Social media and pharmacovigilance: A review of the opportunities and challenges

Richard Sloane,[1,2,5]* **Orod Osanlou,**[2,3]* **David Lewis,**[4] **Danushka Bollegala,**[5] **Simon Maskell**[1,5] **& Munir Pirmohamed**[2,3]

[1]*Department of Electrical Engineering and Electronics, University of Liverpool, L69 3GJ, UK,* [2]*Department of Molecular and Clinical Pharmacology, University of Liverpool, L69 3GL, UK,* [3]*Royal Liverpool and Broadgreen University Hospital NHS Trust, Liverpool L7 8XP, UK,* [4]*Drug Safety & Epidemiology, Novartis Pharma AG, Postfach, CH-4002, Basel, Switzerland and* [5]*Department of Computer Science, University of Liverpool, L69 3BX, UK*

Adverse drug reactions come at a considerable cost on society. Social media are a potentially invaluable reservoir of information for pharmacovigilance, yet their true value remains to be fully understood. In order to realize the benefits social media holds, a number of technical, regulatory and ethical challenges remain to be addressed. We outline these key challenges identifying relevant current research and present possible solutions.

Introduction

Adverse drug reactions (ADRs) are a significant cause of morbidity and mortality worldwide. Approximately 6.5% of UK hospital admissions are caused by ADRs at considerable cost on health, finance and manpower [1]. Even though the number of Yellow Card reports is increasing year-on-year with over 26 000 reports in 2012 [2], the vast majority of ADRs still remain unreported [2]. It has thus been suggested that social media may be a method by which pharmacovigilance could be strengthened given that approximately 72% of internet users search for health information online [3].

Social media have significantly expanded recently with Twitter and Facebook now registering over 300 million and 1.4 billion active users, respectively [4–7]. Social media have already been utilized in different domains illustrating their potential application to pharmacovigilance. For instance, Twitter data were analyzed to determine epicentres of earthquakes in Japan [8] and following the devastating 2010 earthquake in Haiti, Twitter was shown to map trends of the cholera outbreak weeks before traditional reporting methods [9]. The quantity and near-instantaneous nature of social media provides potential opportunities for real-time monitoring of ADRs, greater capture of ADR reports and expedited signal detection if utilised correctly.

The role of social media in pharmacovigilance has been gaining in interest with various social media sources [10–31] used for detecting ADRs, including general purpose social networking sites such as Twitter [6], and health and support networks including PatientsLikeMe [32], DailyStrength [33] and MedHelp [34] (Figure 1). Previous reviews in this area have focused on the approaches that have been taken to analyze social media [35], and the analysis of various pharmacovigilance text sources including biomedical literature, clinical narratives and social media [36]. To complement these existing reviews, we take a broad view to incorporate the technical, industrial, legislative and ethical perspectives of using social media for pharmacovigilance in order to highlight current challenges and potential opportunities. Words in *italics* are described in the 'Glossary of terms' section.

Text mining of social media for pharmacovigilance

Text mining has been applied to various sources of pharmacovigilance data including biomedical literature, clinical narratives and web search logs [36]. To the best of our knowledge, the first application to social media was by Leaman *et al.* in 2010 [31], analyzing postings from the

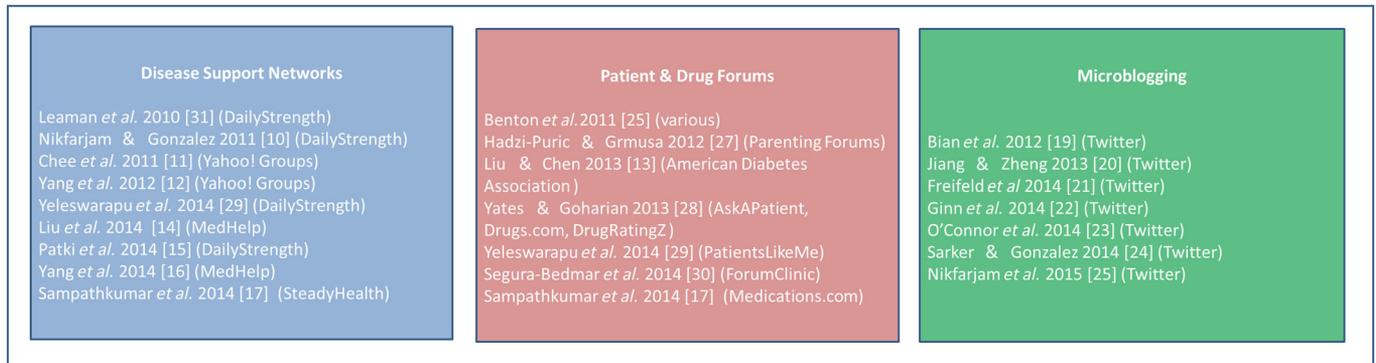| Disease Support Networks	Patient & Drug Forums	Microblogging
Leaman *et al.* 2010 [31] (DailyStrength) Nikfarjam & Gonzalez 2011 [10] (DailyStrength) Chee *et al.* 2011 [11] (Yahoo! Groups) Yang *et al.* 2012 [12] (Yahoo! Groups) Yeleswarapu *et al.* 2014 [29] (DailyStrength) Liu *et al.* 2014 [14] (MedHelp) Patki *et al.* 2014 [15] (DailyStrength) Yang *et al.* 2014 [16] (MedHelp) Sampathkumar *et al.* 2014 [17] (SteadyHealth)	Benton *et al.* 2011 [25] (various) Hadzi-Puric & Grmusa 2012 [27] (Parenting Forums) Liu & Chen 2013 [13] (American Diabetes Association) Yates & Goharian 2013 [28] (AskAPatient, Drugs.com, DrugRatingZ) Yeleswarapu *et al.* 2014 [29] (PatientsLikeMe) Segura-Bedmar *et al.* 2014 [30] (ForumClinic) Sampathkumar *et al.* 2014 [17] (Medications.com)	Bian *et al.* 2012 [19] (Twitter) Jiang & Zheng 2013 [20] (Twitter) Freifeld *et al* 2014 [21] (Twitter) Ginn *et al.* 2014 [22] (Twitter) O'Connor *et al.* 2014 [23] (Twitter) Sarker & Gonzalez 2014 [24] (Twitter) Nikfarjam *et al.* 2015 [25] (Twitter)

