Modeling Regulatory Impacts on Medical Device Supply Chains
by
Melissa Medina
Bachelor of Business Administration, Supply Chain Management, Marketing, International Business, University of Oklahoma, 2010

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ENGINEERING IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUNE 2018

© 2018 Melissa Medina. All rights reserved.
The authors hereby grant to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature redacted

Signature redacted

Certified by..................
Dr. Bruce C. Arntzen
Executive Director, Supply Chain Management Residential Program
Thesis Advisor

Accepted by.............
Dr. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering
Modeling Regulatory Impacts on Medical Device Supply Chains
by
Melissa Medina
Submitted to the Program in Supply Chain Management on May 11, 2018 in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Supply Chain Management

ABSTRACT

Changing regulatory requirements continues to be an increasingly complex issue in the medical device industry. Regulations place stress on regional supply chains across the world. Most recently, the European Parliament issued the Medical Device Regulation (EU) 2017/745 instituting new compliance framework for all devices manufactured, sold, and/or distributed in the European Union. The new framework requires the implementation of unique device identifiers and more stringent conformity assessment procedures. In addition, many device classification types have changed, post-market clinical surveillance has been instituted, and traceability through a centralized IT database is now mandated. While the the act aims to improve patient safety and efficacy across the medical device industry, it poses huge impacts across both the physical and informational flows in supply chains. This research evaluates the regulatory impact across supply chain operations using predictive modeling and machine learning. The model determines how various activities and events in manufacturing and sourcing environments contribute to supply constraints when modified to accommodate new regulatory requirements. The model also determines how product attributes contribute to performance variability. By taking a proactive approach to assess the impacts of regulatory changes, firms can optimize supply chain flows to reduce cost, lead-time, and service level risks.
ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Bruce Arntzen, for providing his guidance and knowledge throughout this research journey. I would also like to thank Pamela Siska for her writing advising, as well as Connor Makowski and Chris Cassa for their analytical input. This thesis would not be possible without the support provided by Jere Butler, the Director of Global Supply Chain Planning and Metrics at the medical device company which sponsored and supported this research.
# Table of Contents

1. **INTRODUCTION** ................................................................................................................................. 5
2. **ISSUE** .................................................................................................................................................. 5
3. **RESEARCH PURPOSE** ......................................................................................................................... 6
4. **LITERATURE REVIEW** .......................................................................................................................... 7
   4.1 **INDUSTRY OVERVIEW** .................................................................................................................... 7
   4.1.1 **Industry Trend- Value Added Services** ......................................................................................... 7
   4.1.2 **Industry Trend- Technology** ......................................................................................................... 7
   4.1.3 **Industry Trend- Acquisitions and Mergers** .................................................................................. 8
4.2 **HISTORY OF REGULATION** ............................................................................................................. 9
4.3 **HISTORY OF PREDICTIVE ANALYTICS IN SUPPLY CHAIN** ......................................................... 9
   4.3.1 **Machine Learning** ...................................................................................................................... 9
5. **DATA AND METHODOLOGY** .................................................................................................................. 10
   5.1 **Data Gathering** .................................................................................................................................. 13
6. **MODEL** .................................................................................................................................................. 15
   6.1 **Model Overview** ................................................................................................................................ 15
   6.1.1 **Program Selection** ...................................................................................................................... 16
   6.1.2 **Technique in Detail** .................................................................................................................... 16
7. **RESULTS AND APPLICATIONS** ............................................................................................................ 18
   7.1 **Results** ............................................................................................................................................ 18
   7.2 **Application** ...................................................................................................................................... 20
8. **CONCLUSION AND NEXT STEPS** ........................................................................................................... 26
9. **REFERENCES** ....................................................................................................................................... 27
10. **APPENDIX** ......................................................................................................................................... ERROR! BOOKMARK NOT DEFINED.
1. Introduction

The medical device industry has been burdened with increasing amounts of regulation for decades. While regulation helps to ensure patient safety and efficacy- it has also created significant burden on the supply chain. In order to stay compliant, manufacturers and distributors have had to reconsider network design. This in turn raises cost and increases product lead times.

