“HOW I LEARNED TO STOP WORRYING
AND LOVE THE INTERWEBSES”

ANALYSIS OF SELF-REPORTED HEALTH OUTCOMES
DATA FROM WEB BASED MEDIA SOURCES

Mark Wolff, Ph.D.
Principal Industry Consultant
Health & Life Sciences Global Practice
SAS Institute
Cary, North Carolina

Michael Wallis, M.S.
Analytical Consultant
Advanced Analytics Division
SAS Institute
Cary, North Carolina

Kenneth Lopiano, Ph.D.
Research Associate
SAMSI
Research Triangle Park, North Carolina
PROMS  PATIENT REPORTED OUTCOMES MEASURES

- Tremendous interest in accessing and using unstructured data to enhance data driven decision making related to patient safety, quality and outcomes.

- Interest not limited to internal/formal data
OBJECTIVE
FROM INFINITY TO SAS TABLES
Twitter Reaction to Events Often at Odds with Overall Public Opinion

By Amy Mitchell and Paul Hitlin
"The resulting dataset contained a high volume of irrelevant information, but provided a useful starting point."

"We did not seek to verify each individual report as truthful, but rather to identify overall associations between Twitter and official spontaneous report data as a preliminary proof of concept."
CONSUMERS …IT’S SILLY, BUT I BELIEVE.

Social Media: Study Shows the Most Trusted Product Reviewer is YOU
Posted by Ken Krause on Tue, Jan 29, 2013

According to a September 2012 study of consumers who had recently made electronic purchases, those surveyed said they paid more attention to customer reviews than professional critics in every category.

http://www.techimage.com
GOOGLE TREND

...RUN AWAY, RUN AWAY, RUN AWAY
OBJECTIVE SAS TEXT ANALYTICS AND DEVICE SAFETY MONITORING

1) Identify websites where individuals are posting publicly available comments on their experience with the device of interest
2) Collect, process, and analyze comments
   a. Identify documents that;
      i. Refer specifically to the device of interest
      ii. Comments that are part of a thread related to the device of interest
      iii. Contain terms identified as known “Adverse Events” (AEs)
         1. AE terms consistent with the product label
         2. Specific AE terms not on the product label
         3. Terms that identify other drugs, substances, or devices
   b. Perform document author “Veracity/Integrity analysis
3) Identify relevant scientific literature PubMed abstracts
4) Identify relevant reports as collected on the MAUDE database
DATA ON LINE DATA SOURCES
SEGMENTATION

COMMENT CLUSTERING

- high interaction, many posts
- low interaction, many posts
- low interaction / few posts

Graphs showing the number of replies received by author for different interaction scores and types of authors:

1. Helpful Authors
2. Unhelpful Authors
3. Typical Authors
NEW APPROACHES

BURST DETECTION

Hierarchical Structure of Bursts For Topic 6

Distribution of posts over time intervals

(a) Normalized frequency by class width for topic 6

<table>
<thead>
<tr>
<th>Time intervals</th>
<th>Frequency normalized by class width</th>
</tr>
</thead>
<tbody>
<tr>
<td>38111, 39324</td>
<td>0.0072</td>
</tr>
<tr>
<td>40929, 40352</td>
<td>0.3913</td>
</tr>
<tr>
<td>40352, 40646</td>
<td>0.6531</td>
</tr>
<tr>
<td>40846, 40782</td>
<td>0.3956</td>
</tr>
<tr>
<td>40702, 40842</td>
<td>0.6071</td>
</tr>
<tr>
<td>40842, 41239</td>
<td>0.3223</td>
</tr>
<tr>
<td>41239, 41331</td>
<td>0.8261</td>
</tr>
</tbody>
</table>
VISUALIZATION

...THIS GOES TO 11
POST-MARKETING SAFETY MONITORING OF SELF-REPORTED SYMPTOM-TREATMENT-OUTCOME MEASUREMENT IN WEB-BASED MEDIA SOURCES

Project Description
FDA is responsible for communicating about the risks and benefits inherent in all the products it regulates.

• The rise of social media on the Internet, including especially user-generated content such as blogs, forums, message boards, wikis and podcasts, has created new opportunities to interface with the public with respect to emerging hazard situations involving FDA-regulated products.

• The increasing presence of social media promises new capabilities to monitor the effectiveness of FDA’s ongoing risk communication efforts. FDA is in need of both historical and "real-time" monitoring and analyses of a representative sample of social media web sites.

The objective of this requirement is to provide FDA with the resources needed to use social media to inform and evaluate FDA risk communications. Specifically, the objective is to provide FDA with:

• Analyses of social media that provide baselines on consumer sentiment prior to FDA communication and that depict changes in social media buzz following FDA communications

• In-house capability for social media monitoring; and Surveillance through social media listening for early detection of adverse events and food-borne illness.

• The scope of work includes social media buzz reports, a social media dashboard, and quarterly surveillance reports related to specific product classes.
“While the individual man is an insoluble puzzle, in the aggregate he becomes a mathematical certainty”

Sherlock Holmes