Figure 1

Social media sources utilized in text-mining studies

social support networking site DailyStrength. Subsequent mining of several social media sources has been comprehensively covered by Sarker *et al.* [35]. A trend is evident in the sources considered with earlier works tending to focus on forums and user groups, and more recent studies focusing on large-scale, more generalized sources such Twitter [19, 23] (Figure 2). Current challenges in extracting ADRs from social media are outlined in Table 1. Challenges 1 and 2 relate to the fact that both drug names and reaction terms can be described in a variety of ways. This is an example of the *named entity recognition* (*NER*) problem in the information retrieval community and is an active research area. Some studies

[19, 20] have addressed the issue of multiple drug terms by performing search queries involving both brand and generic name variants of the relevant drugs. One study generated variations based on drugs' phonetic spellings [37], addressing issues associated with misspelling (challenge 3). Although state-of-the-art methods for automating the identification of name aliases exist [38], these have not been applied to pharmacovigilance.

Addressing the diversity of ways in which ADRs may be expressed is more challenging as this can involve complex phrasing. Several studies [16, 29] used pre-defined lists of adverse reaction terms taken from resources such as the FDA Adverse Event Reporting System (FAERS) [39],
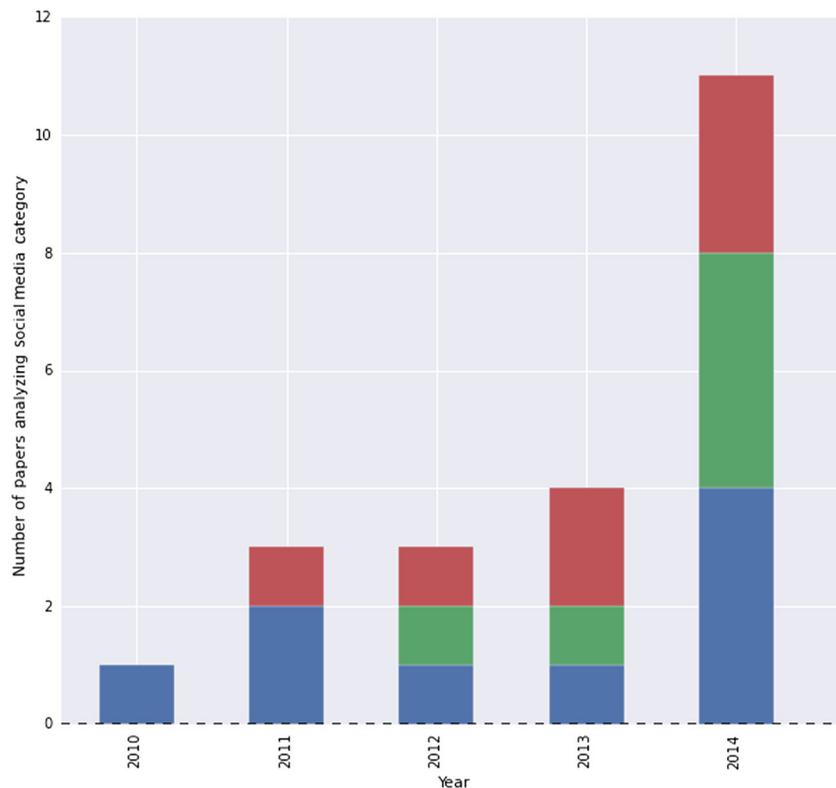

Figure 2

Chart showing the categories of social media analyzed in recent years. Disease support network -, Microblogging -, Patient and drug forums -

Table 1

Key challenges in identifying adverse drug events from social media

1) Drugs may be described by their brand names, active ingredients, colloquialisms or generic drug terms (e.g. 'antibiotic')
2) ADRs may be referred to using creative idiomatic expressions or terms not found within existing medical lexicons
3) The informal nature of social media results in a prevalence of poor grammar, spelling mistakes, abbreviations and slang
4) The existence of a side effect may be clear while the specific side effect experienced remains unclear
5) Discussion of a drug could involve indications, beneficial effects or concerns of an adverse event
6) Supervised machine learning, while powerful, needs training data which requires time-consuming and expensive generation of human-annotated data

Only a small percentage of social media will relate to ADRs

Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART) [40], Collaborative Consumer Health Vocabulary (CHV) [41], MedEffect [42], Unified Medical Language System (UMLS) [43], MedDRA [44] and Side Effect Resource (SIDER) [45]. As noted by Sarker *et al.* [35], such lexicon-based approaches will struggle to capture cases where creative language is used to express an ADR. An example of such creative phrasing is the tweet (generic names added in square brackets) 'Humira [adalimumab] never really worked for me. Orencia [abatacept] was good. Xeljanz [tofacitinib] was the best but ate a hole in my stomach'. In addition to using creative phrasing, this particular tweet identified by Freifeld *et al.* [21] lacks specific details and the treatment received making it difficult to assess the exact nature of the ADR (challenge 4) and assess causality (challenge 5). Thus, while the tweet may describe simple dyspepsia, it may also relate to ulceration and even gastrointestinal perforation, which is potentially life-threatening, and is known to occur with tofacitinib. An ethical issue here is whether regulators should contact such patients to obtain further information.

Supervised machine learning algorithms require substantial amounts of data to be manually curated, often by a domain expert, and are costly in terms of time, money and manpower (challenge 6). This is further complicated by the fact that much of social media data is irrelevant to pharmacovigilance presenting a challenge to algorithms to detect ADRs automatically (challenge 7). To provide a sense of scale, previous studies [21, 22] involved experts annotating 10 822 and 61 401 Twitter posts, respectively. We have been unable to identify significant research on reducing the quantity of annotation required. Ginn *et al.* [22] did investigate the effect of combining training data from medical case reports, DailyStrength and Twitter with encouraging results through the use of carefully selected variables. A recent study [46] explicitly aimed to reduce the amount of training data by use of 'partially supervised' learning based on discussion threads from MedHelp [34]. A current relevant research area relevant to reducing the need for manual annotation is *active learning* [47]. Reducing the quantity of human annotation would also increase the ease with which new signal sources could be incorporated (e.g. foreign language text). Motivated by this, recent work has developed a cross-language similarity measure for biomedical terms [48].

