Journal of the American Medical Informatics Association Advance Access published December 11, 2015

Tweeting back: predicting new cases of back pain with mass social media data

RECEIVED 12 July 2015
REVISED 24 September 2015
ACCEPTED 7 October 2015

Hopin Lee[1,2], James H McAuley[1,2], Markus Hübscher[1,2], Heidi G Allen[1], Steven J Kamper[3,4], G Lorimer Moseley[1,5]

ABSTRACT

Background Back pain is a global health problem. Recent research has shown that risk factors that are proximal to the onset of back pain might be important targets for preventive interventions. Rapid communication through social media might be useful for delivering timely interventions that target proximal risk factors. Identifying individuals who are likely to discuss back pain on Twitter could provide useful information to guide online interventions.

Methods We used a case-crossover study design for a sample of 742 028 tweets about back pain to quantify the risks associated with a new tweet about back pain.

Results The odds of tweeting about back pain just after tweeting about selected physical, psychological, and general health factors were 1.83 (95% confidence interval [CI], 1.80-1.85), 1.85 (95% CI: 1.83-1.88), and 1.29 (95% CI, 1.27-1.30), respectively.

Conclusion These findings give directions for future research that could use social media for innovative public health interventions.

Keywords: back pain, Twitter, public health, social media, case-crossover

BRIEF COMMUNICATION

INTRODUCTION

Back pain is a major public health problem and the leading cause of disability worldwide.[1] The societal burden of back pain is enormous – the annual direct costs of treating back pain are approximately $86 billion in the United States,[2] and indirect costs through lost productivity take the burden well past $114 billion.[3]

Risk factors such as age, low education level, and occupational demands are associated with the development of back pain.[4] However, some of these factors are difficult to modify. Evidence also suggests that preventive strategies that target modifiable risk factors, such as workplace interventions, have not been successful in reducing the societal burden of back pain.[5] Risk factors that are more proximal to the onset of back pain are likely to play a critical role in back pain, because ~80% of patients who experience a new episode of back pain report a sudden onset of symptoms.[6] A recent study showed that exposure to proximal risk factors, eg, manual tasks involving heavy loads, and fatigue, increased the likelihood of the onset of back pain.[7] The authors of that study concluded that proximal risk factors might be suitable targets for interventions aimed at reducing the incidence of low back pain. However, for proximal risk factors, the time-span between exposure and the onset of back pain is brief. New strategies that can reach individuals in a timely manner may offer new directions for public health interventions to reduce the incidence of back pain.

Social media platforms such as Twitter enable rapid, cost-effective communication to mass populations. Every day, there are approximately 100 million active Twitter users who send short messages called "tweets" to a worldwide audience.[8] Many people share their health status via social media,[9–11] and the potential to use such data has been recognized; first by compiling descriptive accounts of what people report,[12,13] and then by using sophisticated analyses to map the spread of infectious diseases, such as influenza.[9,14] Given the high prevalence of back pain in the general population, it is unsurprising that people discuss their back pain on social media.[10] Early identification of individuals who are likely to report back pain via social media could provide useful information to guide online public health interventions for back pain. We used an epidemiological study design (case-crossover) to investigate whether tweets about potential back pain risk factors can predict user tweets about back pain.

METHODS

The University of South Australia's Ethics Committee approved this study.

Sample

The data for this study were collected by a third-party analytics service (PeopleBrowsr, San Francisco, CA) that has access to all archived Tweets posted since 2008.[15] The data used in this study were tweets posted between 2010 and 2012, unrestricted by geographic location. Data were filtered to exclude URLs and retweets by excluding the following terms: "http://," "https://," and "RT." We identified "case tweets" by searching the archived tweets for phrases, in English, that relate to back pain (eg, *painful back*, *pain in my back*, *sore back*, *hurt my back*, *I've got backache*, *injured my back*, *buggered my back*, *my back is killing me*, *I've got back pain*, *put my back out*, *my back hurts*, *back started hurting*). We used first-person pronouns to tailor the search to specifically identify individuals who were tweeting about their own back pain. Those who tweeted about back pain, but had not tweeted about back pain in the previous 3 months, were included in the study. Because our aim was to identify new cases of back pain, we modified the de Vet definition of acute back pain[16] – a single tweet about back pain preceded by 3 months without a tweet that mentioned back pain was defined as a "new episode" of back pain. Those individuals tweeting about ongoing back pain (ie, who had posted more than one tweet about back pain within a 3-month period) or tweeting about back pain in a language other than English were excluded; no

Correspondence to G Lorimer Moseley, Sansom Institute for Health Research, GPO Box 2471, Adelaide 5001, Australia; Lorimer.Moseley@unisa.edu.au; Tel: +610883022454

 For numbered affiliations see end of article.

limits were placed on individuals' age. Demographic information (gender) was collected from the users' profiles.

