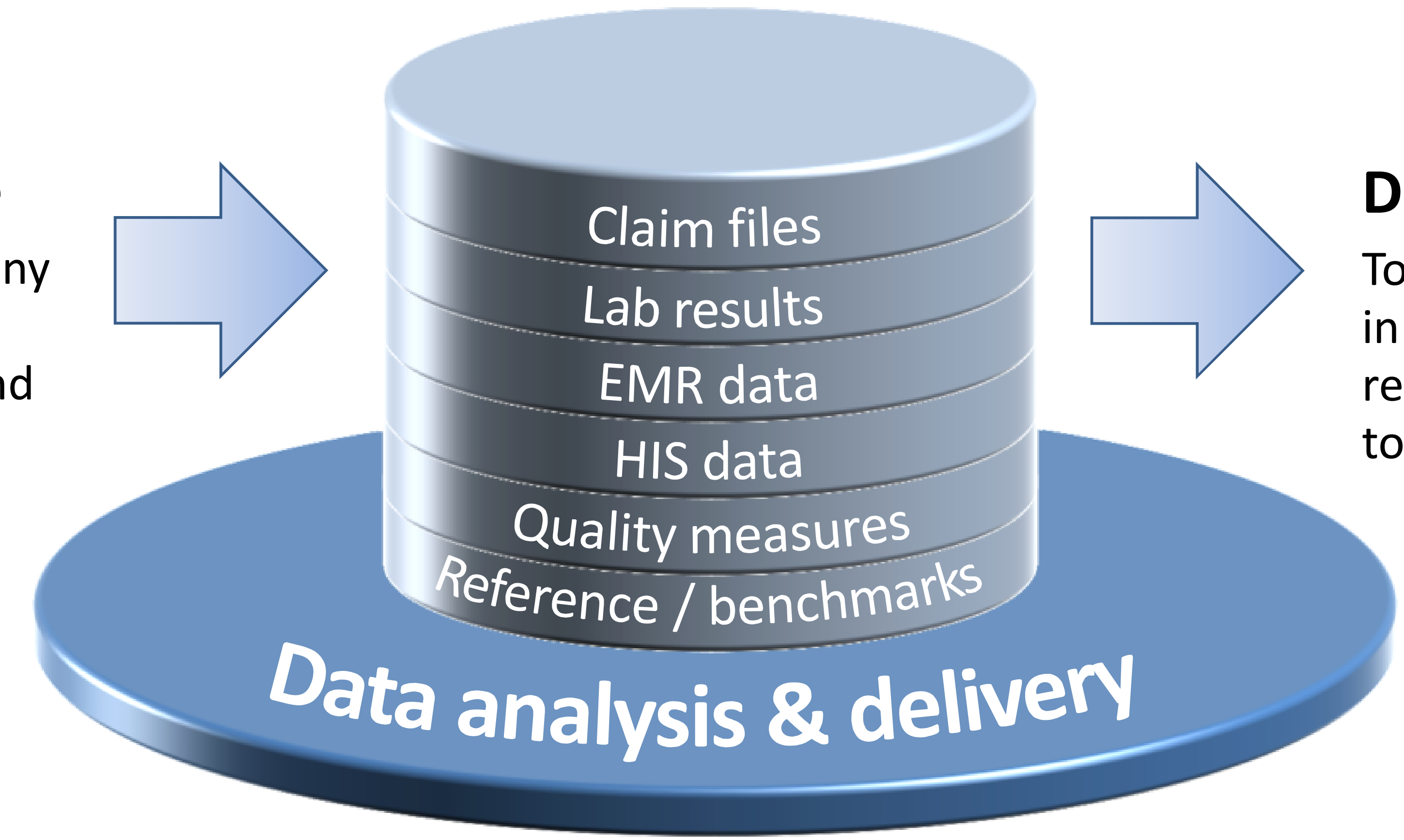
East Bay Labs

Changes in Healthcare Data Links

A paradigm shift – from ‘data warehouse‘ to active data analysis and exchange

Data Capture

Must support many data sources and formats, batch and real-time



Data Exchange

To a wide variety of systems in many formats, batch and real-time – e.g. alerts sent to a care management team.

Active vs Passive – With New Technology

The old data warehouse ‘collect and report’ model is outdated.