The role of statistics in pharmacovigilance and social media

Although Freifeld *et al.* showed a general concordance between consumer-reported ADRs on Twitter and those reported directly to the FDA in the United States [21], careful consideration is needed of the statistical interpretation of processed data and how to utilize it effectively.

Disproportionality ratios take into consideration the occurrence of an ADR against the expected occurrence of the event from a control group [49]. However, such approaches suffer from various sources of bias including:

- Reporting bias, i.e. the population reporting ADRs are a potentially biased sub-population of those experiencing ADRs. It is well established that ADRs are generally under-reported, even in serious cases [50]. Furthermore, it is natural to expect that the reporting medium may affect the reporting rate of a specific event. For instance, a patient suffering from impotence may be less likely to post this on a social media site.
- Notoriety bias may occur as a result of publicity or because the drug has recently been introduced onto the market, a notable example being bisphosphonates and osteonecrosis of the jaw [51]. Sertindole is an atypical antipsychotic that was suspended due to such bias. Despite having a risk of fatal arrhythmias broadly similar to other atypical psychotics, it had an increase in its reporting incidence of nearly 10-fold [52]. This needs to be accounted for when analyzing signals.
- A further source of bias in current disproportionality calculations is that the comparator incidence of the ADR under scrutiny is determined using 'reported' ADRs. As social media users are generally free to discuss whatever they choose, this allows for a new opportunity to estimate the occurrence of each particular event in the absence of 'any' drug. This is an intriguing possibility, which has not yet been investigated with respect to pharmacovigilance.
- While social media provide an opportunity to tackle existing biases in pharmacovigilance, this does not mean that social media are an unbiased medium. It is reasonable to expect that social media user bases are skewed in terms of age, gender, ethnicity and physical location. Studies have been performed to make inferences regarding demographics of Twitter users such as age [53] and ethnicity [54]. Understanding the sources of bias in social media data are integral to correcting for them.

Established statistical techniques exist to deal with cases where information is not readily available regarding the population of interest, such as those experiencing ADRs as opposed to those reporting ADRs. Importance sampling [55] is a mature statistical technique where data from one population can provide unbiased estimates to another population, but has not been applied to either social media analyses or clinical trials. Patient groups such as pregnant women and children are often excluded from clinical trials due to concerns of teratogenicity and ethical issues [56, 57] yet they represent active social media users [58]. Techniques such as importance sampling provide state of the art opportunities for expedited detection of ADRs for such patient groups.

Interpreting a degree of trust, or uncertainty, in the various pharmacovigilance 'sensor' sources is integral to utilizing data effectively. In social media, trustworthiness could be interpreted for a given post with potentially ambiguous language or for a particular user given their reporting history, including duplicate reports. Duplicate reports for a user could include mentioning an ADR several times or re-posting a report from another user. Reposted reports may provide valuable insight as to the trust of a particular user or report. Combining the information these sensors provide will be crucial. The Dempster–Shafer theory [59] is one method to combine evidence from multiple sources while considering the trust in each source. To the best of our knowledge, this approach has not been applied to pharmacovigilance though it has been applied to social media analysis [60]. Harpaz et al. [36] recognize that progress in pharmacovigilance depends on a comprehensive approach to combining different pharmacovigilance signal sources, with previous work combining data from electronic health records and spontaneous reports [61, 62]. We believe these approaches can be augmented by carefully characterizing the bias in each pharmacovigilance source and utilizing more sophisticated techniques to combine source data (e.g. [59]). As introduced earlier, each pharmacovigilance source will be subject to various forms of reporting bias. However, with each source's common dependence on the underlying quantity of interest, this allows for the sources to mutually de-bias one another.

Social media and the pharmaceutical industry

Social media attract much attention from the pharmaceutical industry with sites such as LinkedIn [63], Twitter [6], Facebook [7], YouTube [64], Flipboard [65], Flickr [66] and Pinterest [67] proving valuable in establishing intra- and inter-company stakeholder collaborations and communities. Sponsors within the industry are utilizing social media for commercial purposes to distribute information about diseases and their treatment, medicines and company announcements [68]. Commercial arms of pharmaceutical companies also use social media to monitor conversations led by patients, carers and healthcare professionals regarding marketed medicines [69–71]. A minority of companies also utilize social media in patient engagement for marketing purposes or for recruitment and retention within clinical trials [68–71]. The latter may cause significant difficulty if the subjects enrolled use social media for communication, as it may lead to the unblinding of investigational medicinal products including the product of interest, active comparators and/or placebo.

Social media sites are increasingly being adopted by healthcare professionals and patients for the discussion of adverse experiences with licenced medicinal products. However, the International Harmonized Guideline [72] which provides the definitions and standards for expedited reporting and clinical safety data management was published nearly a decade before the term social media was coined. In 2011, Edwards & Lindquist [73] indicated that 'some years ago' experts within the Council for the International Organizations of Medical Sciences (CIOMS) unanimously concluded that the pharmaceutical industry trawling the Internet for safety information was '…an impossible task, unenforceable and not cost effective'. Edwards & Lindquist [73] however also concluded that the World Wide Web, the information it carries, and the impact of this information cannot be ignored. Rapid technological advancement, the prominence of social media and the change of regulations means that this hypothesis can be extended to social media.

