Paper 1329-2015

Causal Analytics: Testing, Targeting, and Tweaking to Improve Outcomes

Jason Pieratt; Edmund Berry, Ph.D

ABSTRACT

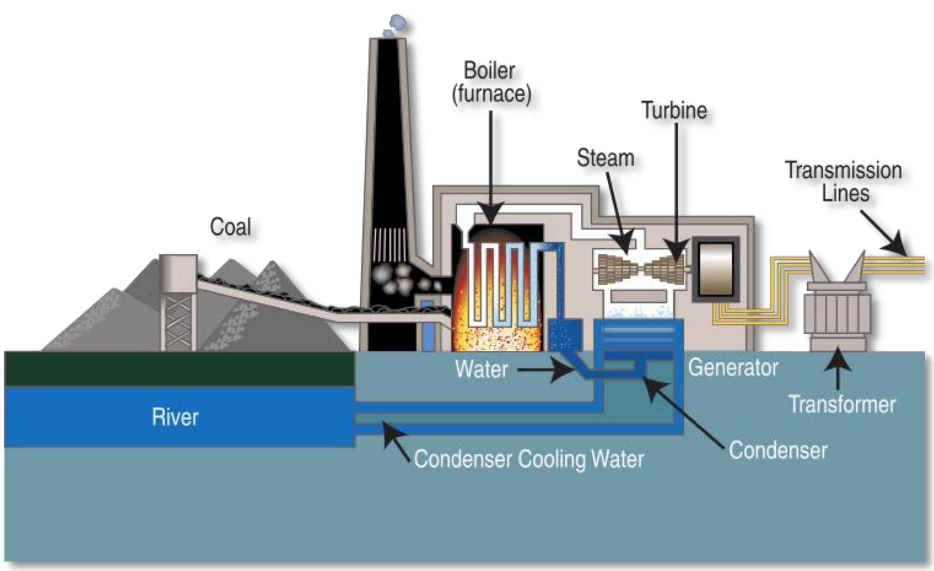
Causal Analytics is the business application of statistical methods that are used for modeling causal relationships between inputs, interventions, and outcomes. Three areas of Causal Analytics are introduced in this paper: "testing" whether outcomes change because of an intervention; "targeting" interventions to have the greatest improvement in outcomes; and "tweaking" an intervention to control system outcomes.

INTRODUCTION

Many business problems are *causation* problems. What prices should we set for each market and product to increase key customers and increase net revenue? Which interventions lead to better outcomes for our customers? Who are the highest priority customers to contact during a retention campaign? Does a new program have a positive return on investment? These problems of net pricing, gaining new customers, retention, customer success, and ROI can all be analyzed (better) using a Causal Analytics framework.

CAUSAL ANALYTICS FRAMEWORK

Causal Analytics is the application of methods that detect and describe causal relationships between *systems*, *interventions* in those systems, and the *outcomes* resulting from interventions. Figure 1 illustrates the pieces of a basic system using the example of a power plant. This system - with inputs, processes, interventions, and outcomes - can be applied to other contexts as well. At the heart of the framework is the idea that "any attempt to *intervene* in some system or process entails the possibility of *gain or loss*" (Causalytics, 2013).



A power plant is a useful way to model systems and interventions. The two main inputs are fuel/coal and water which are processed through the system and transformed into electricity and emissions.

Even seemingly small interventions in this system (e.g. how fine the coal is crushed up) can result in significant changes for the system outcomes.

Image source: http://en.wikipedia.org/wiki/Fossil-fuel_power_station

Figure 1. System Components Illustrated with a Basic Power Plant Diagram

Business problems (especially in health care) usually involve complex issues that make causal analysis difficult. Analytical challenges include internal validity, external validity, selection bias, information bias, known intervention bias, confounding, omitted variable bias, reverse causality, skewed data, rare occurrences, outliers, and more (Academy Health, 2002). It is outside the scope of this paper to discuss each one of these in detail. However, to illustrate the challenges, Figure 2 on the next page provides an example of how these challenges might affect a program evaluation of an intervention for babies born prematurely.