East Bay Labs

Changes in Healthcare Data Analysis

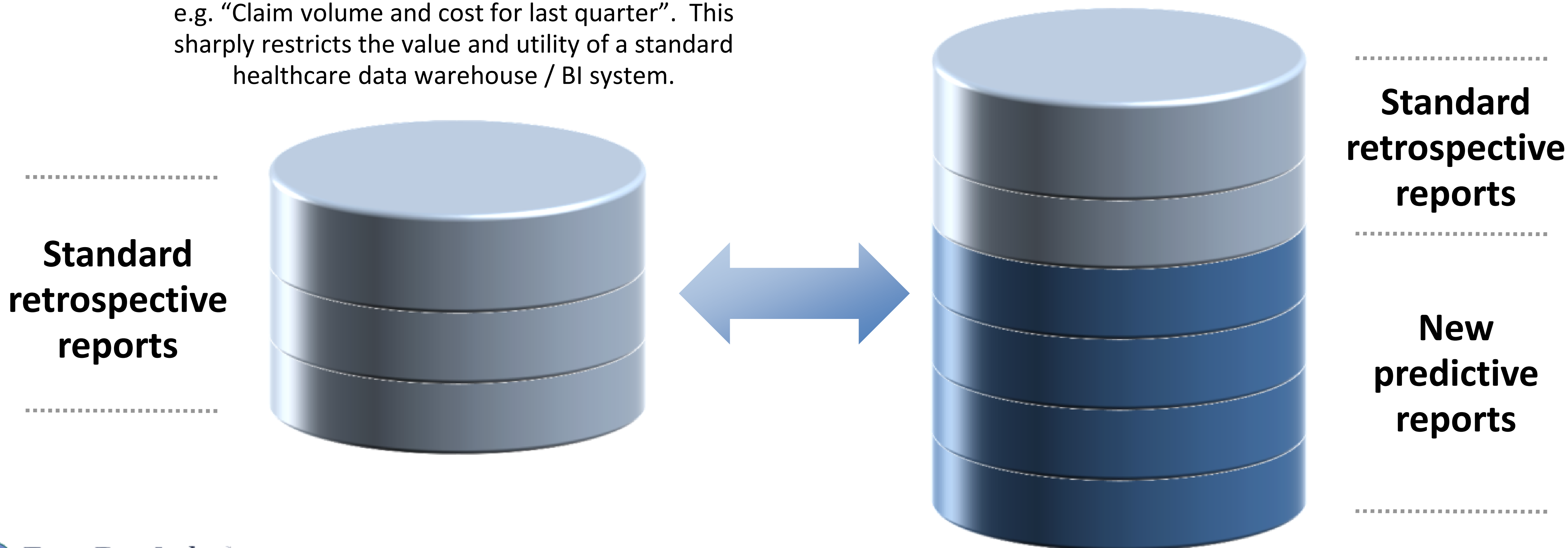
Predictive analysis vs retrospective reports

New 'Look Forward' Approach

Clinical data and new predictive analysis technology convert the 'data warehouse' into an active 'data engine' for improved care and management.

Old 'Look Back' Design

Standard data warehouse reports are retrospective, e.g. "Claim volume and cost for last quarter". This sharply restricts the value and utility of a standard healthcare data warehouse / BI system.



Changes in Healthcare Data Analysis

The ‘predictive algorithm‘ approach is outdated

Simple algorithms

ICD-9 Diagnosis Codes

↓

DxGroups (DxGs)
(n = 784)

↓

Condition Categories (CCs)
(n= 184)

+

$$
\begin{aligned}
\text{PCAL} = {} & \beta_{0agecat} + \beta_1 HCC_1 + \beta_2 HCC_2 + \beta_3 HCC_3 + \ldots + \beta_n HCC_n \\
& + (Age < 2) * (\beta_{i0} + \beta_{i1} HCC_1 + \ldots + \beta_{in} HCC_n) \\
& + (Age < 18) * (\beta_{k0} + \beta_{k1} HCC_1 + \ldots + \beta_{kn} HCC_n) \\
& + (Age > 65) * (\beta_{e0} + \beta_{e1} HCC_1 + \ldots + \beta_{en} HCC_n) \\
& + (\text{Selected interactions of HCCs})
\end{aligned}
$$

= Simple risk scores based on limited data

Changes in Healthcare Data Analysis

Improved predictive accuracy is needed

Results?

Society of Actuaries 2007 comparison of leading predictive modeling products.

(Higher r-squared score is better.)

With little change over five years, it is time for a new approach.

		R-Squared		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**
ACG	Diag	24.2%	23.0%	20.2%
CDPS	Diag	27.4%	24.6%	21.2%
Clinical Risk Groups	Diag	21.5%	20.5%	18.4%
DxCG DCG	Diag	29.7%	26.5%	22.9%
DxCG RxGroups	Rx	30.6%	27.1%	23.4%
Ingenix PRG	Rx	30.9%	27.4%	23.7%
MedicaidRx	Rx	29.7%	26.3%	22.7%
Impact Pro	Med+Rx+Use	29.3%	27.2%	24.0%
Ingenix ERG	Med+Rx	30.0%	26.5%	22.8%
ACG w/ Prior Cost	Diag+$Rx	27.7%	25.4%	22.1%
DxCG UW Model	Diag+$Total	33.1%	29.1%	25.2%
Service Vendor	**Inputs**	**100K**	**250K**	**None**
MEDai	All	35.7%	32.1%	27.6%

East Bay Labs

Changes in Healthcare Data Analysis

Summary

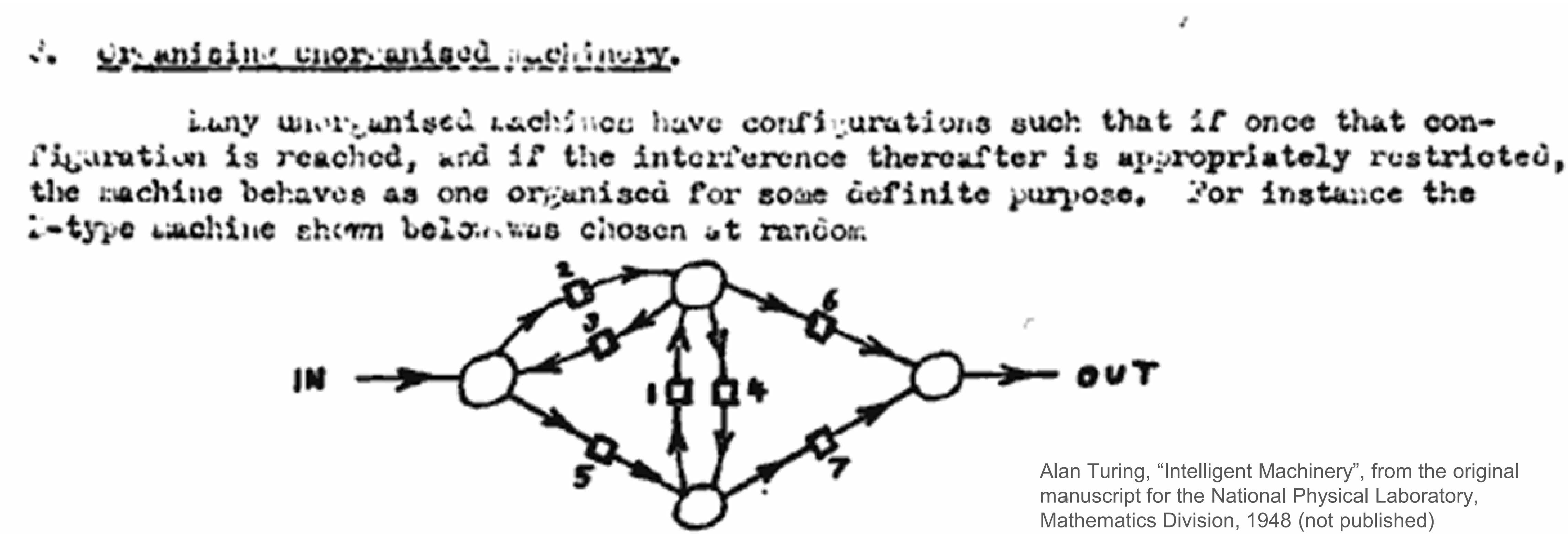
Where are we today?

