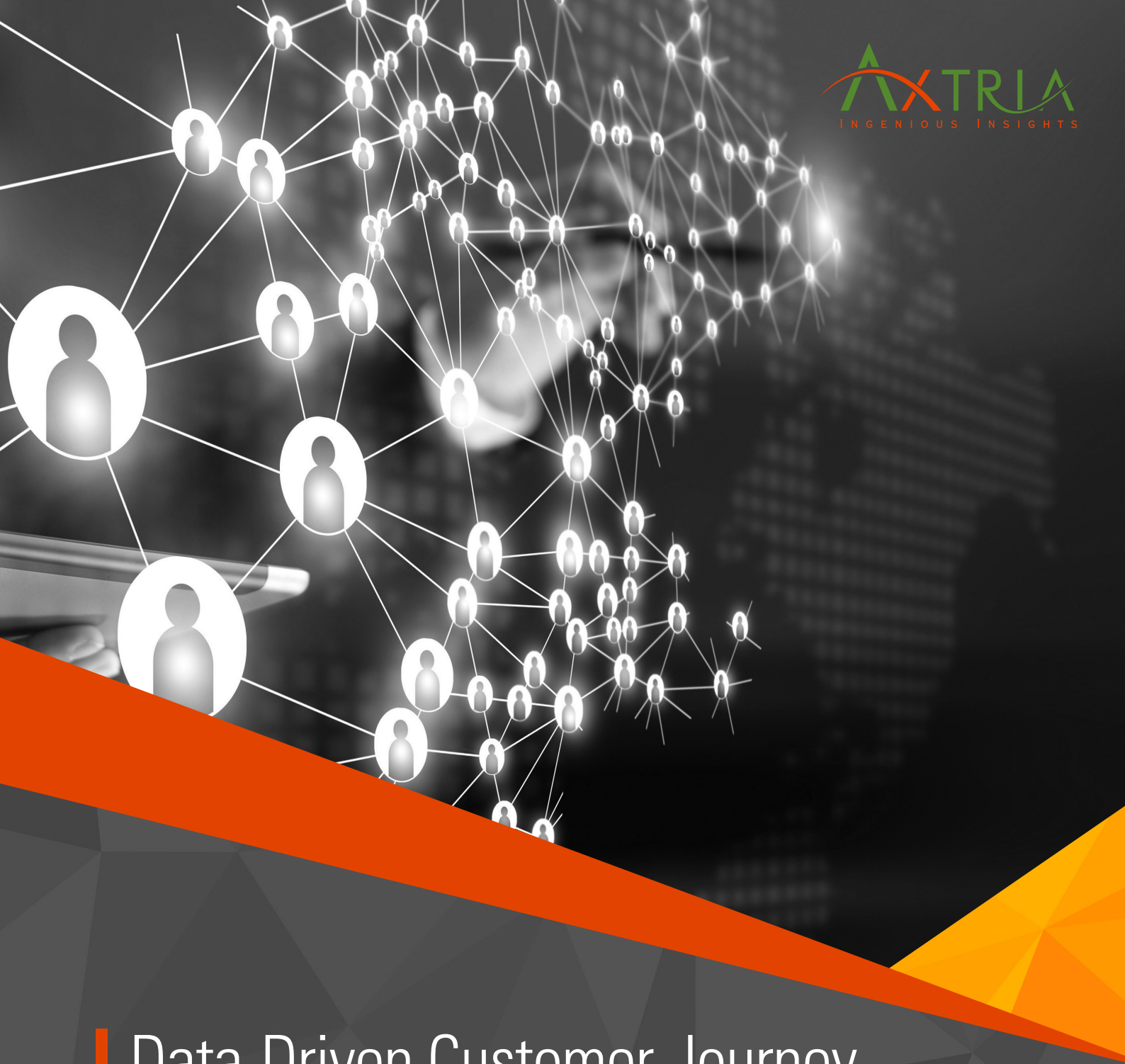


Data-Driven Customer Journey Design Using Bayesian Network Methodology

May 2021

Data-Driven Customer Journey Design Using Bayesian Network Methodology

Ramesh Krishnan, Principal, Axtria Inc.
Teis Kristensen, Project Lead, Axtria Inc.
Kritika Singhal, Senior Associate, Axtria Inc.