This figure illustrates how there are many other factors - *in addition to a NICU case intervention* - that might impact and explain any observed health outcomes. For example, families who choose to participate in a NICU intervention might *also* be more engaged with their personal health when compared to families who choose not to participate (a "self-selection bias" example). It could be that these underlying differences between the groups with "engagement in personal health" are the real cause of any observed health outcomes – not the intervention.

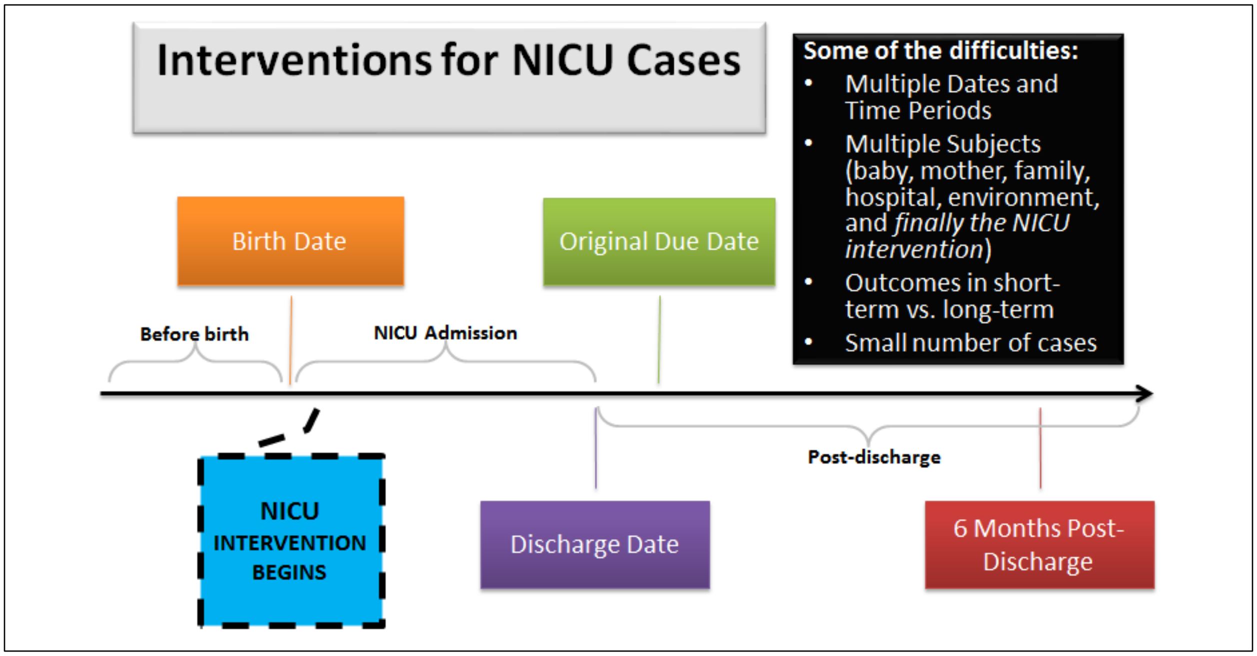


Figure 2. Causal Analysis Challenges Illustrated with a NICU Intervention Program Example

DEALING WITH CAUSAL ANALYSIS CHALLENGES

There are good ways to deal with these challenges of causal analysis. First, consider how an intervention is implemented and how this will create limitations for any causal analysis. It is difficult to prove a certain causal relationship between interventions and outcomes when you only have observational/transactional data, interventions that are self-selected by participants, or interventions provided to everyone all at the same time.

Second, critical reasoning needs to be used throughout the analysis process to determine causal relationships. A useful list for improving causal analysis critical reasoning is Hill's Criteria for Causation below (Hill, 1965):

- **Strength**: There may be a causal effect with small associations. However, the larger the association, the more likely that it is causal.
- **Consistency**: Consistent findings observed by different persons, in different places, with different samples strengthen the likelihood of a causal effect.
- **Specificity**: Causation is more likely if an outcome occurs with a specific population at a specific site and specific disease with no other likely explanation.
- **Temporality**: The effect has to occur after the cause. Also, if there is an expected delay between the cause and expected effect, then the effect must occur after that delay.
- **Biological gradient**: The level of exposure should generally coincide with the level of incidence and/or with the triggering of the effect.
- **Plausibility**: A plausible mechanism between cause and effect is helpful.
- **Coherence**: Coherence between epidemiological and laboratory findings increases the likelihood of an effect.
- **Experiment**: Occasionally it is possible to appeal to experimental evidence.
- **Analogy**: The effect of similar factors may be considered in causal reasoning.