Design

We used a retrospective case-crossover design. Case-crossover studies are suited to identifying proximal risk factors that immediately precede the onset of a condition or injury (in this study, this was identified by a new tweet about back pain). Essentially, case-crossover studies are used to answer the question "was the event/injury triggered by something unusual that happened just before?".[17] This self-matching method, in which the study participants act as their own control population, means that estimates, by design, are adjusted for confounders (eg, age, gender, number of previous episodes).[18] Controlling for these confounders gives case-crossover studies an advantage over traditional case-control studies for studying transient risk factors, as has been demonstrated in studies of myocardial infarctions[19] and occupational hand injuries.[20]

We specified a "hazard period" – the 48-hour window that preceded each case tweet – and two "control periods" of 48-hours each – one that preceded the case tweet by 2 weeks and one that preceded it by 2 months (Figure 1). By limiting our inclusion criteria to individuals who had posted one back pain episode over a 3-month period, we established two control periods, during which individuals did not tweet about back pain. Within the hazard and control periods, we then searched for "risk tweets" that contained predefined phrases about the individuals' physical, psychological, or general health status that represented possible risk factors for back pain. Our aim was to utilize a comprehensive group of risk-related terms to identify tweets that represented three broad, underlying domains: risky physical behavior, a negative psychological state, and poor general health. These domains reflect existing theories that integrate physical, psychological, and biological determinants of pain.[21,22] We used the following terms to detect an exposure to physical risk factors: *gym*, *lifting*, *lifted*, *lift*, *bending*, *bent*, *went for a run*, *jogging*, *went for a jog*, *went running*, *training for a half marathon*, *training for a marathon*, *doing weights*, *treadmill*, *aerobics*, *stepercise*, *dancing*, *power walking*, *on the cross-trainer*, *yoga*, *pilates*, *throwing*, *bowling*, *played golf*, *kicked a footie*, and *played soccer/squash/tennis*; psychological risk factors: *depressed*, *stressed*, *bad mood*, *feeling down*, *angry*, *tired*, *exhausted*, *washed out*, *anxious*, and *fed-up*; and general health risk factors: *sick*, *ill*, *flu*, *crook*, *cold*, *got drunk/tanked/hammered/pissed*, *under the weather*, *shattered*, and *not sleeping*. This list of potential risk factors was derived by an expert group of researchers and clinicians in the field of back pain. Our aim was to create a comprehensive list of possible risk factors that have been implicated in triggering a new back pain episode. This process was guided by existing evidence on proximal back pain triggers and clinical expertise. The list was checked by an independent group of researchers, and, based on their review, additional terms were added.

We measured the frequency of physical, psychological, and general health risk tweets posted during the hazard period and the two control periods of each case tweet. From these measurements, we generated the frequencies of discordant pairs. "Hazard discordant pairs" were the number of individuals who tweeted about a risk factor during the hazard period but not during the control period, and "control discordant pairs" were the number of individuals who tweeted about a risk factor during the control period but not during the hazard period.

Analysis

We compared the exposure (ie, the presence of risk tweets) in the 48-hour period immediately before the case tweet was posted (ie, the hazard period) with the 48-hour control period 2 weeks before the case tweet (ie, the control period). We also made the same comparison using the 2-month control period. We used multiple conditional logistic regression to calculate independent odds ratios and 95% confidence intervals for the physical, psychological, and general health risk tweets. Analyses were conducted using the "*clogit*" command in STATA 13 software (StataCorp., College Station, TX). Because odds ratios can inflate small changes in probabilities, we conducted alternative analyses to estimate positive predictive values (PPVs). The PPVs were calculated as the proportion of individuals who tweeted about a risk factor during the hazard period and tweeted about back pain (ie, frequency of true positive/frequency of positive risk).

RESULTS

Over 2 years 742 028 individuals, who each posted a single case tweet, were identified. In the case tweets, individuals had posted a phrase, in English, that reported a transient episode of back pain, and these individuals had not tweeted about another back pain episode in the previous 3 months. A random sample of case tweets is provided in the Supplementary Appendix. In the total sample, 219 694 individuals were female, 126 052 individuals were male, and 396 282 individuals did not disclose their gender in their profile.

