

//Consumer Research//

Drug Switching Case Study

The Challenge

People often discuss their medical problems, circumstances and experiences in public discussion forums on the web. These archives of public social media forums contain valuable gems of useful consumer knowledge within volumes of useless junk. The challenge is to find them without incurring huge effort and cost.

This poses two important questions:

- Can we use the Web and social media as a marketing research sample?
- Can we quantify the rich, yet unstructured, information consumers post on the web?

Using the web and social media as a consumer research sample creates the following research opportunities:

- A combination of observational and descriptive marketing research
- Permits both qualitative and quantitative information analysis
- Non-invasive (no demand effect)
- Minimizes recall error
- Very rich data
- Sample size is not an issue
- Real time data

The web and social media also create the following challenges:

- Massive amount of data
- Data is all over the Web
- Data is unstructured
- Sample set is not necessarily representative



Our approach allows us to quickly and cost effectively address these challenges by using our proprietary software tools to search out and scan vast amounts of data, mine for our desired content, organize that content and create benchmarks of psychographic information. We can easily replicate these searches on a regular basis to monitor and update our findings.

Drug Switching Study Objective

Our objective for this project was to create a database of diabetes drug information from online forums that contain information on why patients stopped using a particular drug and the replacement drug. Additionally, we sought to better understand the consumer behaviors associated with diabetes drug usage and switching.

More specifically, we sought to answer the following questions by efficiently mining information from **public medical forums**:

1. What diabetes medications people stopped using?
2. The specific reasons they stopped using them.
3. And, with which medications they replaced them.

Summary of Benefits

Using our Consumer Research Platform to mine **public medical forums**, we were able to:

1. Quickly & cost effectively mine information from hundreds of thousands of respondents' data about:
 - Brand switching
 - When respondent started and stopped
 - Reason for starting and stopping
 - Replacement drug
 - Reason for replacement
2. Create a benchmark measure of psychographic information – who are the consumers?
3. Develop an inexpensive yet robust, automatic ongoing database of information that allows you to track information over time and identify factors related to changes
4. Leverage your consumer research budget for any type of product

Input Required

In order to develop this Consumer Research Program we simply required general directional information including:

Minimal key words & concepts:

- What keywords are most associated with the brand and the industry
- What are the concepts you are most interested in learning more about

Target forums and social media sites:

- What online forums and social media sites most likely contain the information we want
- Are there any specialized vertical industry forums?

Specific questions to be answered by public data:

- Brands used
- Reasons for brand switching
- Replacement product
- Reason for replacement

Our Approach/Solution

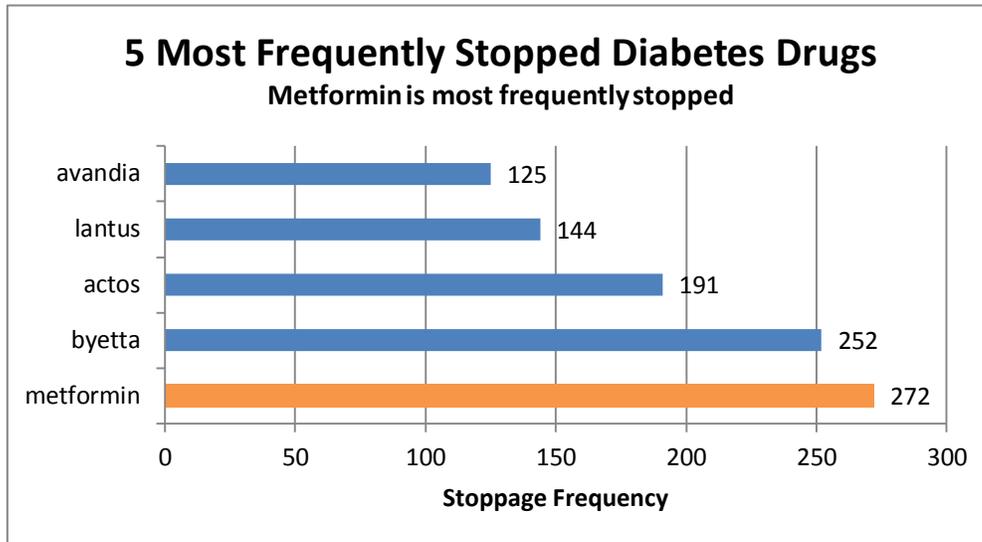
To build our database and capture psychographic insights, we used a patented software-based relation learning and extraction framework (CARE-II-HPSG/VC*) to mine public online medical forums. The software automation process included:

- Building a list of known diabetes drug names by “scraping” drugs.com.
- Creating a diabetes-specific lexicon by pattern learning from 20,000 sentences downloaded from the following forums:
 - Diabetes.blog.com
 - Diabetesdaily.com/forum
 - Diabetesforums.com/forum
 - Healthboards.com/boards
 - Forum.lowcarber.org
- Clustering patterns identified and defining output relations (semi-automated)
- Post-processing of complex relationships

We analyzed 650,000 forum messages consisting of approximately 300,000 non-duplicated sentences totaling 40MB of ASCII text. We found 2,732 distinct instances of diabetes drug replacement information.

The Results

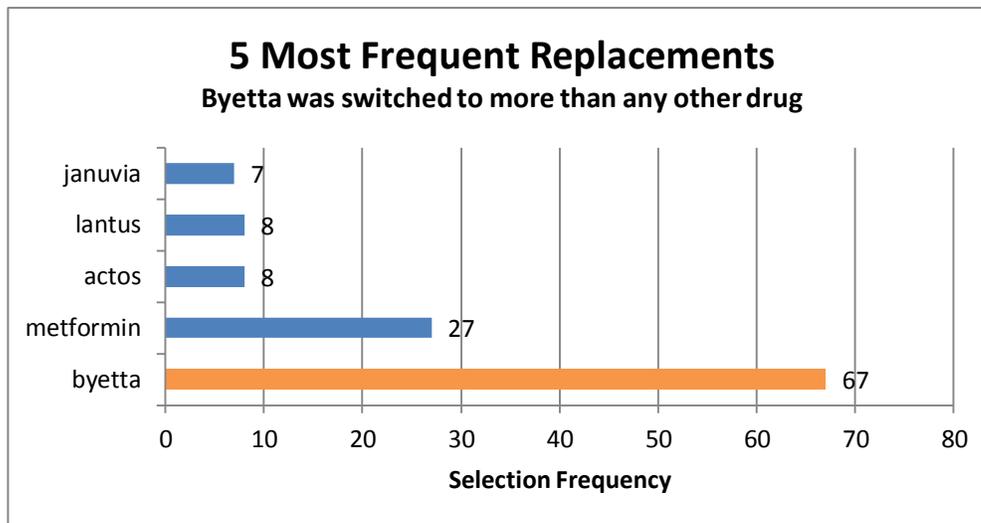
1. We were able to effectively identify the drugs with the most frequent usage stoppage.



2. Create a list of the adverse drug reactions (ADRs) that were frequently mentioned as having a negative relationship with each of the diabetes drugs.

Drug	Adverse Drug Reaction
Metformin	Lactic acid
Metformin	Taste problems
Metformin	Muscle pain
Metformin	Stomach cramps
Metformin	Diarrhea
Metformin	Digestive disorders
Metformin	Leg pain

3. Identify the most frequent replacements for each drug as well. The chart below actually shows the overall replacements regardless of the original drug used.



NOTE: Metformin and Actos are the 2 most prescribed Type 2 diabetes drugs.

Accuracy

Based on a random sample of messages manually checked, our software was:

- **98% accurate** at identified usage-stopped diabetes drugs
- **93% accurate** identifying the replacement drug
- **93% accurate** identifying the reason for stoppage

Summary

Fast, simple to use, cost effective method of collecting valuable information from a large number of data points (650K). It can be used alone and/or with other types of research.

Goes beyond standard text mining and uses patented technology to transform and quantify the rich, yet unstructured, information consumers post on any and every social media platform

Search the web and social media for key consumer insights derived directly from consumer comments online combining the best of observational and descriptive marketing research.

Review and structure massive amounts of data to create a cogent picture of near real-time consumer sentiment that can be *easily* and *cost effectively* replicated.

We believe the efficiency and accuracy of this platform will change the way companies learn from consumers.