EXECUTIVE SUMMARY

Patients and healthcare providers have become embedded in a media ecosystem characterized by strong competition for attention. For pharmaceutical marketers to deliver impactful messages, it is necessary to make the customer the focal point of analysis. A Bayesian Network methodology allows this via a data-driven customer journey analysis. Bayesian networks provide information on the next best action in a customer journey, identify journeys with a high probability of a positive outcome, and the attribution of channels on customer outcomes. Together these insights inform the design of customer journeys, omnichannel orchestration, and aim to facilitate an exploration of the customer experience. The results from a simulated dataset illustrate the type of quantitative insights that can be generated, while the practical implications for pharmaceutical operation teams are highlighted.

1. Crowded Media Ecosystem

The onset of COVID-19 has sped up changes in the pharmaceutical ecosystem. COVID-19 has and continues to affect economies, technology consumption patterns, and social behaviors (1-3). The pandemic has illuminated the link between public health and people's daily lives, and it has brought the importance of medicines to the forefront of people's minds. Pharmaceutical marketers can utilize the reinforced focus on health to drive better healthcare outcomes for their patients. However, when doing so, they are faced with strong competition for attention by other commercial actors, consolidation restricting access to HCPs, and a continuous move towards digital interactions across stakeholder groups (4). Pharmaceutical marketers must stand out in a crowded and noisy media ecosystem in order to inform HCPs and patients of life-improving medicines.

Figure 1. Competitive Media Ecosystem

To assess and navigate the complexities of the pharma environment, marketers have turned to data-driven methodologies that help measure and optimize the impact of promotional initiatives. By using observed data, it enables marketers to make empirical and proactive decisions. Several data-driven tools exist to inform strategic, tactical, and operational decision-making. For example, marketing mix analyses are often used as a strategic tool to inform budget allocation across promotional channels. In contrast, ROI analyses provide a tactical overview of the most profitable approaches to reach forecast objectives. And at the operational level, machine learning models are deployed to predict the next best action at points of contact between HCPs and pharma companies. As analytic methods and capabilities mature, more and more data-driven insights are becoming available.

A focus on website interactions has led to the development of digital attribution methodologies. Digital attribution is a popular data-driven approach used to understand customer journeys in digital customer interactions. Digital attribution describes the impact of interaction points associated with outcomes such as a click, conversion, or website sale. This is achieved by using rule-based methods that identify, for example, the first or last interaction point observed before a specific outcome. The method is used to understand points of interaction that led to a positive outcome but is limited by a focus on single parts of the customer journey.

Simultaneously, a move towards personalizing promotional activities and an increased focus on omnichannel orchestration have gained traction. These initiatives aim to take a customer-centric perspective that emphasizes the holistic experience of health administrators, providers, and patients. A customer journey is the sequence of interactions between an individual and a company over time. The concept has gained traction because it seeks to understand the customer from their first interaction and throughout their complete engagement with a company.

This white paper describes how a Bayesian Network methodology (5) can be used as the engine for a data-driven

examination of customer journeys. The method combines the focus on interactions seen in digital attribution with an emphasis on customer journeys. The method is especially suited to explore customer journeys as the unfolding of events, over time, is the basic unit of observation used in the analysis (6). A Bayesian network calculates the probability of an outcome based on a sequence of observed events (7). The insights generated include the most frequently observed customer journeys, identification of customer journeys with high impact, attribution of outcomes to interaction points in the customer journey, and the next best interaction point following an observed sequence of promotions.

The following sections aim to provide an overview of a data-driven customer journey analysis based on a Bayesian Network methodology. Section 2 describes how a customer journey framework can inform promotion orchestration both within and across channels, reviewing methods used for data-driven examination of customer journeys, and highlights the considerations used in applying a Bayesian Network methodology to generate insights. Section 3 provides the theoretical and mathematical underpinnings of the work. Subsequently, Section 4 presents the Bayesian Network methodology results, specifically the structure of the simulation data, central elements of model development, and the built network's queried results. The last section discusses the benefits and limitations of a Bayesian network method to generate customer journey insights and the broader implications for pharmaceutical marketers.