Lastly, Causal Analytics should use methodologies that overcome causal analysis challenges like selection bias. Figure 3 ranks common methodologies on a scale from "does *less* to help to determine causation" to "does *more* to help determine causation". The reader should think about Figure 3 with a question in mind: "what if this was the ***primary methodology*** that I intended to use for an analysis – how helpful would it be for causal analysis?" To have the best chance of detecting and describing causal relationships, the experimental and quasi-experimental methods should be the primary methodology used.

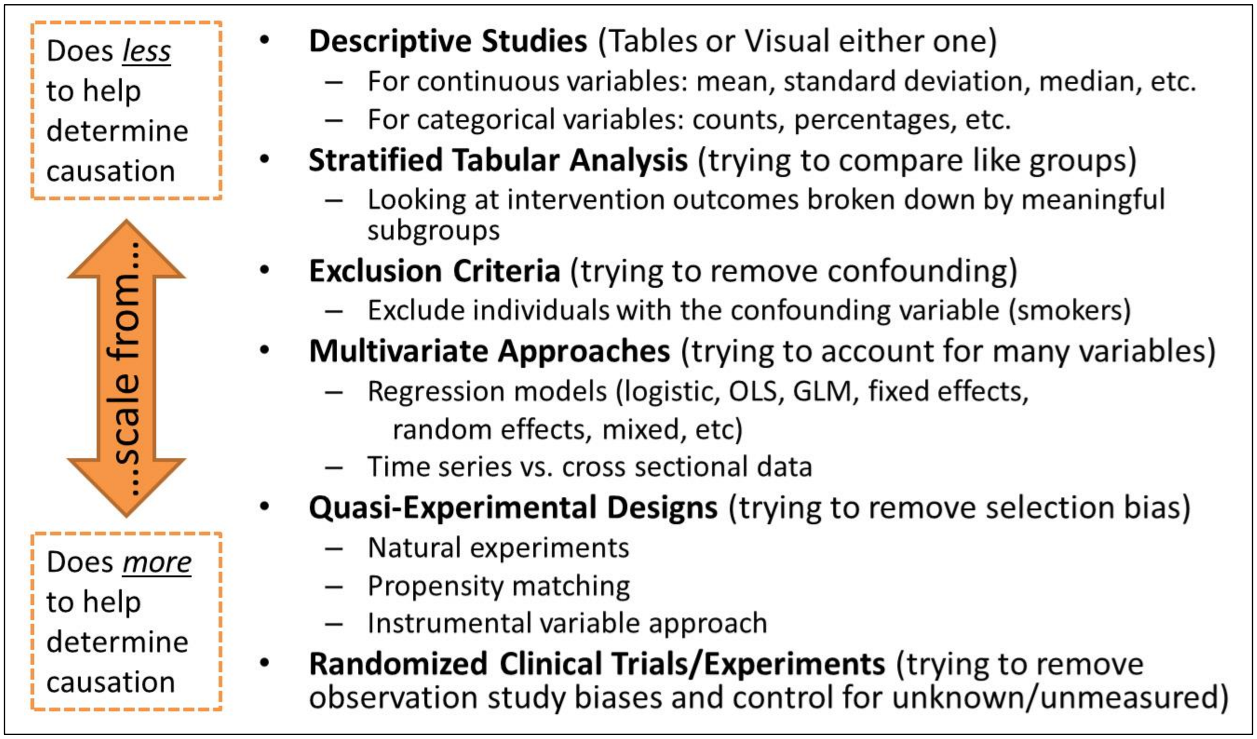


Figure 3. Methodologies Organized by Usefulness in Causal Analysis

CAUSAL ANALYTICS METHODS

Three applications of Causal Analytics are introduced in this section: "testing" whether outcomes change because of an intervention; "targeting" interventions that are likely to see greatest improvement in outcomes; and "tweaking" an intervention to control system outcomes. The "testing" methods are more retrospective and evaluative in nature whereas the "targeting" and "tweaking" methods help more with prospective business analysis and decisions.

