New Predictive Analysis Solutions for Health Care

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

Leveraging HIT and Innovation to Support New Care Delivery Models
Agenda

- Predictive modeling and healthcare
- Advanced predictive technology
- PIDA technology overview
- PIDA results and examples
Why is predictive analysis increasingly important?

ACO and Medical Home financial planning
Fraud & abuse
Provider contracts and risk management
Rising cost of care
Identify high risk patients for care management
Actuarial analysis and plan design

With rising costs and shifting risk, providers and payers need to know “What is coming next year?” vs “What happened last year?”
Who needs predictive analysis products / services?

- **Health plans** – to target intervention programs, model the cost of plan changes and predict future costs / risk
- **Insurers / actuaries** – to model sample member populations, design plan alternatives and assess related risk
- **Self-insured employers** – to identify at-risk members, identify trends and predict cost
- **ACOs and provider organizations** – to identify at-risk members, target intervention programs and predict future costs
- **DM / CM vendors** – to identify at-risk patients, target interventions and predict costs
- **Pharmaceutical companies, PBMs and brokers** – to identify utilization trends, costs and market opportunities

... in summary, any organization that carries risk or has a need to identify at-risk patients, utilization and cost trends
Where are we today with ‘predictive modeling’?

- Rigid: based on medical / pharma claims
- No support for real-time clinical data from EMR and HIS systems
- Limited accuracy blocks the ability to predict risk for individual patients, to improve care and efficiency

Do you have the tools you need?
Who needs predictive analysis products / services?

- **Health plans** – to target intervention programs, model the cost of plan changes and predict future costs / risk
- **Insurers / actuaries** – to model sample member populations, design plan alternatives and assess related risk
- **Self-insured employers** – to identify at-risk members, identify trends and predict cost
- **ACOs and provider organizations** – to identify at-risk members, target intervention programs and predict future costs
- **DM / CM vendors** – to identify at-risk patients, target interventions and predict costs
- **Pharmaceutical companies, PBMs and brokers** – to identify utilization trends, costs and market opportunities

... in summary, any organization that carries risk or has a need to identify at-risk patients, utilization and cost trends
Predictive analysis – the UHealthSolutions project:

- **Advanced technology was acquired in 2008**, based on predictive systems and approaches used in other industries – rather than software designed around standard healthcare claim files.

- **Our 2009 study** compared the Adjusted Clinical Groups Predictive Model (ACG-PM), Chronic Illness and Disability Payment System (CDPS) and Diagnostic Cost Groups (DCG) models, using Medicaid data from the State of Vermont.

- **Analysis continued through 2011**, including a new initiative to include real-time clinical data from EMR, HIS and HIE systems.

- **Our goal**: create a more advanced not-for-profit solution that can tackle the increasing complexity of health care data – with significant improvements in accuracy and lower costs.
Three basic options for predictive ID and forecasting:

- **Data matching and rules**
  Identify procedure codes, care dates and costs, assign groups and weights, and use basic statistics to calculate generic ‘risk scores’. This is the typical vendor approach, e.g. DxCG, Milliman and Ingenix.

- **Statistics**
  Analyze procedure codes, care dates and costs, assign groups and weights and use more complex statistical functions to predict trends and ‘risk scores’ – e.g. with regression analysis. In-house projects often take this path, typically with SAS or other tools.

- **AI: neural networks and fuzzy logic**
  Use claims, clinical data and calculated metrics to feed a neural network engine that identifies relationships and predicts future costs, events and trends. This is the most accurate solution; special technology and expertise are required, however.
We focused on a combination of AI technologies – including neural networks, genetic algorithms, fuzzy logic and rules:

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 is helpful for data analysis.

