White Paper

Algorithmic Forecasting for the Life Sciences Industry

How to bring speed, accuracy, and control to commercial forecasting

RICK JOHNSTON, PhD, Senior Principal, Consulting Services, IQVIA
ALEXANDRA TATARU, MS, MBA, Product Lead, Consulting Services, IQVIA

Table of contents

Executive overview

Imagine being able to integrate thousands of pieces of market and product data into your commercial forecast – in just minutes. Now imagine being able to update that forecast in real-time, as often as you like, with the most current data available. With manual commercial forecasting this would be impossible. But for companies that have implemented algorithmic forecasting, it's the new normal.

Algorithmic forecasting uses real-time data feeds, automated algorithms including machine learning, and an updated forecast workflow to predict commercial performance. Its premise is simple: automate forecast updates as new data becomes available, and reduce human manual labor in the forecast process. This brings speed, accuracy, and deeper insights to commercial forecasts and frees up forecasters to spend their time on strategic tasks like identifying exactly what questions need to be answered, interpreting outcomes, and informing product strategy.

This type of automation and analytics is already being used across the life sciences industry to speed insights and support informed decision-making. While commercial forecasting teams have been slower to join the 'algorithm revolution', the growing pressure to analyze more data faster as well as the potential for greater accuracy is pushing forecast teams to look for better and more streamlined approaches to use automation and machine learning.

This whitepaper offers forecasting teams a framework and roadmap for adopting algorithmic forecasting. Whether they have already begun their digital transition, or are still using spreadsheets to manage data, this paper provides practical steps and best practices for deploying the technology, generating early wins, and convincing stakeholders of the strategic benefits of embracing this real-time forecasting approach.

Forecasters are struggling with manual forecast tools and processes

In the traditional forecast process, data is manually gathered by forecasters and entered into a forecast environment, most often an Excel spreadsheet. During this process, forecasters may find themselves spending days, weeks or months researching, exchanging messages with other team members, and tweaking numbers to see the impact on results. Those numbers then pass through a sequence of review by other groups within the company – management teams, supply chain, and potentially others – each of which may request additional iterations based on their view of what they expect.

There are a number of types of issues that forecasters have expressed with this traditional process:

ANALYSIS TIME

> *"It takes weeks to get results when I ask for an alternative scenario, and I don't know why."*

When so much of the data collection and entry is manual, or potentially involves interactions between teams in different countries or continents, turnaround times can be long. Leadership gets frustrated at the slow response, and forecasters get frustrated when seemingly small requests add up to a large load on their time.

GLOBAL VS. LOCAL

> *"Coordinating forecasting activities between global and local teams has been a big challenge. Local affiliates currently drive the forecast, and global has poor visibility into how assumptions drive the final numbers."*

Global and local teams often struggle to strike a balance with regards to how tailored should the forecast structure be to each market, who owns the forecast, and to what extent expert opinions should be allowed to shift the forecast. Global and local teams typically maintain multiple versions with no single source of numbers to drive decision making.

CONSISTENCY

> *"There's not a lot of consistency in the way the team does forecasting – we have some home-grown models, some from vendors… We don't really have an end-to-end forecasting process."*

Small and medium sized companies often struggle with setting up an infrastructure that enables them to maintain consistency across products in their portfolio. Especially when there's a lot of business development and early stage forecasting, leadership may find that the comparisons are not "apples to apples".

A DELUGE OF DATA

> *"We are under-utilizing existing healthcare data we purchase. For forecasting purposes, we only review our sales data once per year, but it is updated every month."*

The supply of healthcare data has grown faster than forecasting teams' analysis capabilities, and teams are looking for ways to better integrate these sources of information. Anchoring forecasts in real-time data also reduces opportunities for forecast bias – where well-intentioned forecasters or leadership may skew the forecast based on personal expectations.

What is algorithmic forecasting, and why does everyone want it?