2. Data-Driven Customer Journey Design

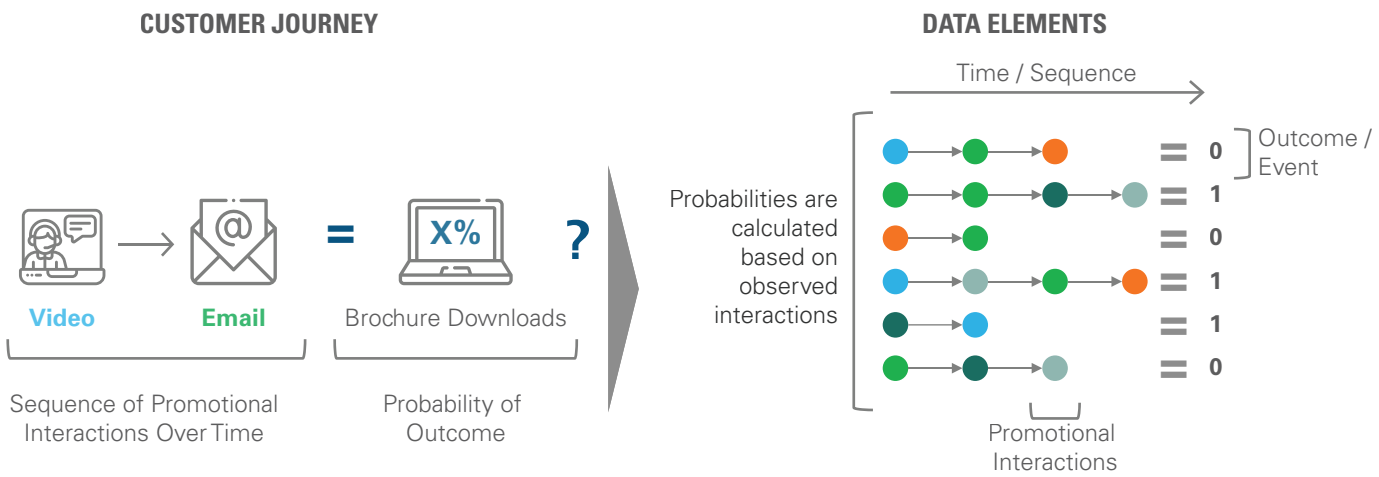
The concept of a customer journey has been used in a multitude of contexts with an equal number of nuances to its meaning. The popularity of the concept has resulted in ambiguity, which requires specificity in defining what a customer journey is. At the core of the concept is the understanding that optimization of promotional initiatives must be done with an emphasis on ensuring customers' positive experiences. The application of a customer journey perspective is done to enable marketers to map pain points and opportunities for improvement. From a Bayesian Network perspective, the points of interaction that unfold over time between customers and companies are the elements that make up a customer journey (8). Figure 2 illustrates how a healthcare provider has multiple interaction points before a contract is signed.

Figure 2. Journey of a Healthcare Provider

When applying a customer journey framework to understand promotional activity, it is essential to clearly identify who is considered a customer and what channels to include in the analysis. Even though these questions may seem trivial, the pharmaceutical industry has multiple stakeholders that could be considered customers, such as patients, providers, and administrators. Therefore, specifying the customer beforehand functions as a funnel that reduces the promotional channels that must be considered. A customer journey can consist of interaction points both within and across channels. For example, a within-channel examination may focus on the sequence of content delivered within a channel, while a cross-channel examination could hone in on interaction points such as executed details, email contacts, etc. The focus of the analysis moves from specific channels towards emphasizing the customers' experience across channels.

Bayesian statistics are suited for customer journey analytics because they encompass time as a central element in the modeling process (5-7). Bayesian methods, such as Markov and Bayesian network models, calculate the probability of an outcome based on a prior sequence of events (5). Specifically, Markov and Bayesian network models ingest a data format that mirrors the points of interaction between companies and customers (6). This approach fits with a customer journey framework as a sequence of events are the points of interaction between a life science company and its customers.