~~Simple claim files~~	⇨	**Complex data** EMR, HIS, HIE, Medical Home, etc.
~~>90 day cycle~~	⇨	**Near real-time** HIE and Medical Home links via the internet
~~Simple algorithms~~	⇨	**Algorithms break** Cannot handle complex and chaotic data
~~Limited accuracy~~	⇨	**Need more accuracy** To support improved care and cost control

The Age of the Algorithm is over!

New Predictive Technology

What do other industries use with similarly complex data?



[illegible]. Organising unorganised machinery.

Many unorganised machines have configurations such that if once that configuration is reached, and if the interference thereafter is appropriately restricted, the machine behaves as one organised for some definite purpose. For instance the B-type machine shown below was chosen at random

Alan Turing, "Intelligent Machinery", from the original manuscript for the National Physical Laboratory, Mathematics Division, 1948 (not published)

… advanced artificial intelligence technology

East Bay Labs

New Predictive Technology

Industries that use artificial intelligence in their products and services:

- **Credit cards:** To identify fraud, suspicious transactions and risk
- **Games:** IBM's 'Deep Blue'™ and Watson™
- **Automobiles**: General Motors™, Chrysler™, Ford™, Aston Martin™, Isuzu™, Mercedes™, Saab™ etc. for engine, transmission, suspension, brake and audio control
- **Consumer products:** Washing machines, cameras, rice cookers
- **Robotics**: Navigation, motion control, sensors, actuators, vision
- **Finance:** Stock & commodities trading, forecasting, bankruptcy prediction
- **Defense:** Target recognition, systems control, battlefield simulation
- **Insurance:** Risk analysis, fraud detection, customer & market analysis

New Predictive Technology

A combination of AI technologies works best with healthcare data

1. **Neural networks**
   Artificial neural networks (ANNs) use an interconnected group of virtual 'neurons' to create complex models of relationships and patterns in data, with the ability to change and 'learn' based on feedback.

2. **Genetic algorithms**
   A genetic algorithm (GA) mimics natural processes with 'search' functions that include selection, mutation and inheritance – to identify the best solutions.

3. **Fuzzy logic**
   With standard 'crisp' logic, the temperature = 72 degrees. With fuzzy logic, the temperature may equal "comfortable" with a membership value of 7. This flexibility improves accuracy for data analysis.

4. **Rule system**
   The most familiar AI technology: a way to capture what we already know.