Figure 1 depicts the two types of discordant pairs: the hazard discordant pair, in which the individual tweeted about a risk factor during

Figure 1: Case-crossover study design. The red oval represents the "hazard period", defined as a 48-hour period immediately preceding the case tweet. The green ovals depict the "control periods", defined as the 2 weeks and 2 months preceding the case tweet.

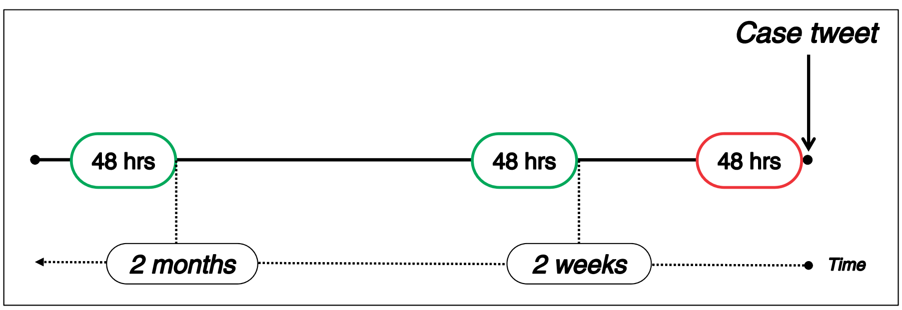
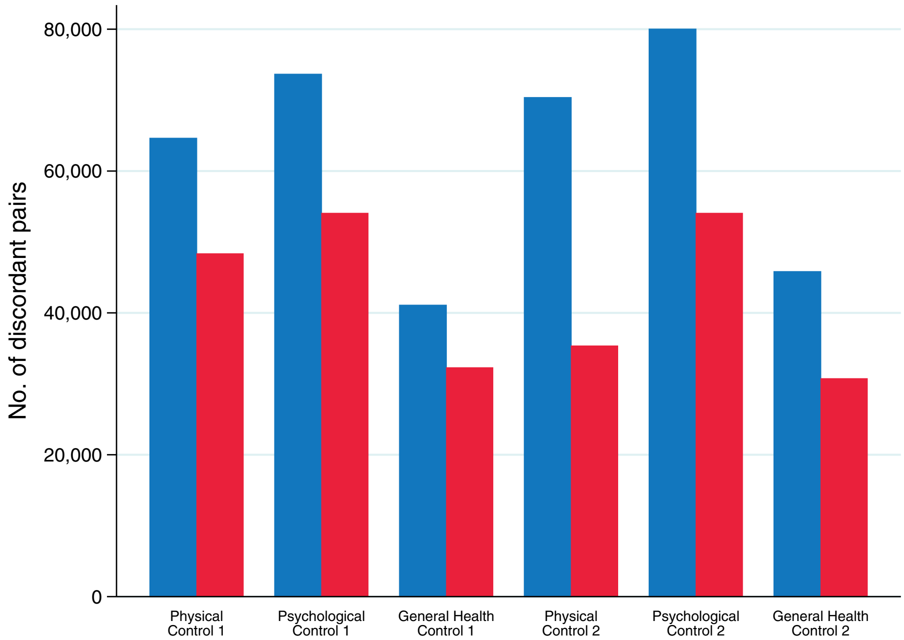


Figure 2: Frequencies of discordant pairs for the 2-week (Control 1) and 2-month (Control 2) control comparisons. The blue bars represent the hazard discordant pairs (ie, a risk factor tweeted during the hazard period but not during the control period). The red bars represent the control discordant pairs (ie, a risk factor tweeted during the control period but not during the hazard period).



the hazard period (ie, the 48-hour period before they tweeted about back pain) but not during the control period (ie, the 48 hour period either 2 weeks or 2 months preceding the case tweet), and the control discordant pair, in which the individual tweeted about a risk factor during the control period but not during the hazard period. Figure 2 presents the frequencies of these discordant pairs for each risk tweet domain (physical, psychological, and general health). We observed higher frequencies of hazard discordant pairs than control discordant pairs for all the risk tweet domains. This shows that there were more individuals who tweeted about a risk factor immediately prior to tweeting about their back pain, than people who tweeted about a risk factor and did not subsequently tweet about their back pain. This trend was observed in both the 2-week control comparison and the 2-month control comparison.