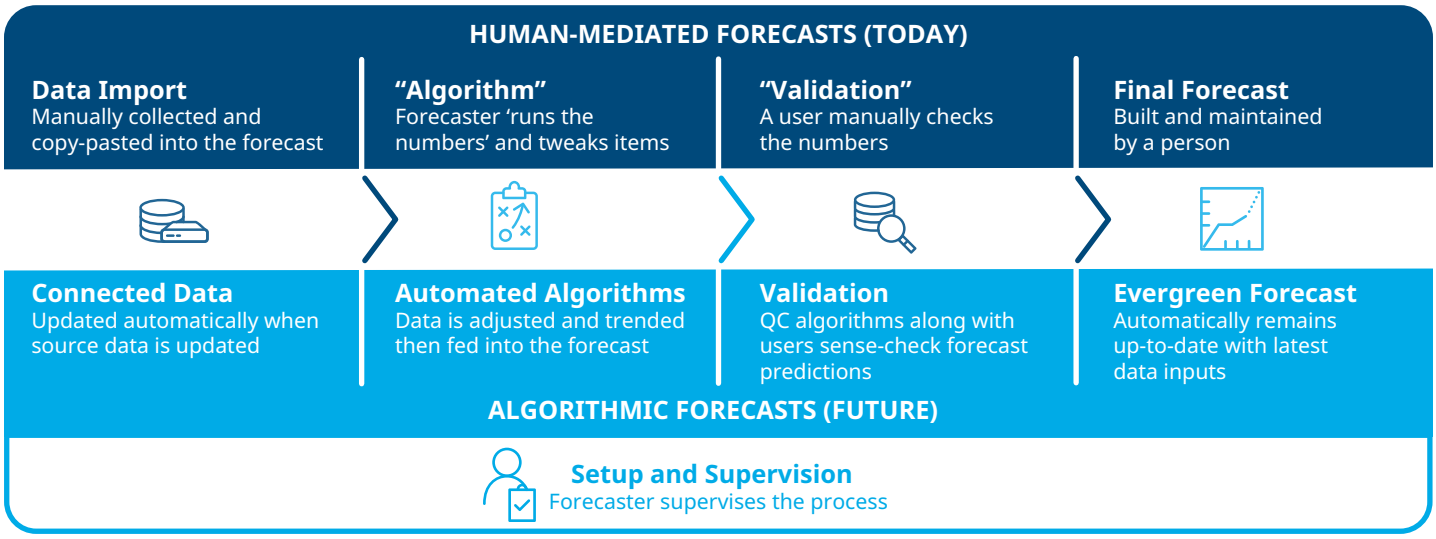
Algorithmic forecasting provides a way to address many of these challenges.

In algorithmic forecasting, a combination of analytics technology and machine learning algorithms replaces most of the manual labor, data chasing and cut-and-paste activities common in a traditional forecast workflow.

Sources of relevant data are directly connected to the forecast engine, allowing the applied algorithms to update the forecast in real-time as new data becomes available. A combination of human users and quality control algorithms validate the auto-generated forecast to guard against errors in the process.

A key element of algorithmic forecasting is that it is not human-mediated.

Forecasters set up and supervise the process, but they are not actively involved in scouring the data and making predictions. That is all up to the algorithm(s), which search the data for meaningful trends, and generate predictions in a fraction of the time it takes to conduct a human review.



Whenever new data is added, or updated, the system automatically re-calculates the forecast. With robust version-control built in, forecasters can track and compare forecast results, and also face a reduced need for a complex submission process. In this environment, the forecast remains 'evergreen' – universally and continually relevant, rather than updated only once each month or quarter.

> *"Evergreening" – where the most accurate forecast is always available, rather than updated once each quarter – would be revolutionary for commercial pharma forecasts. And algorithmic forecasting is the way to do it."*
>
> **RICK JOHNSTON, PhD, SENIOR PRINCIPAL**