New Predictive Technology

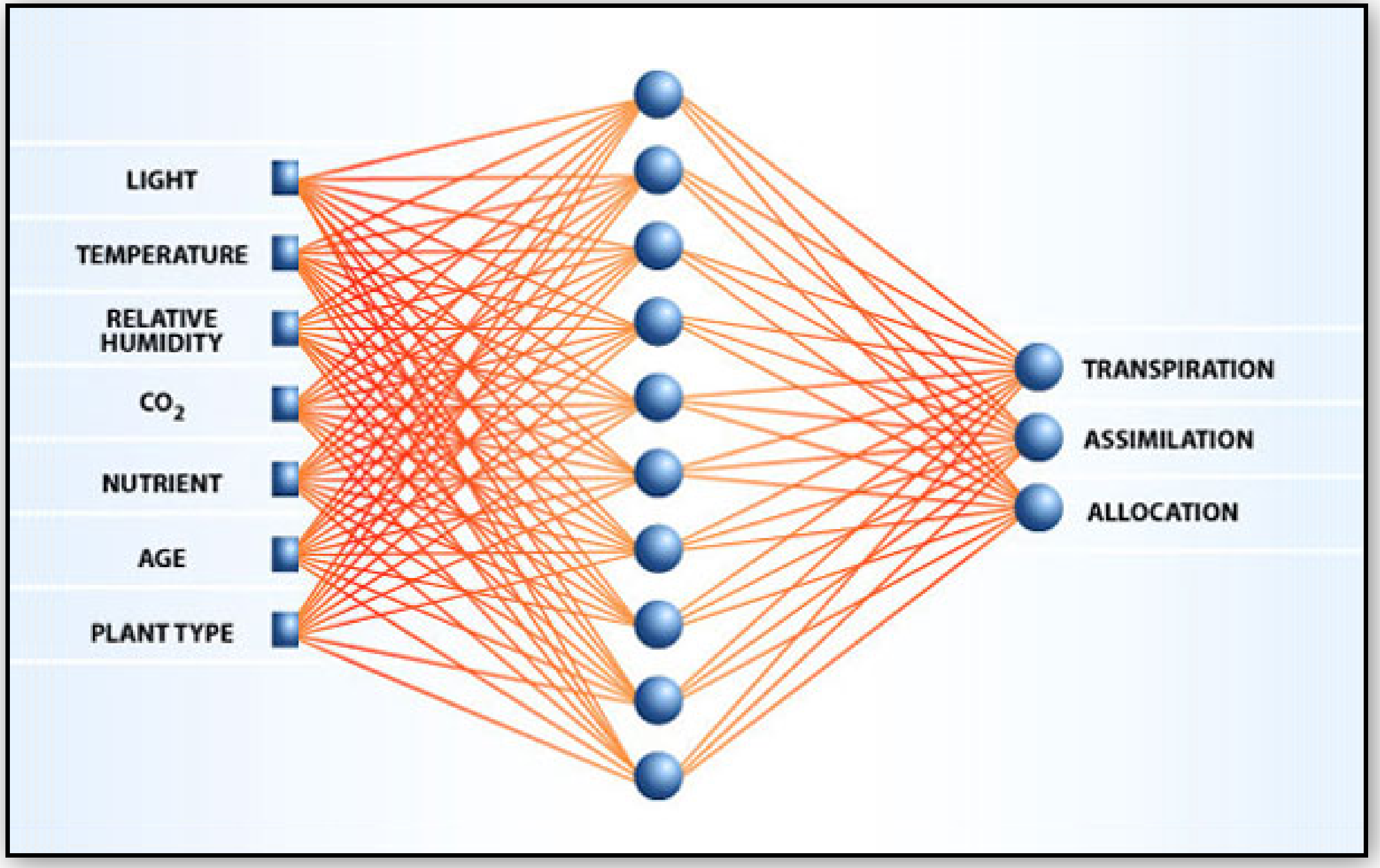
A combination of AI technologies works best with healthcare data

- **Neural networks**

  A software matrix of artificial 'neurons' captures inputs, weighs relationships and impacts, and delivers results.

  This is a diagram of a neural network used in NASA's Advanced Environmental Monitoring and Control program, to manage complex life support systems.

  Data inputs are shown on the left. 'Hidden' neurons weigh the data in the middle, with results delivered on the right.



Courtesy NASA. For more info see: http://aemc.jpl.nasa.gov/activities/bio_regen.cfm

New Predictive Technology

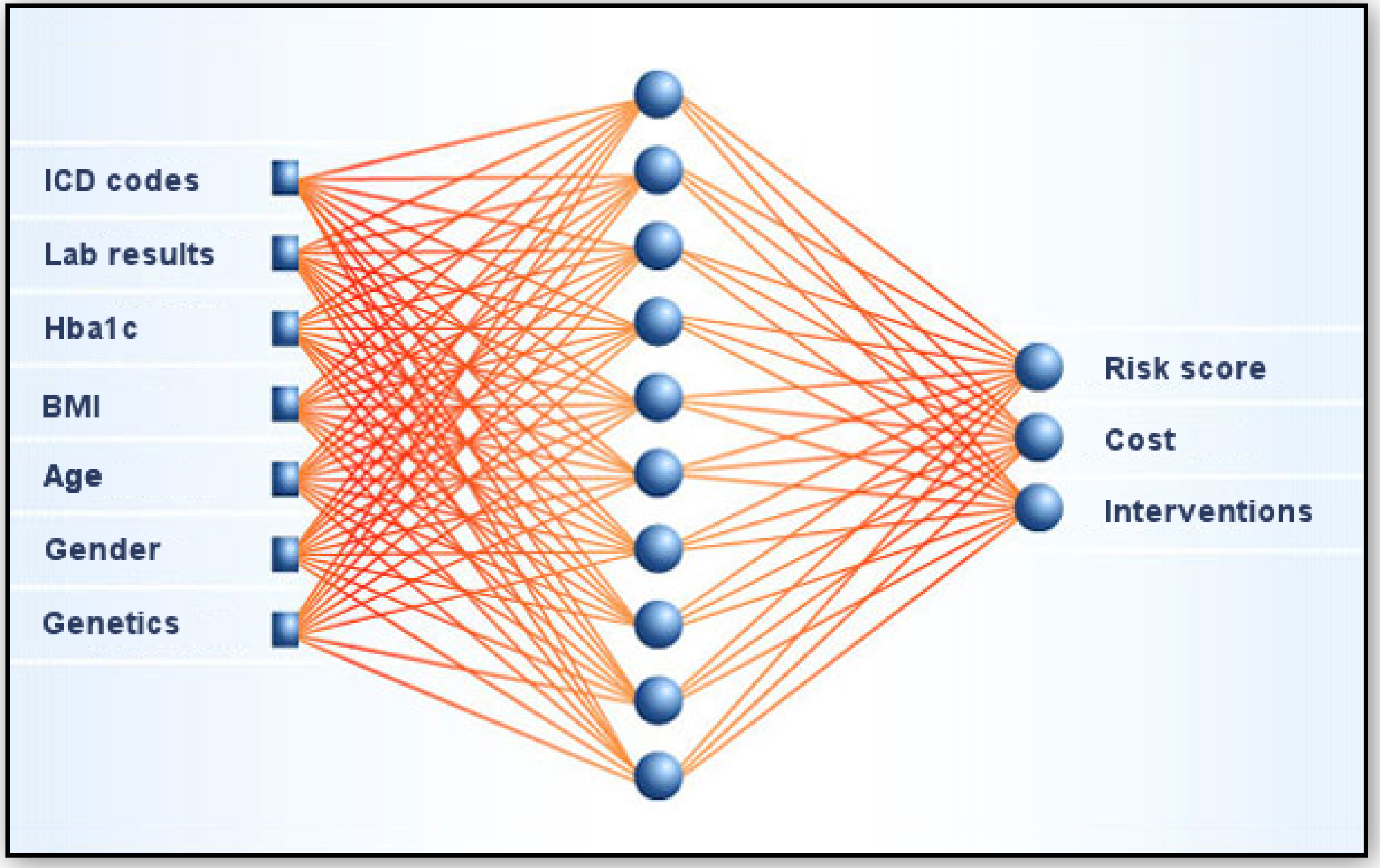
A combination of AI technologies works best with healthcare data

