

Leveraging the ‘Wisdom of the Crowd’ as a biosurveillance tool

Izchak Lichtenfeld¹, and Limor Bar-Hai^{2*}

¹Analytics, First Life Research, Yehud, Israel; ²Medical Informatics, First Life Research, Yehud, Israel

Objective

With a large population sharing experiences regarding health issues and treatments online via social media platforms, generating novel data sets composed of massive unstructured user-generated content of health reports. This collective intelligence is referred to as the ‘Wisdom of the crowd’.

This is a brief overview of data research engaging this unique statistical sample referred to as the ‘Crowd trial’ as an innovative element in health monitoring, enabling early detection and intervention by health professionals, regulators and pharmaceutical companies.

Introduction

With the proliferation of social networks, the web has become a warehouse of patient discussions and reports, estimated at 10 billion records and growing at a rate of 40 percent per year. First Life Research, Ltd. (FLR), has searched and mapped thousands of these discussions and indexed hundreds of millions of reports (currently 960M) and is engaged in building web-based solutions that enable the public and public health practitioners to access massive health-related information and knowledge generated from the crowd.

Methods

FLR’s competency is the ability to identify, analyze, index and aggregate user-generated content by collecting billions of testimonials from social networks. It utilizes cutting edge technologies for massive data aggregation and applies advanced natural language processing (NLP) techniques for continuous analyses, in order to convert this unstructured data into refined information. The insights gained can be used to support and enable better informed decision making processes, both for patients and healthcare providers.

Results

A platform of data investigations utilizing the ‘Wisdom of the Crowd’ focusing on biosurveillance aspects as follows:

1. Pharmacovigilance—brand monitoring and safety alerts: Crowd trial provides a dashboard of parameters on medications, their side effect profile, interactions and drug’s comparative advantage.
2. Social Health at a glance: Temporal overview of prevalence and statistics of the most engaging health issue discussed across the social web, represented by aggregation of the reports (citations) generated by the e-patients (Table 1).
3. Health trends detected by harnessing the social web: This public feedback exists in real-time, large scale and enables ongoing observational studies by tapping into the health reports involving a massive sample size (Fig. 1).

Table 1. A sample of the most prevalent health issues as reflected by the social web (October 2011)

Most Engaging Health Issues_October 2011		
		# Citations
1	Mental Health	64,301,887
2	Neurology	27,264,378
3	Pregnancy & Childbirth	12,112,473
4	Gastro & Digestive	11,940,019
5	Cardio-Vascular Disorders	11,395,631

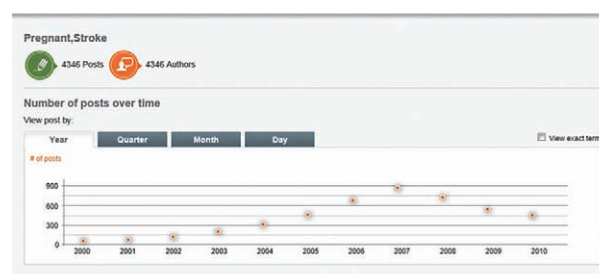


Fig. 1. Pregnancy-related strokes on the rise: a CDC study as it reflect in the Health 2.0. (health social web).

Conclusions

The value of crowd trial to public health is new and complementary to what the existing monitoring processes provide. Moreover, user-generated content contains valuable feedback on medication usage and health information.

Thus the emerging wisdom of the crowd analytics potentially represents a new phase and eventually new tools using data evaluation based on large scale population inputs, and it will benefit greatly all public health environment.

Keywords

Crowd trial; biosurveillance; social web; user generated content; Health 2.0

Acknowledgments

First life research Ltd. (FLR) was founded in late 2007 and is the creator of Treato™ -a comprehensive searchable medical database, based solely on user generated content provided by the e-patients from various websites throughout the world wide web.

*Limor Bar-Hai

Head of medical informatics at First Life Research Ltd
E-mail: limor@firstliferesearch.com