TESTING: DETERMINE IF THE OUTCOMES CHANGE BECAUSE OF AN INTERVENTION

Before going directly into causal "testing" in a business context, there are two concepts that we will define. These are really important to understand in causal analysis.

- **Simpson's Paradox**: What is true for sub-groups of a population is not necessarily true for the aggregated whole population. This occurs when confounding or lurking factors exist among sub-groups in addition to an intervention. The confounding variables also relate to the outcome of interest.
- **Self-Selection Bias**: Self-selection occurs when individuals decide for themselves whether they will receive an intervention or not. Since individuals choose to receive an intervention, there may be systematic differences overall between the group of individuals who choose the intervention compared to those individuals who do not choose the intervention.

From these concepts, we learn that many inputs need to be accounted for - *in addition to an intervention* - so that we can determine how both the intervention and these other inputs relate to observed outcomes. Also, we learn that it is invalid to blindly compare a group that gets an intervention with a non-intervention group. To compare them, the groups would need to be similar or "exchangeable" (Causalytics, 2013).

Figure 4 diagrams the basic process of "testing" in Causal Analytics. The components of this process include a group that gets an intervention and a comparison group that does not get the intervention. These two groups should be similar or "exchangeable" across many factors in the pre-intervention period. Ideally, at the time of intervention, the only difference between Unit "A" and the other unit is that one gets an intervention and the other unit does not. The "testing" actually happens in the post-intervention period and involves comparing the outcomes of the two groups whether they are the same or different. If the two groups have different outcomes, then there is a good chance that the outcome differences can be explained by the intervention since these two units were "exchangeable" prior to the intervention.