- **Neural networks**

  Here is a version of the same network, focused on diabetes risk and cost analysis.

  The data inputs and outputs are different, but the basic approach is the same.

  Nets like these automatically 'learn' and adjust as the data changes. This ability to handle complex data with hundreds of variables is ideal for healthcare data analysis -- which would swamp a standard predictive algorithm.



New Predictive Technology

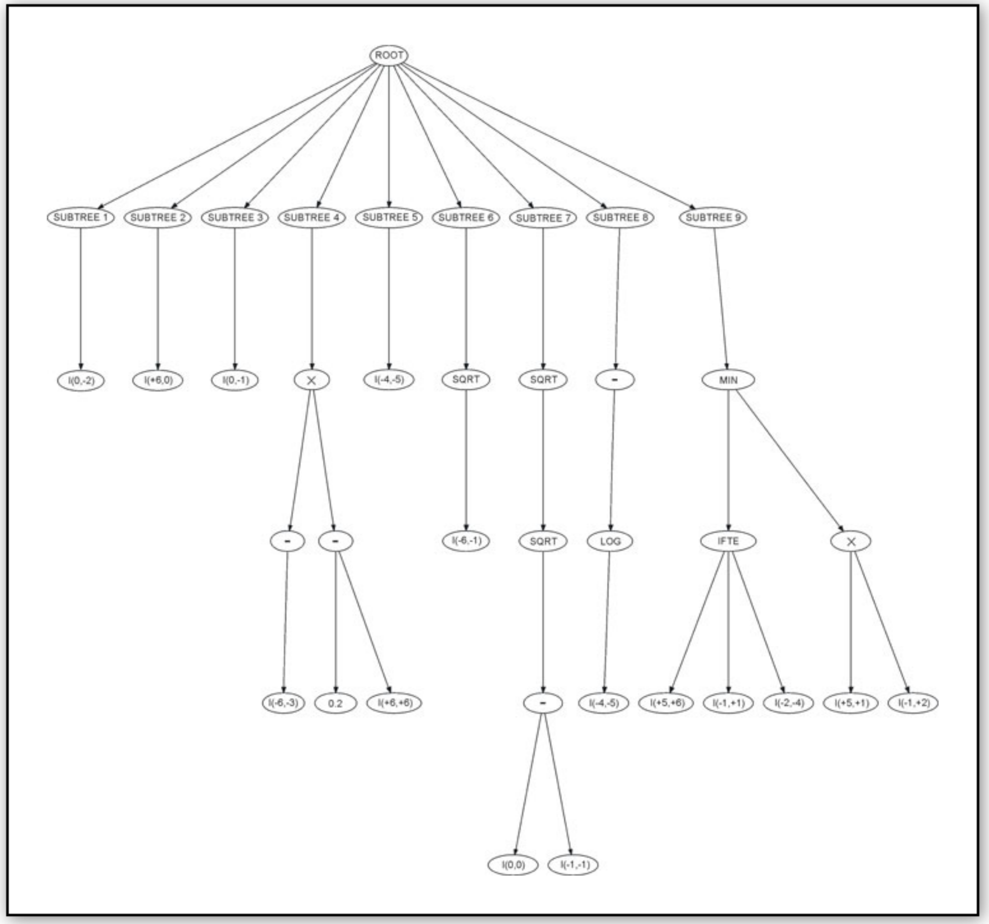
A combination of AI technologies works best with healthcare data

- **Genetic analysis** (GA)

  This approach mimics natural biologic processes with growing clusters of data-driven ‘bacteria’ modeled in software – mutating and reproducing to find the best solutions.

  A basic diagram of a genetic algorithm is shown to the right, used by the University of Sheffield for pattern classification and image analysis.

  GA adds power and flexibility but is computationally intensive. Fast new CPUs and solid state data storage make GA a useful tool, to enrich healthcare data analysis.