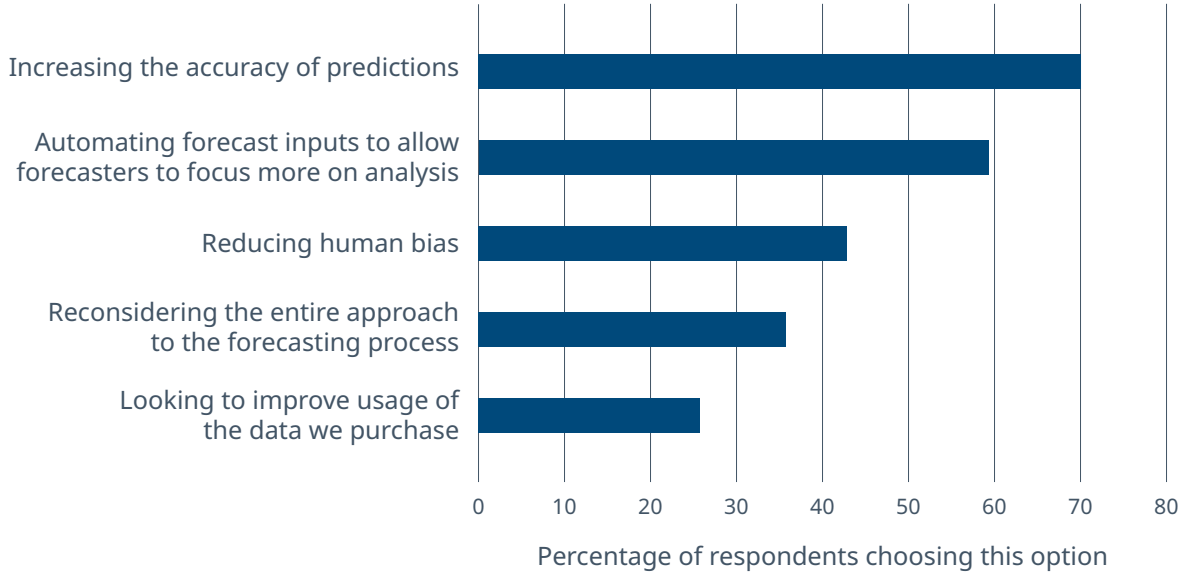
What's in it for forecasters?

With these manual steps becoming more automated, the forecaster is elevated to a more strategic role. Because the forecast inputs are automatically updated, forecasters can instead focus on exploring trends in the data, conducting what-if or uncertainty analysis, identifying key drivers or indicators and developing strategic insights.

These benefits ultimately lead to more accurate forecasts, which is the leading reason why forecasters are interested in algorithmic forecasting, according to a recent survey conducted by IQVIA.

And because the algorithm has no stake in the output, it reduces or removes human bias, which was also ranked among the top three reasons why forecasters are interested in incorporating algorithmic approaches in their forecasts. This real-time capability transforms the forecast from a quarterly static snapshot to a constantly evolving review of trends as they are occurring.

Why are your teams interested in algorithmic forecasting?



Because the forecast inputs are automatically updated, forecasters can instead focus on exploring trends in the data, conducting what-if or uncertainty analysis, identifying key drivers or indicators and developing strategic insights.

Case Study: Excel Falls Short Predicting Biosimilar Trends

A large pharmaceutical company wanted to track the erosion of sales volumes related to the release of several biosimilar products in the previous three years. Sales reps were seeing unusual trends in the market, characterized by some hospitals continuing to prescribe the innovator product while others shifted to the biosimilar.

The company's forecast engine was an Excel spreadsheet that could only project trends in 40 market segments. It wasn't detailed enough to "tease out" patterns of behavior that were occurring. As a result, the company was unable to determine which of its marketing and sales strategies were having the desired impact, and what was causing some customers to switch.

THE SOLUTION

In response, the company deployed IQVIA's Forecast Horizon to generate a more sophisticated, real-time market forecast. With its algorithmic forecasting platform, they were able to analyze data from tens of thousands of individual providers and match it to payer-level data, thereby determining what percentage of the innovator molecule was prescribed versus the competitors. Algorithms were programmed to automatically segment the providers into 'similar groups,' and to aggregate their sales data and remove outliers. The result was a model split into 2000 segments of hospitals with similar patterns. The depth and detail provided through this forecast gave them significantly more predictive results in a fraction of the time.