Most medical device technologies are regulated differently across the globe. The requirements are disparate and inconsistent by region. As such, healthcare suppliers engage in continuous efforts to produce and distribute high quality products at all levels of conformity. This has created disruption and change in the current operational flows. To stay competitive, medical device supply chains must approach safety, quality, and compliance more and more strategically.

2. Issue

In April of 2017 the European Union implemented the Medical Device Directive, (MDD). This legislation represents the largest regulatory change to the European medical device industry in thirty years. The directive implements additional requirements and makes current requirements much more stringent. The MDD will increase the burden of compliance through more stringent conformity assessment procedures, additional clinical data requirements and more extensive post-market surveillance. The EU Commission, Notified Bodies, and the industry will need to invest substantial resources to meet the three-year transition deadline. (BMI, 2017)
The MDD extends legislation beyond requirements on the manufacturer to other economic operators including distributors and importers. Meaning, more players must be more compliant- and these activities must now coordinate across the entire supply chain network.

The MDD also initiates changes to device classification types and the required level of documentation. In addition to more documentation, traceability and tracking with unique device identifiers is now compulsory for most products. Operations must also include additional labeling and packaging in current product portfolios. Each medical device will need to include a patient card and Instructions for Use.

The current structure of supply chains will need to change in order to accommodate these new policies. This will require time, money, and a new strategy for optimizing network flows. Firms will need to develop new processes and systems capabilities. As a result, decentralized systems, forward stocking locations, regional distribution centers, and dual sourcing strategies will all need to be reconsidered. Understanding how these changes will impact service levels, revenue, and cost is a difficult process.

3. Research Purpose

As described briefly in the previous section, the MDD presents a number of challenges that are not clearly outlined, or have been left to medical device industry players to interpret and enact. Decisions will need to be made regarding how to design the new physical and information flows in the supply chain. The purpose of this research is to:

1. Explore different approaches to model the impact of regulatory changes on global supply chains
2. From an industry perspective, develop a model that can predict the impact of regulations on supply chains with different attributes and events

4. Literature Review

4.1 Industry Overview

The availability of medical device technology has exploded over the past century. Healthcare is one of the few industries that continues to grow year over year, independently of economic conditions. Advancements in technology, and aging population, and the growth of the middle class in emerging markets has created a boon in demand and innovation. The global market for medical device technologies reached $521.2 billion in 2017 and should reach $674.5 billion by 2022 (BCC, 2017).

4.1.1 Industry Trend-Value Added Services

While the medical device industry grows, customer demand is shifting toward value delivering products and services. Rising costs are placing a downward pressure on industry profit. (Ibis, 2017) Hospitals now want less inventory, lower cost, and new products. As a result, firms are revisiting design and investments in the supply chain. Firms are increasingly outsourcing manufacturing functions, investing in automation, and reducing investments in R&D.

4.1.2 Industry Trend- Technology

The healthcare industry has become increasingly dependent upon advancements in technology to improve efficiency and reduce cost (BCC, 2017). As a result, healthcare supply chains have been progressively adopting new technologies for decades. A few notable implementations include robotics for manufacturing, RFID for DC’s, and 3D printing...
for production. This analysis leverages the technology of predictive analytics to identify the impact of regulatory activities in the supply chain.

### 4.1.3 Industry Trend- Acquisitions and Mergers

A common trend in this industry is growth through acquisition. Changing parent-subsidiary relationships creates dis-synergies in manufacturing and procurement. Rarely in this business are systems, processes, and networks optimized. In fact, an archetypal medical device supply chain has multiple transformation projects happening at any given time. Most Supply Chain Managers in medical device will acknowledge that the typical supply network is a cluster of complex flows.