For more info see:
www.shef.ac.uk/eee/research/cr/research/feature_extraction

New Predictive Technology

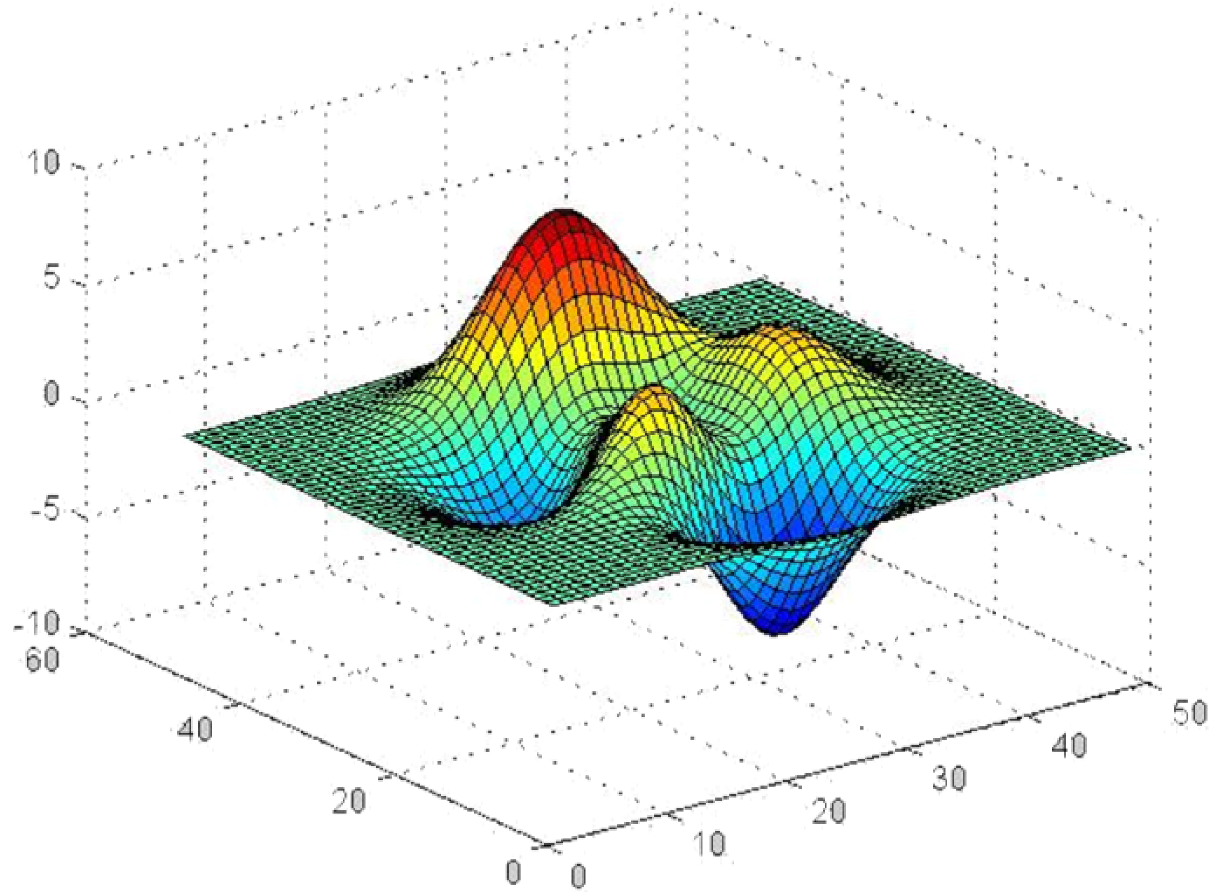
A combination of AI technologies works best with healthcare data

- **Genetic analysis** (GA)

  With the complexity of healthcare data, the genetic approach helps us find key factors that may be hidden in the 'noise' of hundreds of data points and variables.

  For example, imagine that the 3D surface plot on the right maps the risk factors for diabetes. A neural net could quickly identify the main 'peak' highlighted in red, but the small peak and hill in front could be overlooked.

  GA analysis is not efficient, but it is effective for these hidden factors.



Courtesy Mathworks   For more info see:
www.mathworks.com/products/global-optimization/description4.html

New Predictive Technology

A combination of AI technologies works best with healthcare data

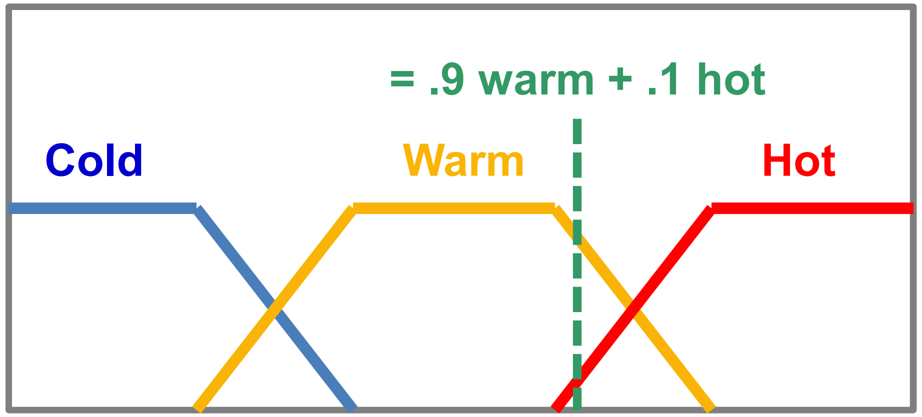
- **Fuzzy logic**

  This technology mimics the way human beings categorize and assess the natural world. A classic example is shown to the right. For a ‘crisp logic’ thermostat, the temperature is 73.5 degrees. A person sitting in the room would say that it was ‘warm’ or ‘very warm’.

  Fuzzy logic captures a broader range of information, adding flexibility to our analysis of healthcare data. This improves accuracy by including more useful data and expanding the links between data elements. Comorbidity factors are one example.

Crisp logic =

Fuzzy logic =



New Predictive Technology

A combination of AI technologies works best with healthcare data

- **Fuzzy logic**

  You probably benefitted from fuzzy technology many times, without knowing it. Auto manufacturers use fuzzy systems to make automatic transmissions shift smoothly and improve brake response in variable road conditions. Fuzzy logic removes the 'jitter' from Canon™ camcorders, makes Mitsubishi™ air conditioners more efficient, delivers perfectly cooked rice, and improves the performance of LG™ clothes driers.