Marketing authorization holders (MAHs) are obliged by European law to establish and maintain a system for pharmacovigilance and record all suspected adverse reactions brought to their attention [74–76]. This legislation is further expanded within good pharmacovigilance practice (GVP) modules [77] which outline the responsibilities of the MAH to record suspected ADRs from digital media. The original GVP module VI from June 2012 was specific in recommending that all individual case safety reports (ICSRs) on digital media be captured, recorded and reported in accordance with the law. A literal interpretation of this would require vast human resource or technological solutions to the challenges in Table 1 which are currently outstanding. In September 2014 GVP module VI was updated to require that MAHs regularly screen websites under their control for suspected adverse reactions (Table 2). The same GVP stipulates that it is considered good practice for the MAH to monitor external sites such as patient support or special diseases group sites with the frequency of the monitoring depending on the risks associated with the medicinal product. GVP suggests that if the MAH is 'aware' of a suspected ADR on external sites, it should evaluate whether the posting qualifies for reporting and if so, handle it in the same manner as a spontaneous report. The definitions of 'monitoring' and awareness' of suspected ADRs are

Table 2

Obligation of marketing authorization holders to report information on suspected adverse reactions from the internet or digital media

> *Marketing authorization holders should regularly screen internet or digital media [see* Definitions*] under their management or responsibility, for potential reports of suspected adverse reactions. In this aspect, digital media is considered to be company sponsored if it is owned, paid for and/or controlled by the marketing authorization holder [see* Definitions*]. The frequency of the screening should allow for potential valid ICSRs to be reported to the competent authorities within the appropriate reporting timeframe based on the date the information was posted on the internet site/digital medium. Marketing authorization holders may also consider utilizing their websites to facilitate the collection of reports of suspected adverse reactions.*

somewhat open to interpretation with unclear legal implications for the pharmaceutical industry and regulators. As previously discussed, algorithms to detect ADRs automatically often require interpreting significant numbers of suspected ADRs to work effectively. This raises some important questions: should the MAH report each suspected ADR each time the data is utilized? Had the MAH conducted research within these data without reporting a serious ADR, would they subsequently be liable for prosecution?

In the US, the Food & Drug Administration (FDA) has published detailed guidelines in draft [78, 79], providing a framework to operating a product safety system covering social media. The FDA has issued at least two warnings to manufacturers in respect of social media activities [80, 81], citing violations of US legislation. In 2009, Neilson [82] found that less than 1% of reports of adverse drug events found on Yahoo Health boards were required to be reported to the FDA. This was due to the conditions required for ICSR reporting and included the inability to identify patients and reporters. While the patient or reporter may not have been immediately identifiable, this raises an important ethical question as to what measures can and should be taken to identify individuals by utilizing additional information such as the geocode location of posting, username and other potentially personally identifiable information.

Ethical implications

Autonomy is a fundamental ethical pillar of modern day medicine and emphasizes the need for people to be involved and in control of their own well-being, if possible. Social media users can generally adjust their privacy settings so that some data are 'private' and that some data are 'open' to the public. Just because someone has not kept their social media posts 'private' does not necessarily mean that they are happy for that information to be used in ways that they have not consented to. Governmental agencies are already known to 'snoop' on the general public for security

reasons, with the 'utilitarianism' argument often cited to justify this [83]. It may be argued that if people are aware of the public nature of their social media behaviour, this is akin to their giving consent [84]. Using social media for pharmacovigilance purposes without informed consent will undoubtedly be a controversial issue and may affect trust in the system.

Non-maleficence is the ethical principle to do no harm, another central component in maintaining trust. However there are scenarios that could potentially bring this into question. As well as the regulatory issues discussed above, it could be argued that there is also an ethical duty on the pharmacovigilance regulators and the MAHs to contact individuals and populations who have suffered ADRs. What happens if an individual suffers haematemesis after taking aspirin? Should it be compulsory to contact individual people if a serious ADR is suspected? Could the regulatory authorities and/or MAHs be liable for negligence and prosecution for not acting [85]? It would be impractical to contact all individuals who have suffered any ADR. However, in order to gain maximal value from reports on social media, it would be beneficial to contact some individuals in order to obtain more information. Where the line is drawn is a difficult decision with potential regulatory, legal and financial implications.

Personal data gathered from the public arena could also potentially have financial ramifications. Insurance companies may invalidate insurance premiums if they find evidence of medical problems in online posts that have not been disclosed [86]. There have been high profile instances of personal data getting in the wrong hands in other industries [83]. As private companies would potentially be interested in these data from an advertising and financial point of view, measures will need to be taken to ensure confidentiality and security of data in order to maintain public trust. Without a transparent system, confidence will be lost by the public and they will not engage with it [84]. Furthermore, it has been suggested that this could affect the doctor–patient relationship [87].

Patient empowerment and improved communication through digital sources has consequences. Healthcare professionals could themselves be at risk of complaints and potential litigation if the public feel that they have suffered from ADRs that they should have been warned about prior to starting the medication. Similar problems could arise for people experiencing ADRs. For instance if a patient tweeted that they had chest pain prior to embarking on a plane which subsequently needed to be diverted for medical reasons, could they be sued for not seeking medical attention as they had symptoms of a potentially serious condition? Similarly could a patient on opioids who reported feeling drowsy on social media prior to crashing their car be liable to prosecution [88, 89]?

From the point of beneficence, if there is potentially important data available in social media, it could be stated there is an ethical obligation to attempt to use this for pharmacovigilance purposes. Using the utilitarianism argument, the duty of care towards these people cannot be ignored. One example is the possibility that Facebook post-analysis could be used to identify and potentially prevent suicides [90]. As only a minority of ADRs are reported, social media could be used to identify signals and generate early warnings more promptly, allowing these to be acted upon before they would be with traditional pharmacovigilance practices. This may reduce morbidity and mortality and also provide invaluable information for certain subsets of society.

An important ethical consideration is that certain subsets of society are less computer literate, lack access to the internet, or have their social media usage censored. Additional measures need to be incorporated to ensure that these subsets are not neglected.

Conclusion and future perspectives

The volume and velocity of data generated from social media sources potentially provide exciting opportunities for advances in pharmacovigilance. However, in order to realize the potential that social media has to offer, a number of challenges remain to be resolved. Some of these are technical in nature while others require careful consideration from regulatory and ethical perspectives to understand fully and yield the benefits that social media have to offer. Indeed, the critical question that needs to be answered is what value social media adds to the current processes of pharmacovigilance, where that value lies, and what processes and regulatory aspects need to be put into place to realize that value.

Qualifying and quantifying the benefit that social media offers to pharmacovigilance will depend on the technological approach to capitalize social media data. In our paper we regard social media as a vast collection of pharmacovigilance sensors, each with their own degree of uncertainty and bias. To realize fully the benefit social media have to offer requires careful combination of each of these data sources to generate significant signals. We present state of the art approaches remaining to be applied for data mining social media for pharmacovigilance in order to achieve this. These same principles can potentially be applied to combine various sources for pharmacovigilance, even in the absence of social media as a signal source.