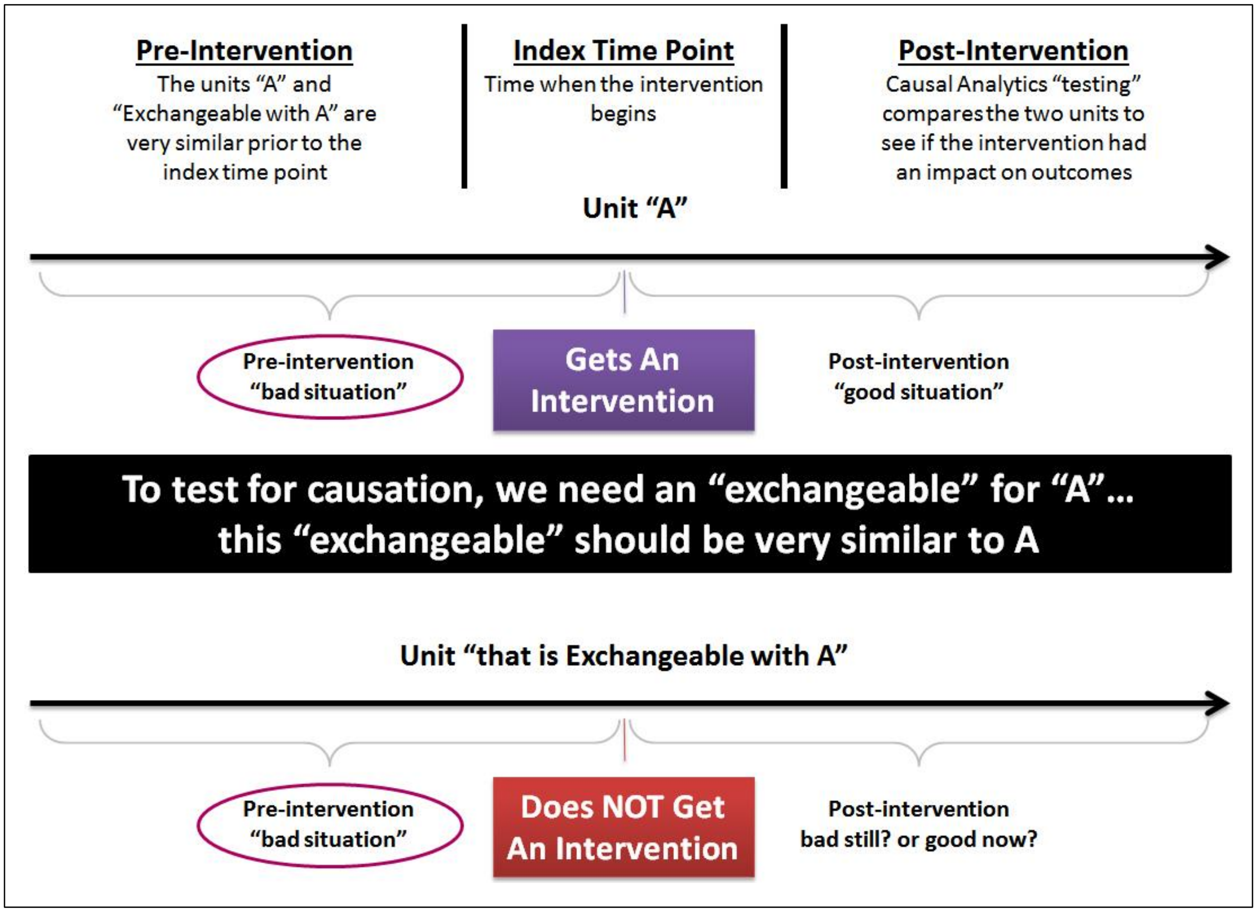


Figure 4. Basics of "Testing" in Causal Analytics

The most difficult part of "testing" in a business context is the task of obtaining an intervention group and non-intervention group that are "exchangeable" or very similar in many ways. From Figure 3 on the prior page, the best scenario would be to use experimental methods by *randomly assigning* some to an intervention group and others to a non-intervention group. This random assignment would decrease self-selection bias since there is no systematic or human selection of the interventions. Random assignment of the intervention would also decrease a bias called Omitted Variable Bias since the two groups would likely be similar to each other across the known/measured variables and the unknown/omitted variables.

There are practical limitations to randomly assigning interventions in business. Occasionally, random assignment might be possible in short-term pilot studies or when it is ethical to withhold a "5-10% sample" of the population from intervention for evaluation purposes. Typically though, random assignment of interventions is not possible in business because the intervention may be tied to transactions in the business process or the intervention is offered or provided to everyone (e.g. sometimes an intervention is given to everyone for ethical reasons). Next, we'll use an example from a hospital post-discharge intervention program to describe a partial workaround to the limitations of randomized interventions. The partial workaround requires the use of quasi-experimental methods such as propensity score matching.

Healthcare organizations (e.g. insurers, hospitals, and primary care physicians) often provide interventions to high-risk customers who have recently been discharged from a hospital – a post-discharge intervention. The outreach is usually completed by a nurse or specialist to assist members with difficulties in adhering to post-discharge instructions, medical equipment needs, medication adherence, scheduling a physician appointment, or other gaps that a customer may face after a hospital discharge. The goal of a program like this is to reduce the number of customers who get re-admitted to a hospital within 30 days. Decreasing 30-day readmission rates is an important quality of health care outcome.

To test if the intervention causes decreased re-admission rates, members who received the intervention can be matched to similar members who did not receive the intervention. The purpose of the matching process is to end with two "exchangeable" groups – with the main difference being that one group received the intervention and the other group did not. Three stages to this matching process are:

1) **Inclusion and Exclusion Criteria**

   The inclusion and exclusion criteria define the cohort of members who were deemed eligible for the intervention and also those who had too many confounding issues. For the post-discharge intervention inclusion/exclusion criteria might include: acute hospital admissions; discharged to home; referred to intervention within 14 days of discharge; at least 90 days of insurance coverage after discharge; exclusion of members in other higher-touch intervention programs.

2) **Exact Matching**

   Exact matching criteria are the characteristics of members that are so important that a matched pair from the intervention and non-intervention groups should be exactly the same on these particular characteristics. Some of the exact matching criteria for our example might include: insurance product type; discharge status code; prior admissions in the past 30 days flag.