An average medical device portfolio in this industry contains thousands of products ranging from consumables to implants and instruments. Larger players typically manage production internally. However, high investment and non-strategic manufacturing operations are done by outside vendors. Depending on the device type this could include metal finishing, anodization, surface coating, sterilization, assembly, and packaging. Highly specialized vendors keep sourcing networks disparate, and changes are often difficult to make due to rigid quality requirements. Optimizing or even consolidating production locations is difficult. New regulatory filings are required for “manufacturer of origin” changes. Approvals can sometimes take years to obtain. Add rampant SKU proliferation into the mix, and the supply chain becomes a difficult web to navigate. These diverse nodes and flows are the subject of this research.
4.2 History of Regulation

Since 1970, regulating medical devices and establishing government oversight has been an ongoing topic of conversation. The US Medical Device Amendment to the Food, Drug & Cosmetic (FD&C) Act marked the beginning of comprehensive device regulation around the world. Since then, major regions around the world, like the European Union, have followed suit.

The primary function of regulation is to ensure safety and efficacy. As technology evolves, so does legislation. Existing directives are reviewed regularly at the administrative level, and new amendments are published as needed. Oftentimes the changes require major business process redesigns in order for a firm to be compliant. It is not uncommon for a company to modify operations multiple times as a result.

4.3 History of Predictive Analytics in Supply Chain

Predictive analytics have gained a significant following in recent years. Improvements in data availability, access, and computational power has created entirely new capabilities. These techniques can draw inferences from large data sets. They have the capability to find hidden patterns and correlations between variables.

4.3.1 Machine Learning

Machine Learning is a specialized sub-field of Artificial Intelligence where algorithms can learn and improve themselves by studying high volumes of available data. This type of application is valuable in areas where there are a lot of variables and variable interactions. Because data mining is such a fast growing application in business, companies are leveraging these techniques to gain competitive advantage. Journals and publications consistently list
areas such as fraud, security, trading, forecasting, and consumer preference tracking as top AI business applications.

Predictive analytics are increasingly being used in front end supply chains for forecasting trends and changes in consumer behavior. However, back end Supply Chain applications have long been overlooked. This thesis applies machine learning to backend activities in the supply chain to identify relationships behind backorders.

5. Data and Methodology

This analysis utilizes twelve months of historical SKU level data from a medium sized orthopedics company. The portfolio includes surgical solutions for upper extremities, lower extremities, and biological products. Included in the analysis are 3,250 items ranging in type from implants to instruments to consumables. The size and features of implants range from very small to very large, packaged and kitted, and sterilized and unsterilized. Product type also ranges from plates and screws to injectables, grafts, and wedges. Figure 1 illustrates the different physical attributes of each device.
Figure 1. Various medical devices, each with different shape, size, weight, and application

Instruments are either single use or reusable with sterilization. A few illustrations of sizers, reamers, curettes, and drills are available in Figure 2.
Products are either sourced externally as finished goods from vendors, internally through in-house manufacturing, or some combination of both. For example, Figure 3, shows an implant that could be produced in house or sent to a vendor. The vendor may perform a specialized service such as colored anodization, enabling it to be easily distinguished by a physician. The part may then be sent to another location, internally or externally for final packaging.

Some products are stocked in centralized pooled global distribution centers, while others are stocked via both regional distribution centers and small forward stocking.
locations. Almost all products are simultaneously present in hospital consignment locations and/or carried by reps in their trunk stock.

Hardware is used in surgical kits and sterilized after individual procedures, while biologics are mostly single use. An example of a surgical kit is provided in Figure 4. Most products have a single regulatory approved location filed as the official “manufacturer of origin” and is globally exported to other international markets. A product’s manufacturing assembly could range from a simple single raw material component to upwards of thirty raw material components. Both sterile and non sterile finished goods exist. Single and bulk pack SKU’s are stocked.