Figure 3. From Customer Journey to Data Components



Both Markov chain and Bayesian network methods allow for the examination and active design of promotional sequences (7). The difference between Markov and Bayesian network methods lies in the assumption made about the relational directionality between model variables. Markov network models assume that only undirected relationships exist, while Bayesian networks are directed acyclic graphs, assuming directionality in the probabilistic impact of events (6). Since there is an assumption of directionality in Bayesian networks, it is possible to a priori specify the direction of relationships that can/cannot exist when these are learned between promotional sequences and outcomes. Only Bayesian networks are expanded upon in this paper due to the ability of the modeling framework to benefit from domain expertise by specifying directional relationships (5).

3. Theoretical Foundation of Bayesian Networks

A Bayesian network is a mathematical representation of the probabilistic relationships between random variables. The objective of a Bayesian network is to model the posterior conditional probability distribution of an outcome variable given a series of observed evidence (6). Figure 4 illustrates the insights generated from a Bayesian network. Bayesian networks have gained traction in multiple contexts, such as health outcomes research and medical decision analyses, as the method supports the examination of uncertainty and causality surrounding events.

Figure 4. Bayesian Network Insight

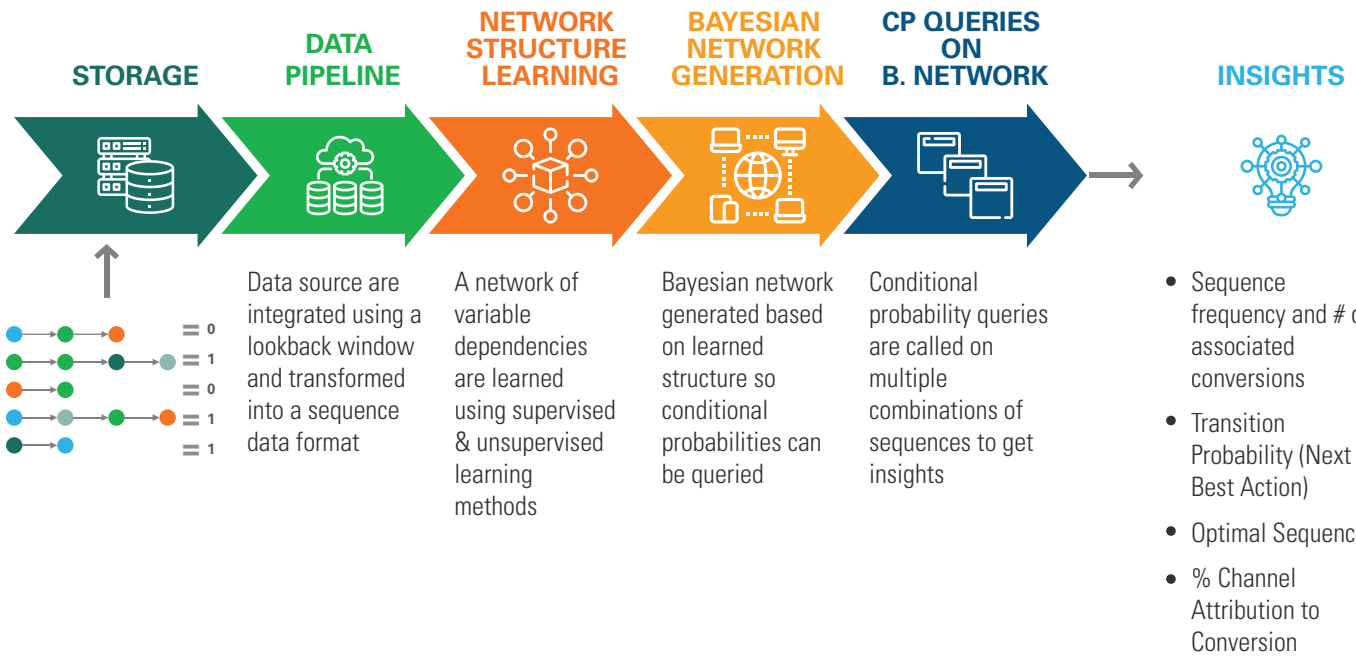
The probability that an event happens or a success criteria is met following a sequences of interactions

Figure X: percent change that an event will happen given a sequence of promotional interactions

A three-stage process is used to generate Bayesian network insights (6-7). First, the relational structures between model variables are learned using a combination of supervised and unsupervised learning. Supervised learning allows for the modeler to specify which relationships can exist, such as no relationships going back in time or only allow outgoing ties for attribute variables. These stipulations of the learning are then applied when unsupervised learning algorithms are applied to identify relational structures in the data.