THE RESULT

A 50x increase in the number of segments made it possible for an incredibly granular view of the behavior of similar hospital groups. With aggregated data illustrating overall patterns by provider, site of care, individual hospital network, and payer trends, forecasters were able to build more accurate forecasts. Every time new data was added to the data lake, the forecast automatically updated to re-calculate the impact, with no human involvement required.

The company used the forecast to make significant changes to its sales and marketing strategy, providing actionable tips to teams that reduced overall attrition for the brand, allowing it to retain a significant market share despite strong competition.

Algorithmic forecasting: reality vs. hype

While life sciences industry vendors have been touting the "wonders" of automated solutions for years, algorithmic forecasting has remained an elusive target for pharmaceutical companies. That has caused some pharmaceutical companies to grow skeptical of the hype.

Part of the problem is that a lot of forecasting applications on the market today are built by companies that try to adapt a forecasting solution from another industry to meet the needs of pharma. But pharma commercial forecasts are different from other markets. The average pharma forecast we encounter has more than 80 variables being input and calculated, while the typical consumer goods forecast has less than 10. The main reason for this is the need to model the complex patient dynamics of most modern medicines: periods of therapy, patients dropping off therapy for a variety of reasons at different times during their therapy, complex dosing patterns, treatment line changes, flow of patients between lines of therapy, and others.

The forecasting tool also has to be able to account for local regulations, emerging healthcare events, and other issues that are unique to this industry.

Practical algorithmic forecasting solutions today are built around three core components:

1. **An ecosystem of data handling tools** that can pull real-time data feeds from a variety of sources and employ data management methods to prepare and combine the data for reliable, robust analysis. In many cases these re-use data infrastructure developed for other analytics and insights tasks within the business.
2. **Automated algorithms using self-service analytics** that have been trained using vast global healthcare databases to identify trends of relevance and provide insights that drive decision-making. Critical to these algorithms is the ability for forecasters to make edits themselves using self-service analytics tools.
3. **Real-time forecast workflows** that allow teams to conduct frequent updates, ask follow-up questions, and model many detailed scenarios to enhance understanding of key trends and monitor changes over time.

This combination of elements makes it possible for forecasters to enhance the quality, detail and reliability of their forecasts, and to make forecasting a more active part of the strategic decision-making process.

1. IT'S ALL ABOUT THE DATA: AN ECOSYSTEM OF DATA HANDLING TOOLS

The first enabling step for teams to implement algorithmic forecasting is to build a data 'landscape' that supports real-time data feeds in a format that can be understood by a forecasting platform. These data handling tools are not just a database or data lake, but include capabilities to:

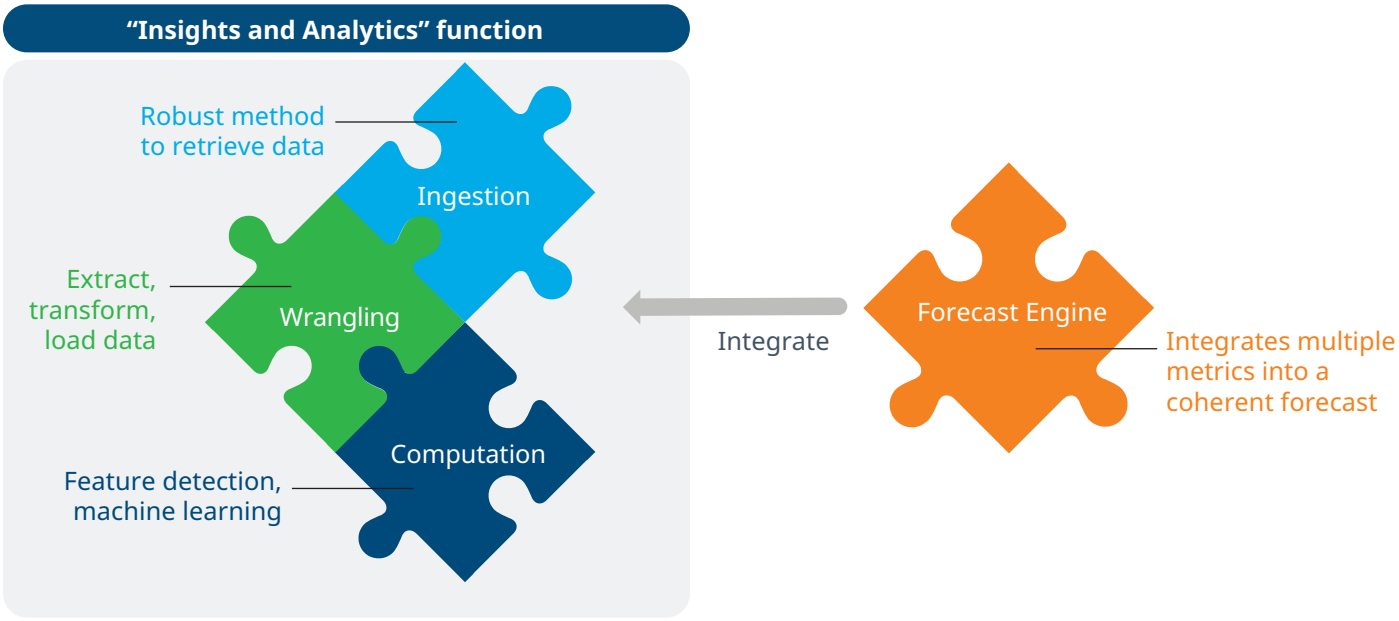
- **Join, filter, and smooth data:** Data handling tools need to be able to join data from disparate data sets, remove outliers and errors, and perform data smoothing tasks. These are commonly considered 'pre-processing' steps that are performed before passing data to a forecasting platform. Examples of algorithms that might be applied include isolation forests, minimum covariance determinate, local outlier factor methods (outliers) and Friedman super smoothing.
- **Save high-fidelity data and allow aggregations:** While time series data is often saved at a daily or weekly granularity, most forecasts operate at a monthly or yearly granularity, so data is usually saved in the more detailed form. Additionally, the ability to aggregate data across segments, such as individual hospital data across a zip code, city, state or country, is required. Pivot table type functionality here can help to aggregate the data to a format that provides a consistent segmentation along with time series data at the right level.
- **Experiment with different approaches:** Data should be available in a form that allows users to quickly see and perform custom queries with it. These tools are typically workbenches that enable data science teams to use open source components like Python and R to do complex data transformations, and then save those scripts to run each time new data is available.
- **Build automation steps for data updates:** To support the algorithmic forecast, data will be coming in at regular intervals. After data has first been identified and transformed, the infrastructure needs to be able to perform a robust scheduled process to push the analytics to the forecast.

Most pharma companies today will already have the capabilities within their data analytics teams to handle many of these tasks. A rapidly evolving ecosystem of tools is also increasingly able to automate many of these data handling steps.

2. AUTOMATED FORECAST ALGORITHMS: THE SELF-SERVICE ANALYTICS REVOLUTION

While Excel has been the backbone of forecast technologies for decades, it is generally insufficient by itself for algorithmic forecasting. There are issues with it having limited flexibility, challenging version control, and sluggish performance on large models that creates headaches for forecasters and reduces the agility and speed of their forecasts. The evolution of algorithmic forecasting can provide far more sophisticated solutions.

IQVIA has managed forecast processes and built many of our own Excel forecast platforms over the years. We have learned that to enable algorithmic forecasting, a new type of forecast enterprise architecture is required.



A key part of that technology is to think about the way that open-source tools like Python and R – as well as the ecosystem of data handling tools around them – can be integrated into the forecast process. IQVIA's approach has been to build a platform that allows users to use their own insights and analytics tools, but to allow the 'hooks' to a forecast engine that brings all those computations together into one user interface.

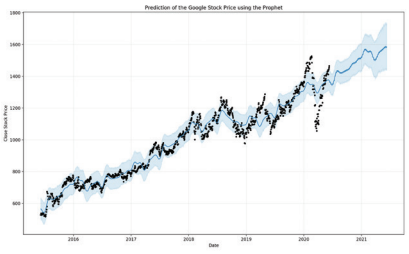
There are a whole host of different algorithms that can be used to process forecast data depending on the type and detail of the data, objective, and capabilities of the algorithm. We list a few categories of algorithms below, but there are also many different publicly available algorithms that can be tuned to your data set.