Figure 4. Example of surgical kit

5.1 Data Gathering

This analysis gathers the following supply chain attributes and performance outcomes as input variables for the model. Included at the SKU level is twelve months of actual historical sales and master data tables. This includes:

ITEM NUMBER- refers to the unique product identification number also referred to as a Stock Keeping Unit or SKU
ITEM CLASSIFICATION- refers to the ABC Classification method indicating whether a SKU is low to high volume, low to high frequency, and low to high value (where value is a function of Cost of Goods Sold [COGS])

FORECAST DEMAND- refers to the monthly forecasted demand for the current month plus one, plus two, and plus three

MANUFACTURING CELL- refers to the machine and series of work order operations that a product goes through in the production facility. This includes: Receiving & Inspection, Cleaning, Packaging, Sterilization, etc.

SALES VOLUME- refers to the percent of sales the specific SKU generates relative to the other SKU’s in its particular size category

ITEM MONTHLY DEMAND- refers to the average unit quantity of sales demand this SKU has generated when analyzed by 12 months of historical sales actuals

ITEM INVENTORY FINISHED GOODS - refers to the finished good item average quantity available on the “shelf”

ITEM INVENTORY SEMI-FINISHED GOODS- refers to the semi-finished good item average quantity available in Work in Progress (WIP) on either Work Orders (WO) or Purchase Orders (PO)

DISTRIBUTOR DEMAND- refers to the sales channel demand type in emerging markets; a binary variable indicating whether the item is made available to be sold and exported in high quantity volumes

INSTANCES OF BACKORDERS- refers to the number of times an item was on backorder in the last twelve months of historical sales actuals
AVERAGE LEADTIME- refers to the average lead time measured in days from the start of the WO or PO until the item is put to shelf

MAKE VS BUY- refers to item the source of supply i.e. external procurement or internal production

MANUFACTURING OPERATION- refers to the machine or work cell which an item is processed through from the start of the WO or PO until the item is put to shelf

MANUFACTURING PLANT LOCATION- refers to the production origin of an item as raw materials are processed and transformed into finished goods at the different facilities.

Facilities included in this data set are Ireland, France, and US

6. Model

6.1 Model Overview

This model utilizes two types of analytical techniques to find relationships in supply chain flows and backorders. The first method, Supervised Learning, looks at those supply chain flows with regard to backorders, and produces an inferred function showing how the item features and performance outcomes are impacting service level. The second method, Unsupervised Learning, also called "Machine Learning," looks at the entire data set without a pre-specified target variable. It then finds hidden patterns and relationships within the data set. Essentially, these models are finding veiled relationships, modeling how they impact backorders, and then predicting whether service level will increase or decrease when a regulatory change is made.
6.1.1 Program Selection

The program used for this analysis is Orange Data Mining and Machine Learning Tools. This open source software platform provides an interface for data analysis workflow. Functionalities include: selecting features, training predictors, comparing learning algorithms, and visualizing data elements. Widget tools are grouped together into six sets for each type and step of modeling. See Figure 5 Orange Tool Capabilities.

<table>
<thead>
<tr>
<th>DATA</th>
<th>VISUALIZATION</th>
<th>CLASSIFICATION</th>
<th>REGRESSION</th>
<th>EVALUATION</th>
<th>CLUSTERING</th>
</tr>
</thead>
<tbody>
<tr>
<td>data input, data filtering, sampling, imputation, feature manipulation, feature selection</td>
<td>box plot, histograms, scatter plot, mosaic display, sieve diagram</td>
<td>random forest, neural network, nearest neighbor, decision trees, random forest, support vector machines</td>
<td>logistic regression, linear regression, ordinary least squares, multivariate</td>
<td>cross-validation, sampling-based procedures, reliability estimation and scoring of prediction methods</td>
<td>k-means, hierarchical, multidimensional scaling, principal component analysis, correspondence analysis</td>
</tr>
</tbody>
</table>

Figure 5. Capabilities of Orange Data Mining and Machine Learning Tools

6.1.2 Technique in Detail

The first Supervised Learning technique applied in this model is Linear Regression. Linear Regression looks at a set of \( n \) inputs (independent variables) and predicts a specific outcome \( y \) (the dependent variable). A general form of a Linear Regression equation is shown in Equation 1.

\[
y = c + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n
\]

Where:

- \( y \) = predicted dependent variable
- \( c \) = a constant
- \( \beta_n \) = regression coefficient of independent variable \( x_n \)
- \( x_n \) = independent variable \( n \)

Equation 1. General form of Linear Regression equation.
In the case of this analysis, the dependent variable $y$ is the service level (i.e. number of instances an item goes on backorder in twelve months). The independent variables $x_n$ are the attributes and of the operations within the medical device supply chain. Linear regression, upon analyzing the historical data inputs, will output a function for future predictions.