Current technological challenges include the difficulty for algorithms to interpret layperson technology. Inter-annotator disagreement during the annotation process on interpreting ADRs in social media [22] makes it extremely challenging, if not impossible, for algorithms to interpret postings with certainty. Currently ICSRs require fairly specific information in order to be reported and studies have shown that only a small minority of social media posts appear to ADRs that would necessitate reporting [82]. However this does not mean the remaining information is not valuable. Combining these data effectively could fully unlock the value social media has to offer. Additionally, thought will need to be had regarding ADR detection among polypharmacy patients and how to incorporate these data with other pharmacovigilance sources to optimize signal detection.

The challenges in utilizing social media for pharmacovigilance are not solely technical and require careful contemplation from a regulatory and ethical perspective. Regulators recognize the importance of legislation and guidance regarding social media but there remain many areas of ambiguity requiring further clarification. The legal aspects of the responsibilities of MAHs and regulators towards both groups and individual people reporting ADRs are a potential minefield and need careful consideration.

It could be argued that there is an ethical obligation to attempt to harvest social media data in order to enhance pharmacovigilance and ultimately improve patient safety. However, this raises important and controversial ethical issues. The boundaries of what constitutes public and private data are not comprehensively defined. Does the capability to determine that individuals are potentially at risk impel us to contact them? In the case of social media it may not be immediately possible to identify an individual, though with sophisticated use of other resources it may be possible. The balance of privacy, duty of care and the greater good is complex.

Internet expansion has seen with it the development of mobile technology. This comes in many forms. There has been a massive rise of 'wearable devices'. Smart phones already have the ability to assess geo-location, heart rate, temperature, blood glucose concentration, electrocardiograph, respiration rate, blood oxygen saturation and whether a person has fallen asleep [91, 92]. This could be a goldmine of information for detecting ADRs. This could potentially be extended to wearable electrochemical sensors, where sweat, tears or saliva pharmacokinetic data could be gathered [93]. Additionally, geo-location data collected via map history, social media post location, or via wearable devices could be used to assess when people are attending general practitioners, hospitals or pharmacies which may be indicative of an ADR. As google trend volume for drug-related keywords provides an accurate approximation of actual community use of those drugs [94], geo-location could also provide real-time information that could lead to the identification of localized ADRs and counterfeit drugs earlier than current methods.

In September 2014 the Innovative Medicines Initiative (IMI) funded the WEB-RADR (http://web-radr.eu/) project [95] to conduct scientific research into the use of social

media networks and to develop dedicated applications (Apps) for reporting ADRs to the National Competent Authorities in Europe. The WEB-RADR consortium comprises patient advocates, regulators, academia and members of the pharmaceutical industry working in collaboration. This project has been reviewed by Sukkar [96] and the leaders of this project contemplated various aspects of the potential impact of social media on drug safety.

The rapid development of technology brings with it the opportunity to harvest vast amounts of data, both from social media and emerging biomedical technologies, to allow earlier signal detection, develop early warning systems, and enhance pharmacovigilance and patient safety. If data is harnessed correctly, the field of pharmacovigilance could potentially undergo a revolution.

Competing Interests

There are no competing interests to declare.

REFERENCES

1 Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, Farrar K, Park BK, Breckenridge AM. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ 2004; 329: 15–9.

2 MHRA. Trends in UK spontaneous Adverse Drug Reaction (ADR) reporting between 2008–2012 [Internet]. MHRA; 2014. Available at http://webarchive.nationalarchives.gov.uk/20141205150130/http://www.mhra.gov.uk/home/groups/pl-p/documents/websiteresources/con408250.pdf (last accessed 1 June 2015).

3 Pew Research Center: Internet ST. Health Online 2013 [Internet]. Pew Research Center; 2015. Available at http://www.pewinternet.org/2013/01/15/health-online-2013/ (last accessed 1 June 2015).

4 Statistica. Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2015 [Internet]. Statistica; 2015. Available at http://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/ (last accessed 1 June 2015).

5 Statistica. Number of monthly active Facebook users worldwide from 1st quarter 2010 to 1st quarter 2015 [Internet]. Statistica; 2015. Available at http://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ (last accessed 1 June 2015).

6 Twitter. Welcome to Twitter [Internet]. Twitter; 2015. Available at https://twitter.com/ (last accessed 1 June 2015).

7 Facebook. Facebook [Internet]. Facebook; 2015. Available at https://www.facebook.com/ (last accessed 1 June 2015).

8 Sakaki T, Okazaki M, Matsuo Y. Earthquake shakes Twitter users: real-time event detection by social sensors. In: Proceedings of the 19th international conference on World wide web. Raleigh, North Carolina, USA: ACM, 2010; 851–60.

9 Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. Am J Trop Med Hyg 2012; 86: 39–45.

10 Nikfarjam A, Gonzalez GH. Pattern mining for extraction of mentions of adverse drug reactions from user comments. In: AMIA Annual Symposium Proceedings: American Medical Informatics Association, 2011: 1019–1026.

11 Chee BW, Berlin R, Schatz B. Predicting adverse drug events from personal health messages. In: AMIA Annual Symposium Proceedings: American Medical Informatics Association, 2011: 217–26.

12 Yang CC, Yang H, Jiang L, Zhang M. Social media mining for drug safety signal detection. In: Proceedings of the 2012 international workshop on Smart health and wellbeing: ACM, 2012: 33–40.

13 Liu X, Chen H. AZDrugMiner: an information extraction system for mining patient-reported adverse drug events in online patient forums. Smart Health: Springer, 2013; 134–50.

14 Liu X, Liu J, Chen H. Identifying Adverse Drug Events from Health Social Media: A Case Study on Heart Disease Discussion Forums. Smart Health: Springer, 2014; 25–36.

15 Patki A, Sarker A, Pimpalkhute P, Nikfarjam A, Ginn R, O'Connor K, Smith K, Gonzalez G. Mining adverse drug reaction signals from social media: going beyond extraction. Proceedings of BioLinkSig 2014; 2014: 1–8.