4. **Rule system**
   This is the most familiar AI technology: a way to capture what we already know.
Artificial Intelligence: an established technology

1943: Warren McCulloch and Walter Pitts wrote “A Logical Calculus of Ideas Immanent in Nervous Activity” for the Bulletin of Mathematical Biophysics

1948: Alan Turing wrote “Intelligent Machinery”, with a detailed description and diagram of a neural network (“B-type machine”):
Artificial Intelligence: an established technology

1951: Marvin Minsky (MIT) designs SNARC (Stochastic Neural Analog Reinforcement Calculator)

1954: Belmont Farley and Wesley Clark (MIT) design a computer simulation of a neural network with 128 neurons trained to identify simple patterns.

1958: Frank Rosenblatt and Charles Wightman (Cornell) build the Mark 1 Perceptron, which performed simple character recognition. John McCarthy (MIT) invents the LISP programming language.

1969: Marvin Minsky and Seyour Papert write “Perceptrons” (MIT Press), which highlights limitations in AI systems. This leads to a long-term decline in AI research and funding.


AI today: in products, services and industries worldwide

- **Credit cards:** Used by vendors to identify fraud, suspicious transactions and risk.
- **Games:** IBM’s ‘Deep Blue’ computer defeated Garry Kasparov in 1997. Now every leading computer game includes artificial intelligence.
- **Automobiles:** Aston Martin, Chrysler, Ford, Isuzu, Mercedes, Saab etc. for engine, transmission, suspension 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, identity mapping
- **Insurance:** Risk analysis, fraud detection, customer & market analysis
- **Healthcare:** Diagnostic tests, ECG, MRI, PET, X-ray, decision support
PIDA Technology Overview

AI functions and options included in PIDA:

**Neural networks:**

- **Structure and training:** PIDA includes a combination of two approaches:
  1. A method based on the Cascade Correlation algorithm. Neurons within the hidden layer are automatically added and trained, to reduce setup time and allow the system to determine the optimum size and structure for the new network.
  2. A “genetic” algorithm that improves the analysis of all of the variables contained within the client’s data. *(Note: The processing time required for genetic algorithms has become a minor factor with improvements in CPUs, RAM and storage technology.)*

Both methods are used, independently and combined, based on the data and target.

- **Classification and ID:** The neural networks within PIDA can also be used to identify and classify, e.g. to identify groups of diagnostic / procedure codes and patient factors linked to increased risk of readmission.
AI functions and options included in PIDA:

Genetic algorithms:

- **Genetic approach:** Many mathematic / statistical functions can be used to model well-behaved systems, but fail when the number of variables and random factors grow. ‘Genetic’ algorithms mimic biologic evolution with models that evolve and change, attempting to find an optimal solution. This is helpful with diverse and chaotic data – a typical challenge in healthcare.

- **Evolution functions:** PIDA includes the ability to evolve a population of potential solutions, with mutation, crossover (“mating”) and diversity (“creep”) functions. The ‘fitness value’ is measured to update the population in the next generation. Dozens of population models can be processed at one time on each server, to identify key factors.

- **Genetic training:** As noted on the previous page, PIDA can use genetic functions to create training values – that match virtually any target that is supported by the client’s data, for any patient segment. This gives PIDA more flexibility and power.
AI functions and options included in PIDA:

Fuzzy logic:

- **Fuzzy logic vs crisp logic:** Standard ‘crisp’ logic is used in most software applications, e.g. “Temperature = 72 degrees”. With fuzzy logic, the temperature could be “warm” with a membership value of 8. In healthcare, for example, ETG (Optum) and MEG (Medstat) groupers use crisp logic to count the exact number of days between events. With fuzzy logic, the number of days can flex to provide more meaningful analysis.

- **A hybrid approach:** The combination of fuzzy logic and neural networks in a hybrid model is now known as a “neuro-fuzzy system” or NFS. The flexibility of fuzzy logic is combined with the learning capability of a neural network. PIDA uses the NFS approach to improve data analysis and identification / classification.

- **Fuzzy clusters:** Data inputs and outputs from processes are combined in clusters that create a composite result for reporting and analysis.
Results: more accurate prediction to match your needs

- **Very accurate:** > 2x compared to leading ‘predictive modeling’ products.
  