In Table 1, we present the odds ratios and PPVs for each risk tweet for the 2-week and 2-month control comparisons. The odds ratios represent the odds of an individual tweeting about a risk factor prior to tweeting about back pain compared to the odds of an individual tweeting about a risk factor and not subsequently tweeting about back pain. For the 2-week control comparison, tweets about psychological risk factors showed the highest odds ratios, followed by tweets about physical risk factors, then tweets about general health risk factors. A similar trend was observed for the 2-month control comparison. The odds ratios for the 2-month control comparison were greater than the odds ratios observed for the 2-week control comparisons. This trend was more pronounced for risk tweets about psychological and physical risk factors. The PPVs represent the likelihood that an individual will tweet about back pain if they tweet about a given risk factor. The PPVs follow a trend that is similar to that of the odds ratios. For the 2-week control comparison, tweets about psychological risk factors were associated with the highest likelihood of individuals subsequently tweeting about back pain, followed by tweets about physical risk factors, then tweets about general health risk factors. For the 2-month control comparison, tweets about psychological and physical risk factors showed similar associations with the likelihood of individuals subsequently tweeting about back pain, followed by tweets about general health risk factors.

DISCUSSION

In this case-crossover study, we used a large database of archived tweets from 742 028 individuals and found that antecedent tweets about psychological, physical, and general health risk factors increased the odds of an individual posting a new tweet about back pain. Our alternative analyses of PPVs showed trends that were similar to the odds ratios we calculated. Although the advantage of a PPV is that it provides an easily interpretable estimate of the predictive value of each risk tweet, the PPVs we calculated should be interpreted with caution, because they do not take into consideration the matched pairs of the case-crossover design.

Recent research has shown that exposure to modifiable proximal risk factors, such as manual tasks involving awkward positioning, and fatigue, were associated with a higher risk of an individual developing back pain.[7] The terms we used to identify physical and psychological risk factors mirrored the key proximal risk factors for back pain onset identified by Steffens et al. – eg, “lifting”, “tired”, and “exhausted”. We found higher odds ratios for psychological risk factors than we did for physical risk factors. This might be due to the fact that people are more likely to tweet about their psychological state than about ongoing physical activities (such as lifting). It is also possible that psychological symptoms aggregate a range of negative sentiments[23] and last longer than physical or general health symptoms. Although this might have increased the point-frequencies of tweets about psychological risk factors

BRIEF COMMUNICATION

Table 1: Odds Ratios, PPVs, and their 95% CIs for Physical, Psychological, and General Health Risk Tweets

Control period	Risk tweet domain	Hazard discordant pairs[a]	Control discordant pairs[b]	Odds ratio (95% CI)	PPV% (95% CI)
2 weeks	Physical	64 618	48 366	1.30 (1.28-1.31)	55.71 (55.45-55.97)
	Psychological	73 646	54 072	1.33 (1.31-1.34)	56.03 (55.79-56.27)
	General health	41 082	32 296	1.22 (1.20-1.23)	54.55 (54.23-54.86)
2 months	Physical	70 356	35 358	1.83 (1.80-1.85)	64.17 (63.90-64.44)
	Psychological	80 006	39 937	1.85 (1.83-1.88)	64.14 (63.89-64.39)
	General health	45 816	30 757	1.29 (1.27-1.30)	58.34 (58.02-58.66)

CI, confidence interval; PPV, positive predictive value. [a]Frequency of individuals who tweeted about a risk factor during the hazard period but not during the control period. [b]Frequency of individuals who tweeted about a risk factor during the control period but not during the hazard period.

and inflated the number of concordant pairs, this does not influence the odds ratios, because they are derived from discordant pairs. Thus, the case-crossover design of our study protects against this source of bias.

Only a few studies, all of which have been descriptive content analyses, have investigated pain-related tweets.[10,23,24] Our study extends the findings of these studies by applying a well-established epidemiological study design to identify proximal risk tweets that predict subsequent tweets about back pain. By using the case-crossover study design, the effect of confounding is minimized, because participants are self-matched. We also limited recall bias, a fundamental problem of most retrospective case-crossover studies,[25] by using an archive of unsolicited, self-reported tweets. These tweets about back pain were searched using phrases that included first person pronouns (eg, “my back is killing me”). Therefore, it is likely that our case tweets were specifically tweets about back pain in the first person context. We also think it is reasonable to suggest that individuals were tweeting about their current back pain in such tweets, rather than a previous episode of back pain. Our selected search phrases and real-time reporting on Twitter suggest that individuals were experiencing back pain at the time of tweeting and were not forecasting future events (see Supplementary Appendix).