IQVIA's approach has been to build a platform that allows users to use their own insights and analytics tools, but to allow the 'hooks' to a forecast engine that brings all those computations together into one user interface.

THREE TYPES OF ALGORITHMS TO CONSIDER

Time Series Predictions

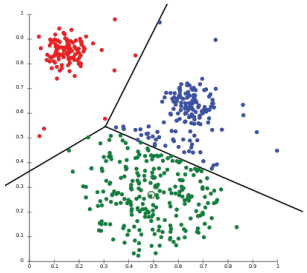
Machine Learning is creating more accurate methods for trending historical data



Example: Prophet, XGBoost

Clustering and Partitioning

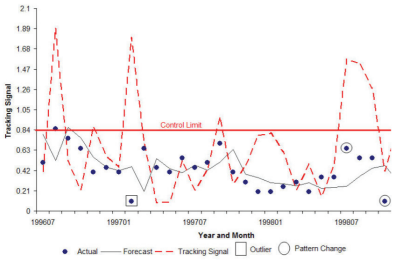
Important for identification of analogues and other 'similarities within groups'



Example: K Means, KNN

Forecast Quality and Tracking

Methods to detect when the algorithm is failing and a person should intervene to see what's happening



Example: Trigg's Tracking Signal

There are three types of algorithms that can deliver early wins for life sciences forecasts:

Time Series Prediction algorithms. In these cases, historical data is used to predict future performance. Algorithms like ARIMA and Holt-Winters are rapidly being replaced by more sophisticated methods that use machine learning as well as more categoric variables to provide better future predictions.

Clustering and Partitioning algorithms. These are suitable for patient identification, market segmentation and analogue selection. In the case of patient identification, these algorithms can be used to identify undiagnosed patients based on similar patient diagnoses or patient journeys. They can also be used for analogue selection, allowing similar products to be found in the huge universe of available marketed products based on a set of features.

Forecast Quality and Tracking algorithms. Once an automated algorithm has been established, it is critical to ensure that the outputs are checked each time new data is received. These quality checks allow forecasters to know if their forecast algorithm is still producing a good result or needs manual intervention.

Sources:
https://towardsdatascience.com/time-series-forecasting-predicting-stock-prices-using-facebooks-prophet-model-9ee1657132b5
https://aws.amazon.com/blogs/machine-learning/k-means-clustering-with-amazon-sagemaker/
https://pdfs.semanticscholar.org/c4e6/6029e5173d25a28be7fc9b811d96c675b9e3.pdf

Our focus so far has been on the infrastructure and technology needed to enable algorithmic forecasting. However, even the best technology can't work unless users commit to changing their workflows and processes to make it part of how they operate. That's a change management issue, not a tech issue.

3. COLLABORATION IS KEY: A REAL-TIME FORECAST WORKFLOW

The primary challenge for many organizations attempting to embrace algorithmic forecasting is siloed responsibilities. Today's forecast processes – the whole system of submissions and reviews – makes it difficult to work cross-functionally in the way that is necessary to implement algorithmic forecasting. In the same vein, there is a lot of momentum behind those processes as they cross many groups in the organization.

The shift to algorithmic forecasting requires re-imagining these processes for the age of multi-user editing and real-time data. Commercial forecasters need to learn 'the basics' of these analytics tools and how they might be used, while insights and analytics teams need to learn to apply their toolset to practical real-world forecasting problems. That includes determining how data can automatically update a forecast without explicit approvals, and how it can automatically check the errors and flag any issues which require human intervention.