Once a network structure is learned, the structure is used to generate a Bayesian network. The Bayesian network contains the structure and strength of relational dependence between variables and can be queried to predict an outcome if prior interactions are given. Queries can be made on observed and unobserved sequences of interactions, as Bayesian networks draw on observed data to estimate the impact of unobserved sequences.

Figure 5. Components of the Analytic Framework



A Bayesian network has three central components; nodes representing variables from the dataset, edges between nodes indicating causal relationships, and the conditional probability distributions associated with each node (7). If a causal relationship exists between two variables, the corresponding nodes in the network have a directed edge between them. The conditional probability distributions of the interaction points (nodes) are determined by Bayes' Theorem: for events ***A*** *and* ***B***.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Equation 1

Given the conditional probabilities of prior events, it is possible to approximate the posterior distributions of the nodes. In other words, the Bayesian network model calculates the dependencies between variables and creates a network of causal relationships that can be queried (5). Each variable is treated as conditionally independent of all its non-descendants so that probabilistic relationships are assumed only along the directed edges of the network, giving:

$$P(X_1, X_2, \cdots, X_n) = \prod_{i=1}^{n} P(X_i | parents(X_i))$$

Equation 2

Here, *X_1, …, X_n* is a set of nodes of the network. Queries are run on the network by moving through the directed acyclic edges. This approach captures the conditional dependencies between variables without the need to prune all possible relationships existing in the Bayesian network. The probability of an outcome is calculated by plugging in the parameters from the conditional probability tables. Var1 and VarX are different variables or nodes of the network:

$$P(Outcome = 1 | X_1 = 1, \cdots, X_n = 1) = \frac{P(Outcome = 1, X_1 = 1, \cdots, X_n = 1)}{P(X_1 = 1, \cdots, X_n = 1)}$$

Equation 3

4. Bayesian Network Results

The following section presents the insights generated from a Bayesian network using simulated data. Bayesian network models were run on a simulated and e-commerce dataset. Only simulated results are reported as the e-commerce dataset comes from a non-healthcare setting. The use of a simulated dataset was done to ensure the validity and reliability of the modeling framework, as a simulated dataset only contains specified relationship patterns that can be compared to model results. A simulated dataset containing 200,000 journeys with five interaction points was generated as the basis for the Bayesian network modeling. The assumptions imposed on the data aimed to mimic observed patterns in the commercial data. Each journey in the data corresponds to the unique customer journey of an individual. Several assumptions were imposed on the generated data, such as an increase in customer journey length resulting in a decrease in frequency and a corresponding increase in outcome probability.

One of the central insights generated from a Bayesian network model is the calculation of outcome probabilities associated with the next interaction point in the customer journey - the next best action. This type of insight can be invaluable for a sales rep in the field, as reps can be provided recommendations on the approach that has the highest probability of a positive outcome. The insights can, for example, be delivered to the sales rep before engaging with a HCP. Table 1 shows a customer journey consisting of three interaction points. Queries on the Bayesian network were run to calculate the probability associated with an additional interaction point in the customer journey, while using the current outcome probability as a reference point. The results show that the currently observed customer journey has a 23% probability of the desired outcome. The probability increases if an additional interaction point is added to the

journey. Specifically, the outcome probability is highest if the next step in the customer journey is channel five. In this way, a Bayesian network can support operational decision-making by helping to establish the next best action for sales reps in the field.

Table 1. Next Best Action in Customer Journey with Three Interaction Points

Observed Customer Journey	Outcome Probability
C1 > C2 > C3	Current Journey
Next Best Action for Customer Journey	-
C1 > C2 > C3 > C1	16%
C1 > C2 > C3 > C2	21%
C1 > C2 > C3 > C3	18%
C1 > C2 > C3 > C4	10%
C1 > C2 > C3 > C5	24%

Note: C is used as an abbreviation of channel interaction points

In addition to supporting pharmaceutical sales reps in the field, it is possible to understand the type of customer journeys most likely to result in a desired outcome. These insights can be used to understand the impact of individual channels depending upon their placement in a customer journey and thus help design the optimal sequence of promotions across channels. Table 2 shows the five most impactful customer journeys from a single interaction point to three interaction points. The results indicate that the most impactful point of interaction is channel five, with a 21.85% probability for a positive outcome. However, the results also illustrate that repeating the same channel over time does not hold the same effect. The benefit of using a Bayesian network to understand optimal promotional sequences is that interaction effects and channel position in a customer journey are considered when outcome probabilities are calculated. The insights allow for a tactical evaluation of the optimal sequence of promotional channels. The knowledge of impactful customer journeys will enable marketers to support the sales force as they engage customers.