The second Supervised Learning technique applied in this model is Logistic Regression. Logistic Regression is used to predict binary outcomes (e.g. 0 or 1, YES/NO, etc.) Logistic regression predicts the probability of an independent variable belonging to a certain class. The general equation for logistic regression is given in Equation 2.

$$ logit(p) = c + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n $$

Where:

- $logit(p) = \ln \left( \frac{p}{1-p} \right)$
- $p = probability \ of \ an \ event \ occurring \ in \ twelve \ months$
- $c = a \ constant$
- $\beta_n = regression \ coefficient \ of \ independent \ variable \ x_n$
- $x_n = independent \ variable \ n$

Equation 2. General form of the logistic regression equation.

The main Unsupervised Learning technique applied in this model is Primary Component Analysis (PCA). PCA converts the set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables. In other words, when a lot of data is available, PCA reduces the dimensionality to only those which are “most informative.” These are the features which account for most of the variability.

The output of the PCA is a transformed dataset with weights of individual instances for each feature, or weights of combinations of those features. Since this is a feature extraction technique, it combines the input variables in a specific way providing the ability to
drop the “least important” variables while still retaining the most valuable parts of all of the variables.

In summary, PCA does 2 things:

1. Reduces the number of variables to the “most informative”
2. Ensures variables are independent of each other

7. Results and Applications

7.1 Results

Per the Linear Regression analysis, given the current performance outcomes and item attributes, the number of instances a SKU will go on backorder is predictable with an $R^2$ of 22.4%. The performance is indicated by the Coefficient of Determination ($R^2$), which measures how close the data are fitted to the regression line. Given the current set of data, realistically predicting the number of instances a SKU will go on backorder with a good level of certainty is not likely. Refer to Figure 6 for the results of the Linear Regression.

<table>
<thead>
<tr>
<th>Evaluation Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
</tr>
<tr>
<td>Linear Regression</td>
</tr>
</tbody>
</table>

Figure 6. Results from Linear Regression of the backorder data.

Since this backorder data was recorded weekly for twelve months the number of possible instances any given SKU could backorder is up to fifty-two times. Predicting any single correct answer out of fifty-two possibilities is would be difficult without significantly more data. However, by switching from linear regression to logistic regression, and transforming the number of instances a SKU will go on backorder to a binary, categorical variable representing a “will or will not go on backorder” prediction, the number of possible
results reduces to two. The odds of forecasting accurately between two options is much
greater than fifty-two options. In modifying the model in this way, and running logistic
regression improves forecast accuracy from 22% to 82.9%. Performance in this case is
indicated by Area Under the Curve (AUC).

Given the current set of data, making a general prediction of whether or not a SKU
will go on backorder, with a good level of certainty, is possible. Using this classification
technique, the prediction would be accurate approximately 78% of the time. Refer to Figure
7.

<table>
<thead>
<tr>
<th>Evaluation Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
</tr>
<tr>
<td>Logistic Regression</td>
</tr>
</tbody>
</table>

Figure 7. Results using Logistic Regression, where the performance is based on the Area Under the Curve (AUC)

While 83% accuracy for predicting a "class" of outcomes (i.e. INSTANCES OF
BACKORDERS vs NO INSTANCES OF BACKORDERS) is an improvement on the 22%
accuracy for predicting a "specific" outcome (i.e. NUMBER OF INSTANCES OF
BACKORDERS), in order to be feasible in a real world setting- more accuracy is needed.

