

MedWatcher: Harnessing emerging technologies for pharmacovigilance

Sumiko R. Mekaru, Clark C. Freifeld, John S. Brownstein
HealthMap, Boston Children's Hospital & Epidemico
sumiko.mekaru@childrens.harvard.edu

Half of Americans take a prescription drug, medical devices are in broad use, and population coverage for many vaccines is over 90%. Nearly all medical products carry some risk of adverse events (AEs), sometimes severe. However, pre-approval trials use small populations and exclude participants by specific criteria. After approval, these products are used by populations not necessarily represented in the original trials; discovery of new AEs is inevitable. Existing post-marketing reporting systems seek to identify these new AEs, but various barriers lead to underreporting of AEs.

MedWatcher is a new system that harnesses emerging technologies for pharmacovigilance in the general population. MedWatcher consists of two components: MedWatcher Social (a text-processing module) and MedWatcher Personal (a crowdsourcing module). In MedWatcher Social, a natural language processing component applies classification algorithms and extracts AE signals from public data from the Internet. The crowdsourcing application provides software to allow consumers to submit AE reports directly.

The MedWatcher Social algorithm for identifying symptoms performs with 77% precision and 88% recall on a sample of Twitter posts. Our machine learning algorithm for identifying AE-related posts performs with 68% precision and 89% recall on a labeled Twitter corpus. For zolpidem tartrate, certolizumab pegol, and dimethyl fumarate, we compared AE profiles from Twitter with reports from the FDA spontaneous reporting system. We find some concordance (Spearman's rho = 0.85, 0.77, 0.82, respectively, for symptoms at MedDRA System Organ Class level). Where the sources differ, milder effects are overrepresented in Twitter. We also compared post-marketing profiles with trial results and found little concordance.

The MedWatcher App was developed with the FDA Center for Device and Radiologic Health to provide the public and healthcare providers a simple and fast way to report adverse events to the FDA. This desktop and mobile app provides an alternative to the outdated processes of tedious PDF forms and complicated documentation to address the safety issue of severely underreported adverse events. MedWatcher Personal saw substantial user adoption, receiving 550 AE reports in a one-year period, including over 400 for one device, Essure. We categorized 400 Essure reports by symptom, compared them to 129 reports from the FDA spontaneous reporting system, and found high concordance (rho = 0.65) using MedDRA Preferred Term granularity. We also compared Essure Twitter posts with MedWatcher and FDA reports, and found rho = 0.25 and 0.31 respectively.

MedWatcher represents a novel pharmacoepidemiology surveillance informatics system; our analysis is the first to compare AEs across social media, direct reporting, FDA spontaneous reports, and pre-approval trials. More than simply submitting reports, users can also make a

quick list of all the prescription medicines, devices, and vaccines their family or patients use and be notified of other adverse event reports or track the latest developments. MedWatcher covers all US drugs, devices, and vaccines. Figure 1 shows back end data flow for adverse event reports, and Figure 2 shows screenshots from the mobile reporting app. This project started in 2012 and the MedWatcher team is seeking ways to work with other departments of the FDA to make the platform more robust and streamlined for non-device products. Expansion into the EU is also in progress.

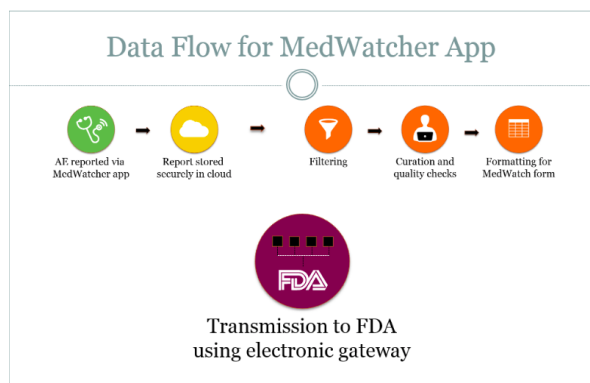


Figure 1: MedWatcher Data Flow

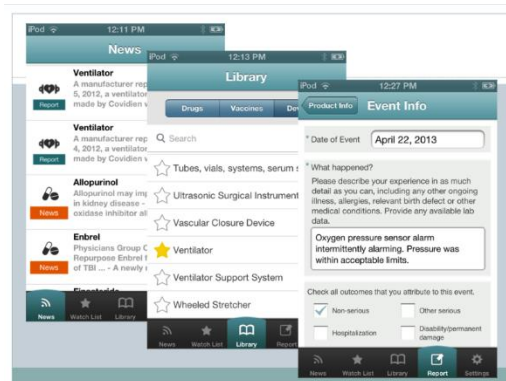


Figure 2: MedWatcher App Screen Shots

MedWatcher Social is a platform built to provide the FDA social media analytics capabilities specifically tailored to adverse event monitoring. MedWatcher Social collects, categorizes, and analyzes massive amounts of consumer-reported health experiences from blogs, forums and other social media venues like Facebook and Twitter. Our automated processes of Natural Language Processing (NLP) and proprietary machine learning algorithms are accompanied by manual curation to effectively distill insights from huge amounts of noise and to ensure accurate output reporting. Capabilities of MedWatcher Social are best described in our recent *Drug Safety* publication. It showed significant correlation between adverse events reported in Twitter and those reported officially through the FDA (Freifeld 2014). This is an ongoing project, started in 2012, with iterative development in collaboration with the FDA’s Office of the Chief Scientist.

The data collected in both MedWatcher products can support numerous types of research questions, including those with a human dynamic or spatial component. As the corpus of collected social media messages grows, text processing can be used to identify potentially relevant factors. For example, if AEs for a medication are often reported in the context of exercise (e.g. “Went for a run but got too light-headed to finish #MedName”), this additional information may be critical for understanding the triggers for the AE. In another example, regional clustering of reports may indicate a multi-factorial cause for certain adverse events. If submitted reports for an AE are clustered in a particular state despite wide spread use of the medication across the country, the clustering could suggest a spatial-dependent location. For example, if the AEs were only in western Nevada, the prevalence of elevated arsenic in drinking water there could play an important role in the AEs. In conclusion, the MedWatcher platform should provide numerous research opportunities over the next several years.