3) **Probabilistic Matching**

   Probabilistic matching uses Propensity Score Matching to account for other many factors of the members through a propensity score model. The model input variables might include inputs such as: age; gender; date; medical costs in the 30 days before discharge; number of admits in the 30 days before discharge; Charlson Comorbidity Index; readmission prediction score; length of stay; admission from the emergency room.

After the matching process, the result is two similar groups: the intervention group and "exchangeable" non-intervention group. To test if the intervention decreases re-admissions, matched pairs statistical tests can be conducted to check for significant differences between the two groups. For further reading on propensity score methods in Causal Analytics "testing", refer to Fahner (2012) and the recommended reading *Analysis of Observational Healthcare Data Using SAS®*.

TARGETING: SELECT CASES THAT ARE LIKELY TO RESPOND WELL TO AN INTERVENTION

With "testing", we were able to estimate the average overall impact of an intervention on the outcomes (estimating an average causal effect). What if we now needed to know more about individuals who benefit the most and who benefit the least from an intervention (estimating an individual causal effect)? With this information, resources could be *targeted* towards those individuals who have the highest chance of responding favorably to an intervention. This links back to our core idea in the Causal Analytics framework section that "any attempt to intervene in some system or process entails the possibility of gain or loss." Causal Analytics "targeting" helps to *gain more and lose less*.

First note that in a simplistic view, there are two types of outcomes: positive outcomes and negative outcomes. Figure 5 separates out the outcome types and the responses to an intervention:

- some will have *positive* outcomes *because of* receiving an intervention (The Persuadables)
- some will have *positive* outcomes *regardless of* the intervention (The Sure Things)
- some will have *negative* outcomes *regardless of* the intervention (The Lost Causes)
- some will have *negative* outcomes *because of* receiving an intervention (The Do Not Disturbs)

Causal Analytics identifies "The Persuadables" and "Do Not Disturbs" through Uplift Modeling methods.

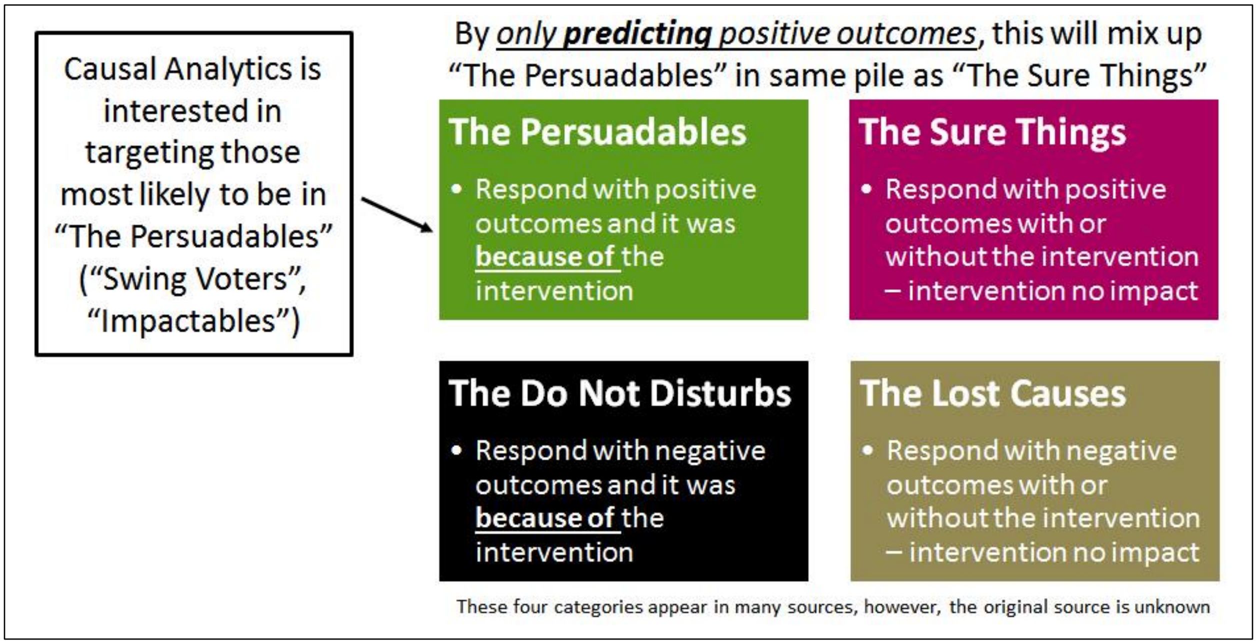


Figure 5. Different Outcomes and Responses to an Intervention

"Targeting" is an area where Predictive Analytics and Causal Analytics are related. However, there are important differences between the two approaches as seen in the three examples below.

