Using Predictive Analytics to Assess Risk and Target and Refine Human Services

Kristin Porter, Richard Hendra & Rekha Balu

MDRC

Presentation at NAWRS
July 31, 2017

The challenge we aim to address:

- **Assess individual's risk of not reaching key milestones**
  - As accurately as possible, taking advantage of all available data
  - Rapidly and iteratively as new information becomes available
  - With an easy feedback loop to staff

A solution: predictive analytics

- To predict each individual’s probability of not reaching a milestone, at a particular point in time.
- To show variation in individuals’ risks.
- To rank individuals’ by their risk levels.

- By capitalizing on a very large numbers of measures.
- By combining subject-matter expertise, with data-driven algorithms.

The value for case management and continuous program improvement:

- **To target individuals** based on their level of risk
  - To increase success of an intervention
  - To save limited resources
  - To reduce burden on some participants

- **To refine human services**
  - Based on characteristics, behaviors etc. of individuals within levels of risk
  - Based on distributions of risk within and across groups and sites

Case study 1: New Visions for Public Schools

- Goal: predict students risk of not reaching key milestones
  - Graduating on time (at start of 12^{th} grade)
  - Passing end-of-year 9^{th} grade Algebra exam required for graduation (at end of 1^{st} semester)
- Data: school records data with frequent updates
  - Test scores, test attempts, course marks, DAILY attendance and tardiness info, behavior measures
  - Longitudinal – going back to 8^{th} grade
- Next steps in continuous improvement: interventions designed based on insights from PM results
  - Particularly targeting students for whom a small intervention could have high impact
  - A focus on schools with high concentrations of students with various levels of risk

Case study 2: Center for Employment Opportunities

- Goal: predict individual’s risk of key milestones at different points in program pathway
  - Job placement (soon after intake, after training)
  - Job retention (at start and midway through employment)
- Data: rich background and administrative records
  - On educational attainment, work history, conviction history, program participation and engagement
  - Goal is real-time data updates and rapid updates of predictions
- Next steps in continuous improvement: interventions designed based on insights from PM results
  - Including behavioral and light touch interventions
  - Implementation study
  - Rapid-cycle RCTs

Limitations to keep in mind:

- Not all outcomes are highly predictable and/or the right measures may not be available or high quality.
- Predictions of risk identify opportunities to intervene, not interventions themselves.
  - PM does not tell us what works for whom.
- Predictions are estimates and come with uncertainty.
- Biases in data can be perpetuated and can have ethical implications.

Are predictive analytics right for an organization?

Yes, if it:

- Wants to improve the accuracy of risk assessments.
- Can implement interventions meant to help individuals at risk.
- Believe that targeting services based on risk will increase the success of an intervention.
- Make the most of limited resources by directing services only to those at risk.

Is an organization ready for predictive analytics?

Yes, if it:

- Has consistent and frequently-updated data readily available.
- Is ready to communicate and act on results.
- Employs a staff and a data system that can sustain predictive work.

For more information:

- kristin.porter@mdrc.org
- richard.hendra@mdrc.org
- rekha.balu@mdrc.org

- www.mdrc.org

Extra slides

Predictive analytics: Questions that identify level of risk vs. questions that measure effectiveness

	Predictive RQ (who)	RCT RQ (what)
Students	"Which youth are most likely to drop out of the program?"	"What is the impact of changing the tutoring/housing/financial support youth receive on program completion?"
Staff	"Which staff/case workers are most likely to leave the profession?"	"Which support services or incentives are most effective for retaining at-risk staff?"
Programs/ Offices	"In which offices or geographic areas are youth more likely to drop out of the program?"	"Do office-wide incentives reduce dropout of youth?"

Darker colors indicate higher true risk

Off-track = Intervention recommended

On-track = No intervention recommended

Two-group approach shows less variation

Continuous measures show more variation

THE COMMON "INDICATORS" APPROACH	MDRC'S PREDICTIVE ANALYTICS APPROACH
Uses a few individual-level summary measures to assess risk. For example, many organizations use the so-called ABC indicators of attendance, behavior, and course performance to understand whether a student is on track to graduate.	Can use hundreds of individual-level measures. For example, it can incorporate daily attendance and tardiness trends, screening data, case notes, and more.
Produces limited risk categories. For example "at risk/not at risk" or "low/medium/high risk."	Produces a continuous measure of risk that reveals variation in risk levels within categories.
Cannot rank which individuals are at more or less risk, can only assign them to categories.	Can rank individuals' risk levels and group them in multiple ways.
Is accurate for the average individual.	Is more accurate for young people who do not have typical risk profiles.
Uses indicators and risk estimates that are static, updated maybe once a year.	Produces risk estimates that can be updated whenever new information is available.
Is typically used to describe the risk of not reaching only one or two milestones, because the research process used to establish risk categories is complicated and subjective.	Can be used to describe the risk of not reaching multiple, successive milestones in a young person's life, because the automated process can be repeated.