More data would help achieve this, but at the risk of over fitting. Therefore, to ensure
relevancy and reduce dimensionality, the data is inserted into a PCA analysis. In the current
dataset, it can be seen that 12 combinations of the total 16 features account for 96% of the
variability.
Figure 8. PCA where 12 components that account for 96% of the variability are selected.

After generating the PCA, logistic regression can then be performed again, but using the primary components as independent variables $x_n$. Now as the amount of available data increases, patterns will compress into combinations of features for better analysis, ensuring independence, preventing over fitting, and losing minimal data. Refer to Figure 9 for the new AUC showing 77%.

<table>
<thead>
<tr>
<th>Method</th>
<th>AUC</th>
<th>CA</th>
<th>F1</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>0.772</td>
<td>0.777</td>
<td>0.746</td>
<td>0.748</td>
<td>0.777</td>
</tr>
</tbody>
</table>

Figure 9. Results using Logistic Regression, where the performance is based on the Area Under the Curve (AUC)

7.2 Application

The Medical Device Directive means product classification changes, new requirements for traceability, and advanced clinical documentation. This section outlines a few of the
supply chain decisions involved with these changes, and illustrates an example of the
models during the decision making process. See Figure 10 below.

<table>
<thead>
<tr>
<th>REGULATORY CHANGE:</th>
<th>SUPPLY CHAIN DECISIONS:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PRODUCT CLASSIFICATION CHANGE</td>
<td>• Dual source for local supply</td>
</tr>
<tr>
<td>Import/Export requires more documentation and extends lead-time</td>
<td>• Manufacture closer to market</td>
</tr>
<tr>
<td></td>
<td>• Centralize stocking locations</td>
</tr>
<tr>
<td>2. NEW REQUIREMENTS FOR TRACEABILITY</td>
<td>• Outsource vs Insourcing Operation</td>
</tr>
<tr>
<td>Now Unique Device Identifiers are needed to be present on the packaging/label/product</td>
<td>• Make or Buy Packaging &amp; Labels</td>
</tr>
<tr>
<td></td>
<td>• Push vs Pull Final Assembly</td>
</tr>
</tbody>
</table>

Figure 10. Regulatory changes with Supply Chain decision implications.

Applying this model to a real-world example involves selecting one SKU out of the
current data set, changing an input variable to represent what it would be after the
regulatory change, and assessing if the classification outcome would improve. Because the
Logistic Regression outperformed the Linear Regression, Logistic Regression would be
used in this case. Below is an illustrated case example:

**Case Example:**

**SKU# 123456A- Ankle Implant** pictured in Figure 11 has experienced a regulatory change
requiring the device to have a Unique Device Identifier (UDI) physically present on the
implant. Unfortunately, the current manufacturing location does not have laser marking
capabilities. So, the supply chain manager will need to decide which manufacturing location
to transfer the product's source of supply without negatively
impacting service level.
Figure 11. Ankle implants that have experienced regulatory changes

This ankle implant is qualified for production in three different locations, two of which have laser marking capabilities.

- The first location uses a Direct Lazer Metal Sintering (DLMS) machine and is located in Europe. This machine will produce a porous surface implant which eliminates the need for a secondary coating operation, reducing the lead time from 53 days to 14 days. However, it has very limited capacity, at approximately 90% utilization it may not be able to fulfill the required demand. An example of this machine is given in Figure 12.

Figure 12. DLMS machine located in Europe
The second machine is an Automated Machining (AM) system located in a central region of the US. The AM will produce the base implant which then goes through an additional outsourced third party operation for surface coating, increasing the lead time from 53 days to 67 days. There is sufficient capacity. An example of this machine is provided in Figure 12.

![Automated Machine located in the U.S.](image)

The third machine is a Turn-Mill Center machine, a much less automated machine than the AM, which requires a human operator. The Turn-Mill is located in the Western part of the US. The Turn-Mill will also require an outsourced third party for a specialized vendor operation. This will increase the lead-time from 53 days to 63 days. There is sufficient capacity. An example of a Turn-Mill Center machine is provided in Figure 14.