The forecast is at the center of a collaborative effort

In the collaborative process

- The forecast is always live and can be updated by any user in real-time
- The brand forecaster "orchestrates" the changes and versions to ensure accuracy
- The mechanics of forecasting are mostly/entirely automated
- The latest numbers are always available, as well as a history of changes
- No formal 'submission' is needed
- Teams have high visibility into the forecast framework, assumptions and processes

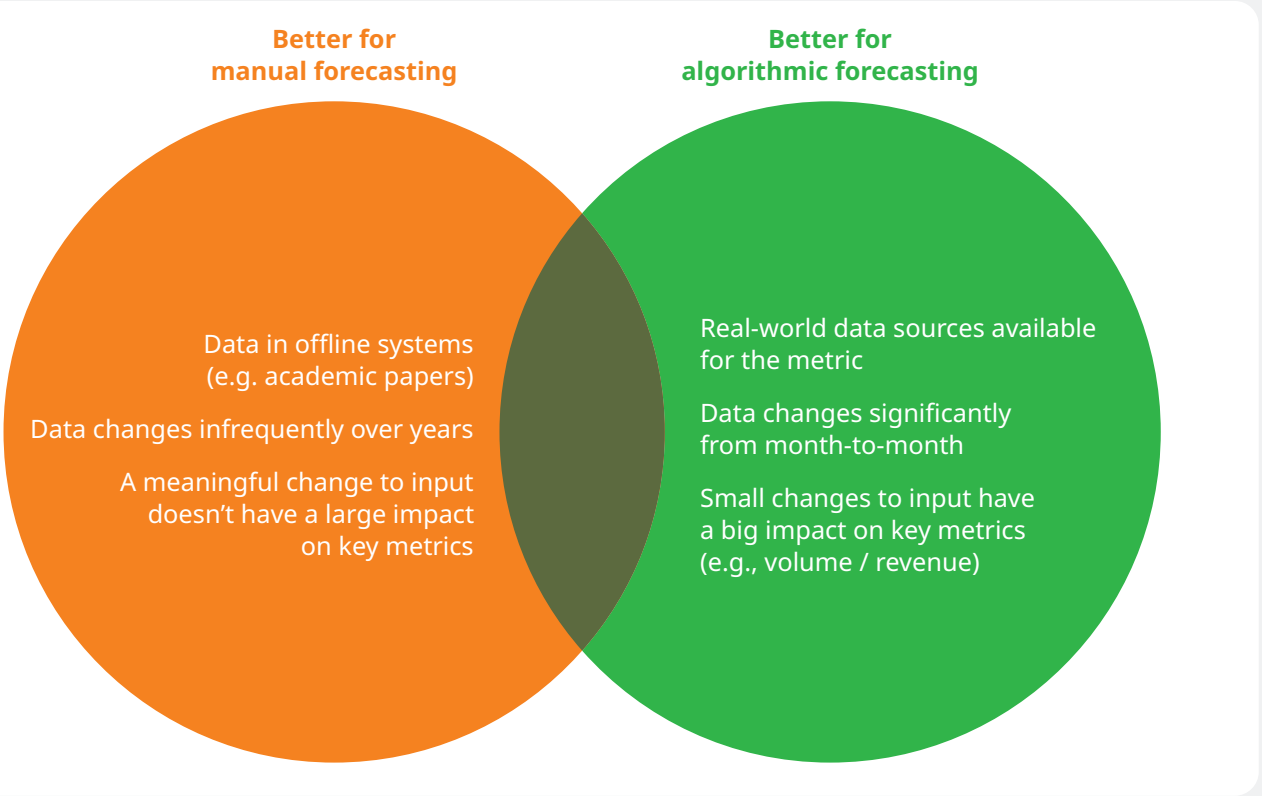
All of this technology already exists today, but the key is to make the business process changes required to make this collaborative environment a reality. That means doing away with cumbersome submission technologies, versioning systems, review processes and other paradigms designed for the age of Excel. Instead, a forecaster becomes the conductor of a team of people all providing forecast data and analytics in real-time, 'orchestrating' the entire process from end-to-end.

The results of this type of process speak for themselves – eliminating differences between the 'official' and the 'latest' versions of the forecast, enabling all team members to access one source of truth which is always kept up to date with the latest numbers, and removing sources of bias introduced by different groups during the hand-off process.

Where do I begin? The algorithm journey

Algorithmic forecasting will not be a simple process for most companies. IQVIA recommends an incremental approach where parts of the forecast funnel are transferred to algorithmic-type approaches, while others are still manually forecast. This approach focuses algorithmic forecasting on areas where it delivers the most value.