  *Tested against DxCG, Ingenix, MedAI and ACG, using standard SOA models.*

- **Flexible targeting:**
  
  We can tune the PIDA engine for specific populations, e.g. “women with diabetes over the age of 40.” This is critical for efficient care / disease management.

  *Competing products are inflexible and produce generic risk scores.*

Screen cap of utilization analysis using Medicaid claims data. R2 scores range between .7 and .9 depending on the cost slice and target – significantly higher than every competing vendor, e.g. Verisk, Ingenix, Milliman and Elsevier.
Sample results:

1. **Predict claim costs for next year**

   This data set includes 23K Medicaid members with an average age = 72.94. We predicted costs for Year 2 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.

<table>
<thead>
<tr>
<th>r2 Scores by Actual Claim $ in Year 2:</th>
<th>r2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 to $1K</td>
<td>0.854</td>
</tr>
<tr>
<td>$1K to $5K</td>
<td>0.827</td>
</tr>
<tr>
<td>$5K to $20K</td>
<td>0.830</td>
</tr>
<tr>
<td>$20K to $50K</td>
<td>0.754</td>
</tr>
<tr>
<td>$50K to $100K</td>
<td>0.685</td>
</tr>
<tr>
<td>$100K to $500K</td>
<td>0.912</td>
</tr>
<tr>
<td>$500K+</td>
<td>0.997</td>
</tr>
<tr>
<td>Average</td>
<td>0.990</td>
</tr>
</tbody>
</table>

Note: An Excel file is available with this PowerPoint, with all of the results noted above, based on anonymized data.
PIDA Results and Examples

Sample results:

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

This counter-intuitive and more difficult example attempted to identify members with high costs and utilization who appear to be high risk, but will have sharply decreasing costs next year.

4,837 Medicaid 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.

Note that accuracy increases along with the projected decrease in costs, providing the highest level of accuracy for members we need to identify.

Accuracy is low where the cost shift is low – dropping to zero where there is minimal impact.

<table>
<thead>
<tr>
<th>r² for delta groups:</th>
<th>N</th>
<th>r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>-$200K to -$100K</td>
<td>282</td>
<td>0.884</td>
</tr>
<tr>
<td>-$100K to -$50K</td>
<td>708</td>
<td>0.552</td>
</tr>
<tr>
<td>-$50K to -$10K</td>
<td>2867</td>
<td>0.457</td>
</tr>
<tr>
<td>-$10K to $0</td>
<td>977</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Note: An Excel file is available with all of the results noted above, with anonymized data.
Recap: significant features and benefits

- **Accuracy:** PIDA is >2x more accurate than the leading ‘predictive modeling’ solutions. This is particularly important for cost forecasts and member-level risk analysis, to target care management efforts.

- **New data sources:** Traditional predictive modeling systems were designed based on healthcare claims. With the shift to EMR, HIS and HIE systems, clinical data is becoming more important. PIDA can support any data you can provide.

- **Available data:** PIDA does not require multiple years of claims history to achieve significant results. With high coverage churn in the market, this is a key advantage.

- **Improved focus:** You can focus on specific chronic conditions, groups, risks and financial goals. This improves accuracy and value compared to generic ‘risk scores’.

- **Lower cost:** PIDA does not require a fixed PM/PM or PM/PY fee, unlike standard predictive modeling products. As a not-for-profit organization, we work with clients to match their requirements.
For more information please contact:

Mark Hays

CTO, Commonwealth Medicine and UHealthSolutions

Phone: 508.421.5802

Email: Mark.Hays@UHealthSolutions.org

“Healthcare Explorer”, “PIDA” and matching logos are trademarks of UHealthSolutions. The Commonwealth Medicine logo, Umass / University of Massachusetts logo and name are trademarks of the University of Massachusetts. This document and all of the contents contained herein are copyright 2010 and 2011, UHealthSolutions, Inc.