![Turn-Mill Center Machine.](image)
In the existing process, prior to regulation change, the average manufacturing lead-time for a product to go to shelf is 53 days, and the probability that there will be instances of backorders within the year is 74%. In order to analyze the service level differences between the first, second and third options, the SKU attribute “AVERAGE LEADTIME” of 53 days is substituted with 14, 67 and 63 days, respectively. Logistic regression is then performed to determine if there will be (YES/NO) instances of backorders over a twelve-month period.

**Existing Process**: prior to regulation change = 74% probability that there will be instances of backorders over a twelve-month period.

<table>
<thead>
<tr>
<th>High v Low BO</th>
<th>Item_id</th>
<th>Logistic Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>200011901.0</td>
<td>1.0 0.257 0.743</td>
</tr>
</tbody>
</table>

**Option Number 1**: Probability that there will be instances of backorders within the year decreases from 74% to 62%

<table>
<thead>
<tr>
<th>High v Low BO</th>
<th>Item_id</th>
<th>ItemName_Name</th>
<th>Logistic Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>200010901.0</td>
<td>INBONEa...</td>
<td>1.0 0.382 0.618</td>
</tr>
</tbody>
</table>

**Option Number 2**: Probability that there will be instances of backorders within the year increases from 74% to 77%

<table>
<thead>
<tr>
<th>High v Low BO</th>
<th>Item_id</th>
<th>Logistic Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>555</td>
<td>200011901.0</td>
<td>1.0 0.232 0.768</td>
</tr>
</tbody>
</table>
Option Number 3: Probability that there will be instances of backorders within the year increases from 74% to 76%

Using Logistic Regression to measure SKU performance as a binary “YES or NO” for “INSTANCES OF BACKORDERS” has moderately high predictive power with approximately ~76% accuracy. Meaning, this data can somewhat accurately be used to assist in the decision making process. Out of the three outcomes, option number one is the obvious choice of location for manufacturing this item. This option reduces the probability of the SKU experiencing “INSTANCES OF BACKORDERS” by more than ten percentage points. Options number two and three, on the other hand, increase the probability of the SKU experiencing “INSTANCES OF BACKORDERS” by 3% and 2% respectively. In other words, options two and three both decrease service level—thus negatively impacting the business. Furthermore, the difference of only one percentage point is so minuscule that choosing either option won’t realistically impact the business.

The model’s results should lead to two actions from the supply chain. First, the management team, upon seeing these outcomes, should focus on increasing capacity for the DLMS machine in option number one. If they do not, service level will decrease resulting in backorders and lost revenue. Second, the management team, upon seeing these results, should focus on reducing lead time and driving efficiencies with the automated machine and
turn-mill machine in options two and three. This should both help reduce current backorders and positively impact the customer experience.

8. Conclusion and Next Steps

The example above illustrates several key considerations. First, the supply chain does not need to wait until a new medical device directive is issued to focus on initiatives like freeing up capacity or analyzing current service level improvements.

Second, data quality and availability is integral to the effectiveness of both the model and the outcome. As data availability improves over time, so will the model. However, in order to make this a sustainable, repeatable process in day to day operations some investment needs to be made in data cleansing and master data maintenance. A significant amount of work went into making such a large data set usable for programming. Even more work went into filtering it down to only usable parts. This is partially because both master data and sales history were dispersed among several different ERP systems. It would be beneficial to include more historical data of the same nature, as well as a wider set of attributes.

While the application of predictive analytics can bring valuable information to the supply chain, there is a law of diminishing returns. Without appropriate data integrity, the results are not informative. Without the correct level of data attributes available, significant hours are put into creating a model with very little predictive power. Moving forward, with some systems maturity and integration, this model has the potential to become much more robust. The benefits of which are improved supply chain efficiency, profitability, and a positive customer experience.
9. References


10. Appendix

**MDR Could Hinder Europe's Leading Position**

Global Medical Devices Risk/Reward Index

Scores out of 100. Higher Score = Lower Risk. Source: BMI Risk/Reward Index
11. Appendix

Orange Programming Data Analysis Workflow