HOW DO WE DETERMINE IF A FORECAST INPUT SHOULD BE ALGORITHMIC OR MANUAL?



Algorithmic approaches are favored for inputs where:

- Detailed real-world data is available and regularly updated
- Data is changing significantly from month-to-month
- A small change in that input would cause a significant change to forecast outputs, such as net revenue or total volume

Common examples include market share, total volume, diagnosed or treated patients.

Prioritization of those metrics with greatest impact can be informed with a tornado-style analysis that looks at potential ranges for those variables and the effect on net revenue.

Many forecasters ease into algorithmic forecasting by taking a 'manual-lite' approach, where algorithms suggest a solution, but confirmation is required from the forecaster to actually apply the algorithm in each case. For example, the algorithm may suggest a 'naïve' forward-looking trend to the user – lowering the chance of bias in their estimations – but it would still be up to the forecaster to accept the recommendation. Analytics-assisted inputs like these still deliver many of the benefits of algorithmic forecasting, without total automation.

The complexities of forecasting in the healthcare space, as well as the importance of accuracy means pharma companies may struggle to find forecasting technology capable of meeting their needs. That's exactly what IQVIA's Forecast Horizon and Pipeline Architect platforms were built to do.

These two solutions were built by experts in data analytics, global healthcare regulations, and digital technology to ensure the platforms could meet the exact needs of pharmaceutical industry stakeholders across the globe.

IQVIA's **Pipeline Architect**[1] platform, designed for early stage products, uses algorithmic forecasting to pull information on hundreds of trials from IQVIA's global databases and combine that with public data to predict an asset's development costs and timelines, as well as to inform high-level development timing and commercial expectations. These insights lead to better forecast development and more accurate timeline predictions.

IQVIA's **Forecast Horizon**[2]**,** designed for later stage or inline products, is a commercial forecasting tool which can be used to build an algorithmic forecast. The platform can pull in data and from clinical and commercial data resources and apply out-of-the box or custom-built algorithms to inform forecast inputs.

Users can specify complex patient dynamics such as persistence (time on therapy) and flow between lines of therapy, reflect the patient journey across different demographics, and compare the performance of different products across treatment categories. These automated insights enhance the depth, accuracy, and precision of forecasts, while minimizing time spent on manual updates.

Both platforms are helping life sciences companies across the industry navigate forecasting challenges and allowing them to evolve their approach to real-time, evergreen forecasts.

Conclusion

Commercial forecasters are looking to embrace algorithmic forecasting as a method to achieve better, faster, and more accurate results.

However, many life sciences companies are uncertain about where to begin. The good news is that it isn't as hard as it sounds.

With this framework, forecasters can begin discussions to put together their data infrastructure, determine what algorithms suit their objectives, and begin discussing the potential benefits of a real-time forecasting process. And even if they are not ready to jump in fully, there are ways to benefit from algorithmic forecasting, like automating a smaller number of inputs or requiring manual approval, which can be put in place today.

1 https://www.iqvia.com/solutions/research-and-development/consulting/pipeline-architect
2 https://www.iqvia.com/solutions/commercialization/brand-strategy-and-management/forecast-horizon

About the authors

RICK JOHNSTON, PhD
Senior Principal,
Consulting Services, IQVIA

Dr. Johnston is a Senior Principal and Software Lead within IQVIA's Consulting Organization. He is an accomplished technical leader with more than 15 years' experience in building software that solves big data, machine learning and advanced analytics problems in pharma. Software he built is used by more than 90% of the world's 25 largest pharmaceutical companies. Dr. Johnston has a PhD in operations research and bioinformatics from UC Berkeley.

ALEXANDRA TATARU, MS, MBA
Product Lead,
Consulting Services, IQVIA

Alexandra is a Product Lead for IQVIA's Consulting Services team. She focuses on the blend of technical and business concepts within the life sciences space, with over 9 years of experience in healthcare across a mix of product management, management consulting, and engineering. Alexandra holds a BS and MS from Stanford University in Mechanical Engineering as well as an MBA from Columbia University.

CONTACT US

ForecastHorizon@iqvia.com

iqvia.com