16 Yang CC, Yang H, Jiang L. Postmarketing Drug Safety Surveillance Using Publicly Available Health-Consumer-Contributed Content in Social Media. ACM Trans Mgt Inf Sys 2014; 5: Article 2, 21 pages.

17 Sampathkumar H, Chen XW, Luo B. Mining adverse drug reactions from online healthcare forums using hidden Markov model. BMC Med Inform Decis Mak 2014; 14: 91. 1–18.

18 Knezevic MZ, Bivolarevic IC, Peric TS, Jankovic SM. Using Facebook to increase spontaneous reporting of adverse drug reactions. Drug Saf 2011; 34: 351–2.

19 Bian J, Topaloglu U, Yu F. Towards large-scale twitter mining for drug-related adverse events. Association for Computing Machinery. Proceedings of the 2012 international workshop on Smart health and wellbeing 2012: 25–32.

20 Jiang K, Zheng Y. Mining twitter data for potential drug effects. In: Advanced Data Mining and Applications, eds Motoda H, Wu Z, Cao L, Zaiane O, Yao M, Wang W. Berlin Heidelberg: Springer, 2013; 8346: 434–43.

21 Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, Dasgupta N. Digital drug safety surveillance: monitoring pharmaceutical products in twitter. Drug Saf 2014; 37: 343–50.

22 Ginn R, Pimpalkhute P, Nikfarjam A, Patki A, O'Connor K, Sarker A, Gonzalez G. Mining Twitter for adverse drug reaction mentions: a corpus and classification benchmark. LREC BioTexM. Proceedings of the fourth workshop on building and evaluating resources for health and biomedical text processing, 2014; 2: 1–8.

23 O'Connor K, Pimpalkhute P, Nikfarjam A, Ginn R, Smith KL, Gonzalez G. Pharmacovigilance on Twitter? Mining Tweets for adverse drug reactions. AMIA Annual Symposium Proceedings: American Medical Informatics Association, 2014: 924–933.

24 Sarker A, Gonzalez G. Portable automatic text classification for adverse drug reaction detection via multi-corpus training. J Biomed Inform 2015; 53: 196–207.

25 Nikfarjam A, Sarker A, O'Connor K, Ginn R, Gonzalez G. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. J Am Med Inform Assoc 2015; 22: 671–82.

26 Benton A, Ungar L, Hill S, Hennessy S, Mao J, Chung A, Leonard CE, Holmes JH. Identifying potential adverse effects using the web: A new approach to medical hypothesis generation. J Biomed Inform 2011; 44: 989–96.

27 Hadzi-Puric J, Grmusa J. Automatic drug adverse reaction discovery from parenting websites using disproportionality methods. Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012): IEEE Computer Society, 2012: 792–97.

28 Yates A, Goharian N. ADRTrace: detecting expected and unexpected adverse drug reactions from user reviews on social media sites. In: Advances in Information Retrieval, eds Serdyukov P, Braslavski P, Kuznetsov S, Kamps J, Rüger S, Agichtein E, Segalovich I, Yilmaz E. Berlin Heidelberg: Springer-Verlag, 2013; 7814: 816–19.

29 Yeleswarapu S, Rao A, Joseph T, Saipradeep VG, Srinivasan R. A pipeline to extract drug-adverse event pairs from multiple data sources. BMC Med Inform Decis Mak, 2014; 14: 13.

30 Segura-Bedmar I, Revert R, Martínez P. Detecting drugs and adverse events from Spanish health social media streams. Proceedings of the 5th International Workshop on Health Text Mining and Information Analysis (Louhi), EACL, April 2014: 106–15.

31 Leaman R, Wojtulewicz L, Sullivan R, Skariah A, Yang J, Gonzalez G. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. Proceedings of the 2010 Workshop on Biomedical Natural Language Processing, Uppsala, Sweden: Association for Computational Linguistics, 2010: 117–25.

32 Patientslikeme. Patientslikeme [Internet]. Patientslikeme; 2015. Available at https://www.patientslikeme.com/ (last accessed 1 June 2015).

33 DailyStrength. Online Support Groups and Forums at DailyStrength [Internet]. DailyStrength; 2015. Available at http://www.dailystrength.org (last accessed 1 June 2015).

34 MedHelp. All Ask a Doctor Forums and Medical Commmunities - MedHelp [Internet]. MedHelp; 2015. Available at http://www.medhelp.org/forums/list (last accessed 1 June 2015).

35 Sarker A, Ginn R, Nikfarjam A, O'Connor K, Smith K, Jayaraman S, Upadhaya T, Gonzalez G. Utilizing social media data for pharmacovigilance: A review. J Biomed Inform 2015; 54: 202–12.

36 Harpaz R, Callahan A, Tamang S, Low Y, Odgers D, Finlayson S, Jung K, LePendu P, Shah NH. Text mining for adverse drug events: the promise, challenges, and state of the art. Drug Saf 2014; 37: 777–90.

37 Pimpalkhute P, Patki A, Nikfarjam A, Gonzalez G. Phonetic spelling filter for keyword selection in drug mention mining from social media. AMIA Summits Trans Sci Proc 2014; 2014: 90–95.

38 Bollegala D, Matsuo Y, Ishizuka M. Automatic discovery of personal name aliases from the web. IEEE Trans Knowledge Data Eng 2011; 23: 831–44.

39 Food and Drugs Administration. FDA Adverse Event Reporting System (FAERS) [Internet]. Food and Drugs Administration. Available at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm (last accessed 1 June 2015).

40 COSTART. Coding Symbols for a Thesaurus of Adverse Reaction Terms - Summary [Internet]. NCBO BioPortal; 2015. Available at http://bioportal.bioontology.org/ontologies/COSTART (last accessed 1 June 2015).

41 Initiative CHV. Open Access, Collaborative Consumer Health Vocabulary Initiative [Internet]. Available at http://www.consumerhealthvocab.org/ (last accessed 1 June 2015)

42 Medeffect Canada. MedEffect Canada [Internet]. Health Canada; 2015. Available at http://www.hc-sc.gc.ca/dhp-mps/medeff/index-eng.php (last accessed 1 June 2015).