Example 1: Causal Analytics vs. Predictive Analytics

- Predictive Analytics Targeting: predict which customers are likely to complain
- Causal Analytics Targeting: predict which customers are likely to complain <u>and would instead not complain if they received a proactive intervention program</u>

Example 2: Causal Analytics vs. Predictive Analytics

- Predictive Analytics Targeting: predict patients trending towards an adverse event
- Causal Analytics Targeting: predict patients trending towards a <u>preventable</u> adverse event <u>for which an intervention program might stop the adverse event from happening</u>

Example 3: Causal Analytics vs. Predictive Analytics

- Predictive Analytics Targeting: predict patients at-risk for missing clinic appointments
- Causal Analytics Targeting: predict patients at-risk for missing clinic appointments <u>who would instead keep their clinic appointment if contacted with a reminder call the day before</u>

With limited time and resources to deliver interventions, there is a larger return using targeting methods based on the Causal Analytics framework. Recall from the "testing" section, we discussed an intervention program for members with a recent discharge from the hospital. As part of the evaluation of this program, uplift modeling methods discussed in Causalitics (2013), Khargharia (2013), and Kubiak (2012) can provide the program with the characteristics of members who respond most favorably to the intervention program.

The analysis generated from Causal Analytics "targeting" can be used in a couple of ways: 1) scoring future cases for whether a case has a high chance of being in The Persuadables; or 2) program improvements to try and influence any "Do Not Disturbs" towards a more favorable response to the intervention.

TWEAKING: MANIPULATE AN INTERVENTION IN ORDER TO CONTROL THE OUTCOMES

The last area of Causal Analytics is “tweaking” interventions to control outcomes. Any intervention in a system has features that can be manipulated to have different causal effects on the outcome of interest. These manipulation features generally can be categorized as dosages, delivery method, and timing.

Dosages are the level of exposure to an intervention (“how much”, “how many”). Some examples include the amount of a drug, the level of out-of-pocket expenses on different insurance plans, interest rates on credit cards, or the number of calls per week for a telephonic intervention. Even the basic “yes got an intervention” or “no did not get an intervention” would be considered two dosage levels.

Delivery methods involve the way a subject comes in contact with an intervention (“what”, “where”). For example, a retention campaign might contact some members by phone, others by direct mail, and others with an in-home visit. A drug might be given as an injection, pill, or nasal spray.

Timing an intervention manipulates the order and sequencing of the intervention (“when”, “what time”, “how long”). A drug might be prescribed twice-per-day in small amounts or once-per-day in a larger amount. A telephonic intervention might have followup calls two weeks from the initial call or thirty days from an initial call and this followup timing might impact the outcomes in different ways.

Causal Analytics “tweaking” encompasses methods used for process control, optimization, and decision-making by modeling how dosages, delivery methods, and timing impact the outcomes.

Consider an example from the field of higher education for Causal Analytics “tweaking” (one of the authors has prior experience as a higher education researcher). Financial aid leveraging “is an analytical tool that enables admissions and financial aid administrators to estimate the amount of financial aid (regardless of formal need formulas) that would be necessary to increase the probability that a student with a specified set of characteristics would enroll.” (Hossler, 2000, p. 83). In other words, financial aid leveraging is about “tweaking” the scholarship dollars offered to admitted students to control who enrolls at a university. A university should not offer large scholarships to students who would enroll with smaller scholarships. Similarly, a university may want to increase enrollments of out-of-state or “high quality” students and will need to know the scholarship amounts needed to achieve these enrollment goals.