Table 2. Optimal Sequence of Promotion by Customer Journey Length

Journey Length of One		Journey Length of Two		Journey Length of Three	
Journey	Outcome Probability	Journey	Outcome Probability	Journey	Outcome Probability
C5	21.85%	C5 > C4	34.33%	C5 > C1 > C4	83.33%
C4	20.26%	C4 > C5	29.41%	C4 > C5 > C1	80.00%
C2	19.93%	C2 > C4	28.57%	C4 > C3 > C1	66.67%
C3	18.97%	C1 > C1	26.23%	C1 > C2 > C5	66.67%
C1	18.74%	C5 > C3	25.64%	C3 > C1 > C2	60.00%

Note: Only the five most impactful combinations of promotion are included.

Lastly, a Bayesian network can be used to attribute the impact of a channel to an outcome variable. This is done by using the difference in probability between configurations of promotion sequences. The method of the calculation follows the principles applied to game theory methods for coalition attribution (8). Axioms to derive at a Shapley value were also tested. Table 3 reports the attribution results. The impact of a channel was defined as the loss of probability if the channel was excluded from the customer journey. The results show the percentage of outcomes attributed to each channel. The

results show that channel five continues to have a strong impact on the probability of an outcome to happen across our insights. Channel five can be attributed to 21.52% of the outcomes, compared to the 18.57% attributed to channel one. These insights can form the basis for further analyses. For example, the financial impact of a change in the sequence of promotion for a number of HCPs can be calculated. The insights facilitate the strategic planning of promotional activities by providing a quantifiable understanding of the customer journey and supporting omnichannel orchestration dynamics.

In addition to simulated and observed data, several machine learning models were used to test and validate the Bayesian network approach. Logistic regression, random forest, gradient boosted regression, and naïve Bayes models were tested on the simulated data. The machine learning models predictive power was hindered by skewness in the data towards non-conversion entries. The best machine learning models, tested on multiple transformed datasets, had a classification error rate of approximately 50%, which suggests a Bayesian network approach is more appreciate for solving the problem of understanding the impact of sequences on outcomes. Traditional machine learning models may be more applicable with data structures less focused on promotional sequences. The strength of a Bayesian network approach is the ability to better understand the dynamics between promotional channels by highlighting the impact of specific promotion sequences from a probabilistic lens.

Table 3. Bayesian Network-Based Channel Attribution

Channel	Attributed Outcomes	Attribution Percent
C1	4116	18.57%
C2	4261	19.23%
C3	4453	20.09%
C4	4562	20.59%
C5	4768	21.52%

Note: Game theory axioms are used to calculate attribution that follows the assumptions of cooperative games

5. Practical Implications of using Bayesian Network Insights

The use of a conceptual framework based on probability acknowledges the uncertainty with which an outcome happens. By taking a probabilistic approach to understanding customer journeys, time and sequence of events are included as fundamental components in the statistical modeling process. The probabilistic queries run on Bayesian networks can be used to provide operational, tactical, and strategic insights that can help the orchestration of omnichannel efforts.

Examining a customer journey using Bayesian networks provides insight into how different combinations of promotional interactions can impact the probability of an outcome. The outcome sought can be any success criteria associated with a customer over time, such as clicking an email link, downloading an information brochure, and placing Rx orders. The generated insights about customer interactions enable an active design of customer journeys that ensure HCPs have the information and inventory they need to serve their communities.

Bayesian networks should be seen as a way to provide insights that are underpinned by a historical understanding of customer interactions while allowing for the exploration of unobserved promotion sequences and how these sequences can increase the probability of a positive outcome. The combination of a customer-centric approach with Bayesian network modeling can help optimize intermediate campaign successes, Rx, and customer satisfaction by quantifying the impact of promotional interactions. Bayesian network modeling informs the next best action in a current customer journey, allowing the mapping of the impactful sequences of promotions and the attribution of promotional channels to outcomes. The goal is to create a data-driven approach to customer journeys that can help facilitate and guide discussions among commercial operations teams.

