

**The Food and Drug Administration's (FDA's)  
2015 ORSI Science Symposium  
April 27, 2015  
SPEAKER ABSTRACTS AND BIOGRAPHIES**

**Session 1: Centers for Excellence in Regulatory Science and Innovation (CERSIs) Presentations – 8:35-11:30 AM**

**University of California-San Francisco/Stanford University CERSI**

CERSI	<b>University of California-San Francisco (UCSF)-Stanford University</b>
Speaker	<b>Nigam H. Shah, MBBS, PhD</b>
Title and Location	<b>Assistant Professor Medicine (Biomedical Informatics) Stanford University, Stanford, CA</b>
Biography	<p><b>Dr. Nigam Shah is assistant professor of Medicine (Biomedical Informatics) at Stanford University, Assistant Director of the Center for Biomedical Informatics Research, and a core member of the Biomedical Informatics Graduate Program. Dr. Shah's research focuses on combining machine learning, text-mining, and prior knowledge in medical ontologies to enable use cases of the learning health system.</b></p> <p><b>Dr. Shah received the AMIA New Investigator Award for 2013. Dr. Shah integrates teaching into his advanced research work and was recognized with the Biosciences Faculty Teaching Award for outstanding teaching contributions in his graduate class on “Data driven medicine” (Biomedin 215). He holds an MBBS from Baroda Medical College, India, a PhD from Penn State University and completed postdoctoral training at Stanford University. More at: <a href="https://med.stanford.edu/profiles/nigam-shah">https://med.stanford.edu/profiles/nigam-shah</a></b></p>
Presentation Title	<b>Improving ADR Signal Detection by Combining Signals from Spontaneous Reports and Electronic Health Records</b>
Presentation Abstract	<p><b>Adverse drug events (ADEs) are undesired harmful effects resulting from use of a medication, and occur in 30% of hospitalized patients. We propose to augment ADR signal detection by combining datasets that capture complimentary dimensions about drug safety profiles.</b></p> <p><b>Our approach has two parts: 1) We characterize the relative gain in signal detection accuracy by combining FAERS and EMR data instead of using them in isolation. 2) We then build a machine learning system that operates on EMR data, and alerts clinicians to situations where submission of a FAERS report should be considered. In this presentation, we will review initial results on the machine learning system, which uses the text from 9.5 million clinical notes, along with prior knowledge of drug usages and known adverse drug events, as inputs. These inputs are used by a discriminative classifier which outputs the probability that a given drug-disorder pair represents a valid adverse drug event association. We evaluate our method by assessing support for the predictions in other curated data sources, including a manually curated, time indexed reference standard of label change events. Our classifier achieves an area under the curve (AUC) of 0.94 on a held out test set, and predicts 240 high-confidence, well-supported drug-AE associations. 36% of the predictions are supported in at least one of the resources which have information that was not available to the classifier; demonstrating the feasibility of systematic post-marketing surveillance for ADEs using machine learning electronic medical records.</b></p>