A model to estimate the causal relationship between scholarships and enrollment decisions might include variables such as: residency of the student; distance to the university; market; gender; race; first generation flag; financial need level; date of admission; high school GPA; ACT or SAT score; and scholarship amount. Although many variables are included in the model, the main information of interest is the statistical relationship between scholarship dollars and enrollment decisions. Figure 6 provides a mock-up example of a decision tool built on this statistical relationship. The tool allows testing of new scholarship awards and then seeing how multiple outcomes are impacted from the scholarship tweaks.

This mock-up financial aid leveraging decision tool takes the statistical relationship between “scholarship awards” and “enrollment decisions” and provides a look at estimated yield rates, headcount, out-of-state makeup, high school GPA average, short and long term net revenue, and more.

A tool like this can be used to decide which levels of scholarships best achieve student quality and net revenue goals of a university.

Cohort Monitor	MaxTopNet	Baseline	Test	Change T-B
Yield Rate	52.0%	45.5%	48.4%	2.9%
Enroll Headcount	4050	4050	4050	0
%Out-State	31%	24%	26%	3%
HSGPA	3.61	3.49	3.55	0.05
ACT Composite	25.29	24.60	24.87	0.27
First Fall GPA	2.90	2.75	2.82	0.07
Retention Rate	81.3%	79.0%	80.0%	1.0%
Yr1 Tuition Revenue	$43,475,197	$41,066,651	$42,044,749	$978,098
Yr1 Award Costs	$14,076,343	$10,110,905	$11,289,003	$1,178,098
Yr1 Discount Rate	32%	25%	27%	2%
Tot1Yr Net Revenue	$29,398,853	$30,955,745	$30,755,745	-$200,000
Tot2Yr Net Revenue	$52,683,693	$54,736,426	$54,751,687	$15,262
Tot3Yr Net Revenue	$73,388,893	$75,266,836	$75,857,809	$590,973

Figure 6. Example Decision-Tool for Financial Aid Leveraging Example

CONCLUSION

Causal Analytics "testing" evaluates the overall impact of an intervention, "targeting" identifies who is impacted the most and least, and "tweaking" provides insight for optimizing an intervention. Attention is given to inputs, interventions, and the outcomes. Integrating Causal Analytics into business processes can help improve ROI, design personalized interventions, and improve customer success programs.

REFERENCES

Academy Health. (2002). "Does X Really Cause Y? [White paper]." Accessed from http://www.academyhealth.org/files/FileDownloads/DoesXCauseY.pdf.

Bogard, M. (2014). "Big Data: Don't Throw the Baby Out with the Bathwater." Accessed March 11, 2015 from http://www.econometricsense.blogspot.com/2014/05/big-data-dont-throw-baby-out-with.html.

Causalytics. (2013). "Cadit Modeling for Estimating Individual Causal Effects [White paper]." Accessed March 11, 2015 from http://www.causalytics.com/.

Fahner, G. (2012). "Estimating Causal Effects of Credit Decisions." *International Journal of Forecasting*, 28: 248-260.

Hill, A. (1965). "The Environment and Disease: Association or Causation?" *Proceedings of the Royal Society of Medicine*, 58(5): 295-300. Accessed March 11, 2015 from http://www.edwardtufte.com/tufte/hill.

Hossler, D. (2000). "The Role of Financial Aid in Enrollment Management" *New Directions for Student Services*, 89: 77-90.

Khargharia, J. (2013). "How Incremental Response Modeling Can Help You Reach the Right Target Group More Precisely." Accessed March 11, 2015 from http://blogs.sas.com/content/subconsciousmusings/2013/07/12/how-incremental-response-modeling-can-help-you-reach-the-right-target-group-more-precisely/.

Kubiak, R. (2012). "Net Lift Model for Effective Direct Marketing Campaigns at 1800flowers.com". Accessed March 11, 2015 from http://support.sas.com/resources/papers/proceedings12/108-2012.pdf.

Lee, T., Zhang, R., Meng, X., Ryan, L. (2013). "Incremental Response Modeling Using SAS® Enterprise Miner™". Accessed March 11, 2015 from http://support.sas.com/resources/papers/proceedings13/096-2013.pdf.

RECOMMENDED READING

- *Analysis of Observational Healthcare Data Using SAS®*

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Jason Pieratt
jason.pieratt@gmail.com

Edmund Berry, Ph.D.
berryedmunda@yahoo.com