Reach out to Axtria if you want to know more about data-driven approaches to customer journey analytics. Bayesian networks provide insights based on the lived experience of customers that facilitate the design and orchestration of promotional activities across channels.

References

1. Cohen O, Fox B, Mills N, & Wright P. COVID-19 and commercial pharma: Navigating an uneven recovery. Available from: https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/covid-19-and-commercial-pharma-navigating-an-uneven-recovery [Accessed July 15, 2020]
2. Baum MA, Ognyanova K, Chwe H, et al. *The state of the nation*: a *50-state covid-19 survey - report #14: misinformation and vaccine acceptance*. Report number:14, 2020. Available from: http://www.kateto.net/covid19/COVID19%20CONSORTIUM%20REPORT%2014%20MISINFO%20SEP%202020.pdf [Accessed October 22, 2020]
3. Mullin, E. The Most Important Takeaways from the FDA's Big Covid-19 Vaccine Meeting. *Medium*. https://coronavirus.medium.com/the-most-important-takeaways-from-the-fdas-big-covid-19-vaccine-meeting-a635bdfc9be4 [Accessed September 5, 2020]
4. Deloitte. The Great Consolidation: The potential for rapid consolidation of health systems. Available from: https://www2.deloitte.com/us/en/pages/life-sciences-and-health-care/articles/great-consolidation-health-systems.html [Accessed July 7, 2020]
5. Murphy K. *A Brief Introduction to Graphical Models and Bayesian Networks*. Weblog. Available from: http://www2.denizyuret.com/ref/murphy/intro_gm.pdf [Accessed September 20, 2020]
6. Scutari M, Strimmer K. Introduction to graphical modelling. *arXiv*. [Pre-print] 2010. Available from: https://arxiv.org/pdf/1005.1036.pdf
7. Koller D, Friedman N. *Probabilistic Graphical Models*: *Principles and Techniques*. Cambridge: MIT Press; 2009.
8. Anderl E, Becker I, Von Wangenheim F, Schumann JH. Mapping the customer journey: Lessons learned from graph-based online attribution modeling. *International Journal of Research in Marketing*. 2016;33(3): 457-74.
9. Shapley LS. Stochastic games. *Proceedings of the National Academy of Sciences*. 1953;39(10): 1095-1100.

Ramesh Krishnan
Principal, Axtria Inc.
300 Connell Drive, Suite 5000
Berkeley Heights, NJ 07922
E: ramesh.krishnan@axtria.com

Teis Kristensen
Project Lead, Axtria Inc.
300 Connell Drive, Suite 5000
Berkeley Heights, NJ 07922
E: teis.kristensen@axtria.com

Kritika Singhal
Senior Associate, Axtria Inc.
300 Connell Drive, Suite 5000
Berkeley Heights, NJ 07922
E: kritika.singhal@axtria.com

Contact Us

+1-877-9AXTRIA
info@axtria.com

Disclaimer

Axtria® understands the compliance requirements behind personalization and we do not work with any personally identifiable data that can identify an end-customer of a business.

We have the strictest data security guidelines in place as we work with businesses to improve the experience for their customers.

www.axtria.com
facebook.com/AxtriaInc/
info@axtria.com
linkedin.com/company/axtria
@Axtria

Founded in 2010, Axtria is a global provider of cloud software and data analytics to the Life Sciences industry. We help Life Sciences companies transform the product commercialization journey to drive sales growth and improve healthcare outcomes for patients. We continue to leapfrog competition with platforms that deploy Artificial Intelligence and Machine Learning. Our cloud-based platforms - Axtria DataMAx™, Axtria SalesIQ™, Axtria InsightsMAx™ and Axtria MarketingIQ™ - enable customers to efficiently manage data, leverage data science to deliver insights for sales and marketing planning, and manage end-to-end commercial operations. We help customers in the complete journey from Data to Insights to Operations.

For more information, visit www.axtria.com

Follow Axtria on Twitter, Facebook and LinkedIn