43 UMLS. Unified Medical Language System® (UMLS®) [Internet]. US National Library of Medicine; 2015. Available at http://www.nlm.nih.gov/research/umls/ (last accessed 1 June 2015).

44 MedDRA. Welcome to MedDRA [Internet]. MedDRA. Available at http://www.meddra.org (last accessed 1 June 2015).

45 SIDER. SIDER Side Effect Resource [Internet]. SIDER; 2012. Available at http://sideeffects.embl.de (last accessed 1 June 2015).

46 Yang M, Kiang M, Shang W. Filtering big data from social media - Building an early warning system for adverse drug reactions. J Biomed Inform 2015; 54: 230–40.

47 Zhang C, Chen T. An active learning framework for content-based information retrieval. IEEE Trans Multimedia 2002; 4: 260–68.

48 Bollegala D, Kontonatsios G, Ananiadou S. A cross-lingual similarity measure for detecting biomedical term translations. PLoS One 2015; 10e0126196. doi:10.1371/journal.pone.0126196.

49 Hauben M, Zhou X. Quantitative methods in pharmacovigilance. Drug Saf 2003; 26: 159–86.

50 Hazell L, Shakir SA. Under-reporting of adverse drug reactions: a systematic review. Drug Saf 2006; 29: 385–96.

51 de Boissieu P, Kanagaratnam L, Abou Taam M, Roux MP, Drame M, Trenque T. Notoriety bias in a database of spontaneous reports: the example of osteonecrosis of the jaw under bisphosphonate therapy in the French national pharmacovigilance database. Pharmacoepidemiol Drug Saf 2014; 23: 989–92.

52 Moore N, Hall G, Sturkenboom M, Mann R, Lagnaoui R, Begaud B. Biases affecting the proportional reporting ratio (PPR) in spontaneous reports pharmacovigilance databases: the example of sertindole. Pharmacoepidemiol Drug Saf 2003; 12: 271–81.

53 Nguyen D, Gravel R, Trieschnigg D, Meder T. "How old do you think I am?" A study of language and age in Twitter. Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media: AAAI Press, 2013; 439–48.

54 Mislove A, Lehmann S, Ahn Y-Y, Onnela J-P, Rosenquist JN. Understanding the demographics of Twitter Users. ICWSM 2011; 11: 5. 1–4.

55 Huang J, Gretton A, Borgwardt KM, Schölkopf B, Smola AJ. Correcting sample selection bias by unlabeled data. Advances in neural information processing systems, 2006: 601-08.

56 Blehar MC, Spong C, Grady C, Goldkind SF, Sahin L, Clayton JA. Enrolling pregnant women: issues in clinical research. Womens Health Issues 2013; 23: 39–45.

57 Field M, Berman R. Institute of Medicine (US). Committee on Clinical Research Involving Children. Ethical conduct of clinical research involving children. Eds Field M, Berman R. Washington, DC: National Academies Press, 2004.

58 Pew Research Center: Internet ST. Demographics of key social networking platforms. [Internet]. Pew Research Center; 2013. Available at http://www.pewinternet.org/2013/12/30/demographics-of-key-social-networking-platforms/ (last accessed 1 June 2015).

59 Yang J-B, Xu D-L. Evidential reasoning rule for evidence combination. Artif Intell 2013; 205: 1–29.

60 Kim YA, Ahmad MA. Trust, distrust and lack of confidence of users in online social media-sharing communities. Know-Based Syst 2013; 37: 438–50.

61 Harpaz R, Vilar S, Dumouchel W, Salmasian H, Haerian K, Shah NH, Chase HS, Friedman C. Combing signals from spontaneous reports and electronic health records for detection of adverse drug reactions. J Am Med Inform Assoc 2013; 20: 413–9.

62 Harpaz R, DuMouchel W, LePendu P, Shah NH. Empirical bayes model to combine signals of adverse drug reactions. In: Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining. Chicago, Illinois, USA: ACM, 2013; 1339–47.

63 LinkedIn. Join the world's largest professional network [Internet]. LinkedIn; 2015. Available at https://www.linkedin.com/nhome/ (last accessed 1 June 2015).

64 YouTube. YouTube GB [Internet]. Google; 2015. Available at https://www.youtube.com/ (last accessed 1 June 2015).

65 Flipboard. Flipboard is Your Personal Magazine [Internet]. Flipboard; 2015. Available at https://flipboard.com/ (last accessed 1 June 2015).

66 Flickr. The home for all your photos. [Internet]. Yahoo; 2015. Available at https://www.flickr.com/ (last accessed 1 June 2015).

67 Pinterest. What's Pinterest? [Internet]. Pinterest; 2015. Available at https://about.pinterest.com/en-gb (last accessed 1 June 2015).

68 Twitter. Pharmaceutical News (@PharmaIQ) on Twitter [Internet]. Twitter. Available at https://twitter.com/pharmaiq (last accessed 1 June 2015).

69 Twitter. Novartis (@Novartis) on Twitter [Internet]. Twitter. Available at https://twitter.com/novartis (last accessed 1 June 2015).

70 Twitter. Pfizer Inc. (@pfizer) on Twitter [Internet]. Twitter. Available at https://twitter.com/pfizer (last accessed 1 June 2015).

71 Twitter. Roche (@Roche) on Twitter [Internet]. Twitter. Available at https://twitter.com/roche (last accessed 1 June 2015).

72 ICH Expert Working Group. ICH Harmonised Tripartite Guideline. Clinical safety data management: definitions and standards for expedited reporting E2A. European Medicines Agency. 27 October 1994. Available at http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf (last accessed 1 June 2015).

73 Edwards IR, Lindquist M. Social media and networks in pharmacovigilance: boon or bane? Drug Saf 2011; 34: 267–71.

74 European Commission. Commission Implementing Regulation 520/2012. Off J Eur Union 2012; 55: 5–25. Available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:159:0005:0025:EN:PDF (last accessed 1 June 2015).

75 European Commission. Directive 2001/83/EC of the European Parliament and of the Council. Official J Eur Comm 2001; 44: 67–128. Available at http://www.edctp.org/fileadmin/documents/ethics/DIRECTIVE_200183EC_OF_THE_EUROPEAN_PARLIAMENT.pdf (last accessed 1 June 2015).