  This versatile technology is also used worldwide in critical industrial control systems, avionics and mil / defense applications.

New Predictive Technology

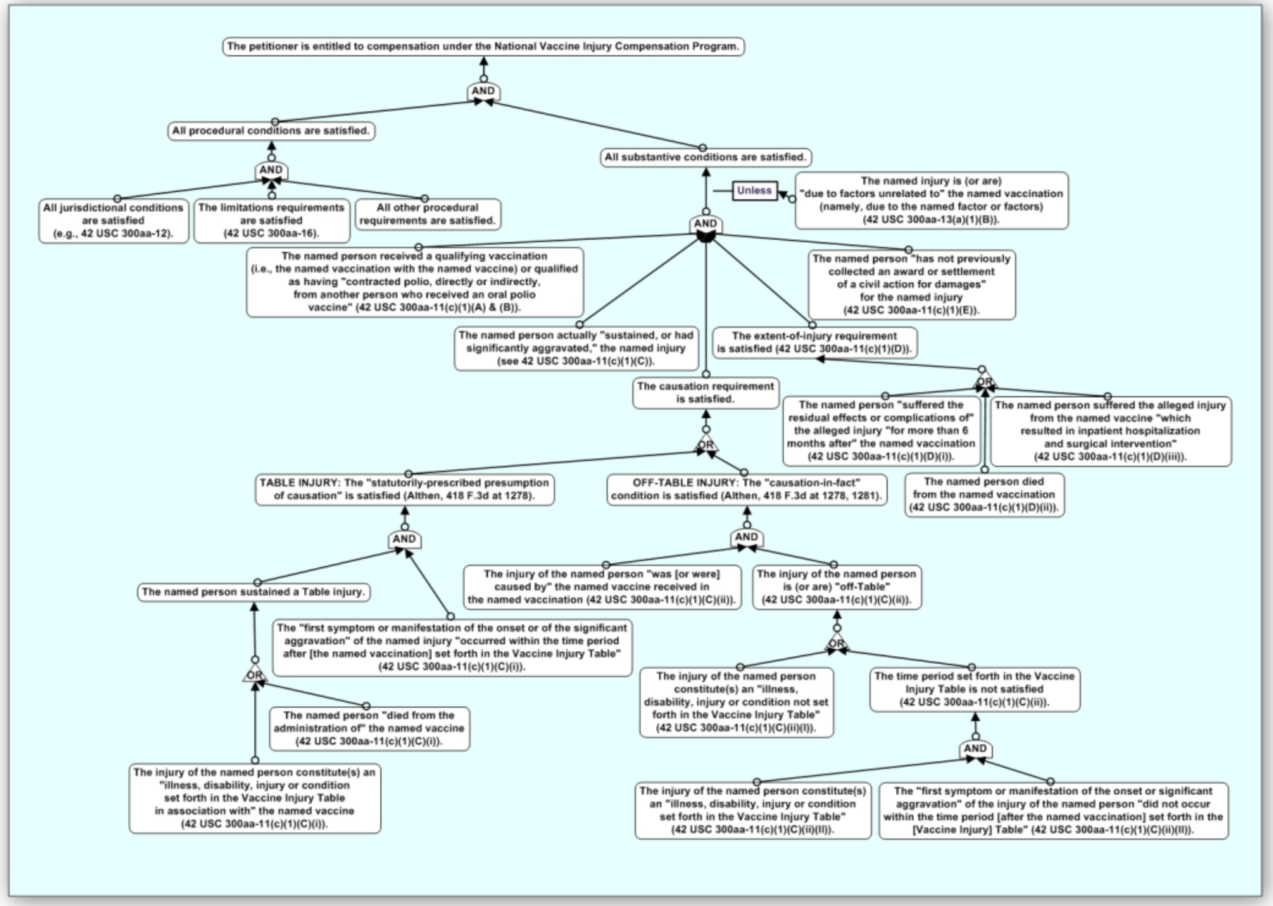
A combination of AI technologies works best with healthcare data

▶ Rule Based Analysis

This AI tool is easy to understand, and advanced technology is not required. A rule based system simply gives us a way to capture what we know about health care, and apply it in our analysis.

This diagram was created by the LLT Lab at Hofstra University, to map the rules embedded in regulations for the National Vaccine Injury Compensation Program.

Rules for health care data can include simple logic or complex formulas related to disease, lab results, etc.

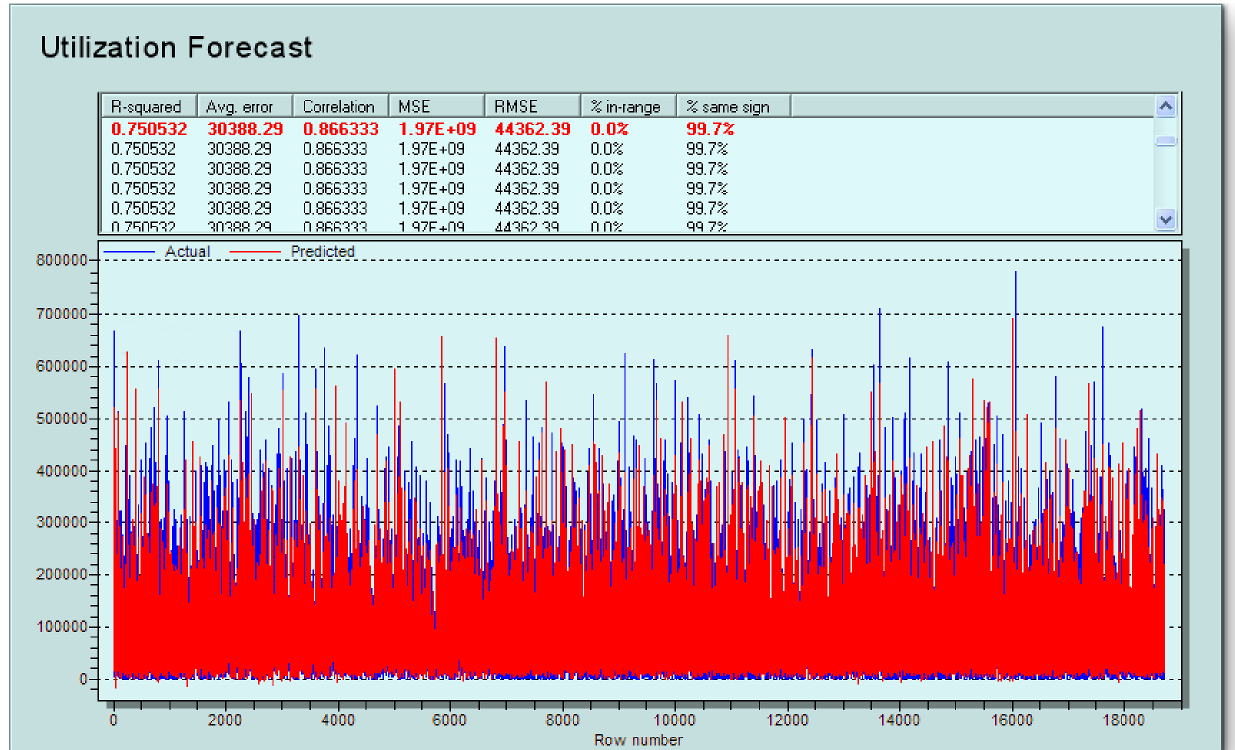


For more info see: www.lltlab.org

New Predictive Technology

Sample results: more accurate prediction with all of your data

- More accurate: > 2x compared to typical 'predictive modeling' products. This is helpful for actuarial analysis and critical for care / disease management.
- Complex data: AI technology can analyze complex data that would overwhelm an algorithm, including EMR, HIS and other data sources.
- Changing data: AI can shift to match new and changing data patterns.
- Flexible targeting: You can 'tune' AI systems for specific populations, e.g. "women with diabetes over the age of 40." This is critical for efficient care and disease management.



Screen cap of AI-based utilization analysis using Medicaid claims data. R2 scores range between .7 and .9 depending on the cost slice and target – significantly higher than standard PM systems.

New Predictive Technology

Sample results: more accurate prediction with all of your data

- **Predict claim costs for next year**

  This data set included 23K members with an average age = 72.94. Total costs for Year 2 were predicted based on the claims history in Year 1. Predicted costs were compared to the actual costs in Year 2, to create the r2 accuracy scores.

  r2 scores were sliced by cost. Note that accuracy is highest (>.91) for members with costs above $100K – the key group we need to identify for intervention and cost control.

Forecast vs actual cost in Year 2, based on Year 1 claims data

Total claim cost level	r2	Population
$1 to $1K	0.854	N = 1,857
$1K to $5K	0.827	N = 5,194
$5K to $20K	0.830	N = 7,565
$20K to $50K	0.754	N = 4,433
$50K to $100K	0.685	N = 2,556
$100K to $500K	0.912	N = 1,752
$500K+	0.997	N = 51

Note: An output file is available with all of the results noted above, based on anonymized data, in Excel format.

New Predictive Technology

Sample results: more accurate prediction with all of your data

- **Identify high-risk members who will have sharply *decreasing* costs**

  This more difficult example attempted to identify members with high costs who *appear* to be high risk, but will actually have sharply decreasing costs next year.

  4,837 members were identified with an average age of 76.15. Prediction was based on claims in Year 1, then compared to the actual costs in Year 2. r2 scores were sliced by cost range, as shown to the right.

Claim cost band	Population	r2 score
-$200K to -$100K	282	0.884
-$100K to -$50K	708	0.552
-$50K to -$10K	2867	0.457
-$10K to $0	977	0.001

The groups we were trying to identify are highlighted at the top – members who will have sharply decreasing costs. Accuracy is high, with an r2 of .552 to .884. Note that accuracy increases along with the projected decrease in costs, and drops to near zero where there is minimal change in cost. These 'focused' results are typical for an AI based system.

New Predictive Technology

Tools and resources

- **Wolfram Mathematica™ :**
  www.wolfram.com/products/applications/neuralnetworks/
  www.wolfram.com/products/applications/fuzzylogic/
- **MathWorks MATLAB™ :**
  www.mathworks.com/products/neural-network/index.html
  www.mathworks.com/products/fuzzy-logic/index.html
- **SAS™ :**
  www.sas.com/technologies/analytics/datamining/miner/neuralnet/index.html
- **SPSS™ :**
  www-01.ibm.com/software/analytics/spss/products/statistics/neural-networks/
- **An introduction to neural networks:**
  www.amazon.com/Neural-Networks-Applied-Sciences-Engineering/dp/084933375X
- **An introduction to fuzzy logic:**
  www.amazon.com/First-Course-Fuzzy-Logic-Third/dp/1584885262

New Predictive Technology

For more information

Please contact:

East Bay Labs™

Mark Hays

CEO

Phone: 508.661.9733

Email: MarkHays@EastBayLabs.org

East Bay Labs is a social benefit corporation dedicated to improving the quality and effectiveness of healthcare.