76 European Commission. Regulation 726/2004. Off J Eur Union 2010; 47: 1–33. Available at http://ec.europa.eu/health/files/eudralex/vol-1/reg_2004_726/reg_2004_726_en.pdf (last accessed 1 June 2015).

77 European Medicines Agency. Good Pharmacovigilance Practices, Module VI, Revision 1, Management and reporting of adverse reactions to medicinal products. EMA/873138/2011. 08/09/2014. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/09/WC500172402.pdf (last accessed 1 June 2015).

78 Food and Drugs Administration. Draft guidance: Internet/Social Media Platforms with Character Space Limitations— Presenting Risk and Benefit Information for Prescription Drugs and Medical Devices. [Internet]. Food and Drugs Administration; June 2014. Available at http://www. fda.gov/downloads/drugs/guidancecomplianceregulatory information/guidances/ucm401087.pdf (last accessed 1 June 2015).

79 Food and Drugs Administration. Draft guidance: Internet/Social Media Platforms: Correcting Independent Third-Party Misinformation About Prescription Drugs and Medical Devices. [Internet]. Food and Drugs Administration; June 2014. Available at http://www.fda. gov/downloads/drugs/guidancecomplianceregulatory information/guidances/ucm401079.pdf (last accessed 1 June 2015).

80 Food and Drugs Administration. Inspections, Compliance, Enforcement, and Criminal Investigations. Zarbee's, Inc. [Internet]. Food and Drugs Administration; 2014. Available at http://www.fda.gov/ICECI/EnforcementActions/ WarningLetters/2014/ucm403255.htm (last accessed 1 June 2015).

81 Food and Drugs Administration. Tirosint letter [Internet]. Food and Drugs Administration; 2014. Available at http:// www.fda.gov/downloads/Drugs/GuidanceCompliance RegulatoryInformation/EnforcementActivitiesbyFDA/Warning LettersandNoticeofViolationLetterstoPharmaceuticalCompanies/ UCM388800.pdf (last accessed 1 June 2015).

82 Nielsen. Listening to consumers in a highly regulated environment. How pharmaceutical manufacturers can leverage consumer-generated media. [Internet]. The Nielsen Company; August 2008. Available at http://www. nielsen.com/content/dam/corporate/us/en/newswire/ uploads/2009/11/Nielsen-Online-Healthcare-Practice_Social- Media-Adverse-Event-Reporting_nov09.pdf (last accessed 1 June 2015).

83 Barnard-Wills D. UK news media discourses of surveillance. Sociol Quarterly 2011; 52: 548–67.

84 Swirsky ES, Hoop JG, Labott S. Using social media in research: new ethics for a new meme? Am J Bioeth 2014; 14: 60–1.

85 Taylor HA, Kuwana E, Wilfond BS. Ethical implications of social media in health care research. Am J Bioeth 2014; 14: 58–9.

86 Financial Ombudsman Service. Ombudsman News. [Internet]. Financial Ombudsman Service; May/June 2005. Issue 46. Available at http://www.financial-ombudsman.org. uk/publications/ombudsman-news/46/46.pdf (last accessed 1 June 2015).

87 Farnan JM. Connectivity and consent: does posting imply participation? Am J Bioeth 2014; 14: 62–3.

88 Department for Transport. Drug drive legislation: am I fit to drive? [Internet]. UK Government; February 2015. Available at https://www.gov.uk/government/news/drug- drive-legislation-am-i-fit-to-drive (last accessed 1 June 2015).

89 Kress HG, Kraft B. Opioid medication and driving ability. Eur J Pain 2005; 9: 141–44.

90 Kailasam VK, Samuels E. Can social media help mental health practitioners prevent suicides? Anecdotal evidence suggests that analyzing Facebook posts can lead to earlier intervention. Curr Psych 2015; 14: 37–51.

91 Pandian PS, Mohanavelu K, Safeer KP, Kotresh TM, Shakunthala DT, Gopal P, Padaki VC. Smart Vest: wearable multi-parameter remote physiological monitoring system. Med Eng Phys 2008; 30: 466–77.

92 Yilmaz T, Foster R, Hao Y. Detecting vital signs with wearable wireless sensors. Sensors 2010; 10: 10837–62.

93 Bandodkar AJ, Wang J. Non-invasive wearable electrochemical sensors: a review. Trends Biotechnol 2014; 32: 363–71.

94 Simmering JE, Polgreen LA, Polgreen PM. Web search query volume as a measure of pharmaceutical utilization and changes in prescribing patterns. Res Social Adm Pharm 2014; 10: 896–903.

95 WEB-RADR. WEB-RADR: Recognising adverse drug reactions. [Internet]. WEB-RADR; 2015. Available at http:// web-radr.eu/ (last accessed 1 June 2015).

96 Sukkar E. Searching social networks to detect adverse events. Pharmaceut J 2015; 294: 75–78.

Glossary of terms

Text mining is the process of using algorithms to derive high quality useful information from text.

Machine learning is the process of constructing algorithms that can learn from input data typically with a view to make decisions based on previously unseen data. Machine learning algorithms broadly fall within two categories: supervised learning and unsupervised learning.

Supervised machine learning involves the process of a human 'teacher' providing the desired outcome for each item of input data. This would typically involve a desired target variable for the data and within the context of pharmacovigilance could simply be a binary value indicating whether or not a specific sentence contains an ADR.

Active learning is where the sequence of annotation tasks is defined interactively with the machine learning algorithm to explicitly remove ambiguities perceived by the machine learning algorithm.

Unsupervised machine learning is when the algorithms seek to find some form of structure within the data without being specifically trained by a human as to the desired outcome. In the context of text mining this could

be the process of organising documents into collections which appear to be similar in nature based on their content.

NER Named entity recognition is an information extraction task that seeks to locate and classify elements in text into predefined categories. In the context of text mining social media this could involve finding all drug aliases for an active ingredient including colloquialisms and slang.

ICSR Individual case safety report refers to the format and content for the reporting of one or several suspected adverse reactions in relation to a medicinal product that occur in a single patient at a specific point of time. A valid ICSR should include at least one identifiable reporter, one single identifiable patient, at least one suspect adverse reaction and at least one suspect medicinal product.