There are some limitations to our study, if one attempts to generalize our findings to real cases of back pain. Those individuals who report back pain via Twitter in an unsolicited manner may not truly reflect back pain cases as defined in traditional studies. We did not verify whether individuals who tweeted about back pain actually had back pain (eg, via medical records); such data are not yet available for tweets about back pain. This limitation also applies to risk tweets and demographic data. In other fields, such as human immunodeficiency virus (HIV) research, validation work has shown that databases of tweets about HIV represent geographically defined databases of real cases.[26] Similar work in the back pain field would improve our understanding about the feasibility and validity of using social media data, to generalize our findings to real back pain cases. It is also possible that Twitter users may not truly represent the general population of individuals who experience back pain. This limits the generalizability of our findings. We recognize that our risk factor terms for possible back pain triggers may not be exhaustive and that this might have introduced bias and underestimated the observed effects in our study. Future work in this area could evaluate the predictive sensitivity and specificity of individual risk factors. Reverse causality (ie, tweets about back pain causing tweets about risk factors) and the progression of back pain (via linguistic analysis) are other issues that warrant further investigation. In this study, it is plausible that previous episodes of back pain may have influenced the frequencies of observed risk tweets.

In other areas of healthcare, epidemiologists have called for innovative approaches that use online social media to deliver preventive health interventions.[27–30] Recent meta-analyses have also shown that behavioral interventions delivered through social media have positive effects on health outcomes.[31] For the prevention of back pain, informatics-based interventions could utilize real-time surveillance methods to identify “at risk” individuals, then direct them to online educational material[32] to improve their awareness and to address potential risks that are associated with back pain. For those individuals who develop back pain, targeted guideline advice[33] and reassurance[34] could be delivered via direct messaging services, accompanied by URL links to online interventions with dedicated web pages[35,36] and social media groups.[37,37] Refined data-filtering algorithms[14,39] and geolocation analyses[14,40] would further advance this area of research, and prospective-retrospective studies will help us understand the causal sequence, and appropriate timing to guide online interventions.

CONCLUSION

Considering the massive societal burden of back pain, this study provides a way forward for identifying “at risk” individuals through social media, which lays the platform from which we might develop targeted interventions that can be delivered via social media.

CONTRIBUTORS

Study concept and design: G.L.M., J.H.M., H.L., M.H.; Data acquisition: H.A., G.L.M., J.H.M.; Statistical analysis: H.L., M.H., S.J.K., J.H.M.; Interpretation of data: H.L., M.H., S.J.K., J.H.M., G.L.M.; Drafting of initial manuscript: H.L.; Critical revision of the manuscript for important intellectual content: All authors; Administrative, technical, or material support: H.A.; All authors approved the final version of the manuscript.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

COMPETING INTERESTS

H.L.: Grants (National Health and Medical Research Council of Australia); G.L.M.: Consultancy (Grunenthal, Australian Institute of Sport, Grunenthal, Kaiser Permanente California, Pfizer, Return to Work SA); Grants (National Health and Medical Research Council of Australia, ID 1061279; 0630431; 1008017; 1047317); Payment for lectures including service on speakers' bureaus (lectures for various companies on pain); Royalties (Explain Pain, Painful Yarns, Graded Motor Imagery Handbook, The Explain Pain Handbook: Protectometer, Noigroup Publications); S.J.K.: Consultancy (AO Spine: providing methodological advice for research projects); Employment (National Health and Medical Research Council of Australia); J.H.M.: Grants (National Health and Medical Research Council of Australia, ID 1008003 and 1043621). The remaining authors have no conflicts of interest to declare.

ACKNOWLEDGEMENTS

PeopleBrowsr (San Francisco, California, USA).

SUPPLEMENTARY MATERIAL

Supplementary material is available online at http://jamia.oxfordjournals.org/.

REFERENCES

1. Vos T, Barber R, Bell B, *et al*. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet.* 2015;386:743–800.
2. Martin BI, Deyo RA, Mirza SK, *et al*. Expenditures and health status among adults with back and neck problems. *JAMA.* 2008;299:656–664.
3. Maetzel A, Li L. The economic burden of low back pain: a review of studies published between 1996 and 2001. *Best Pract Res Clin Rheumatol.* 2002;16:23–30.
4. Hoy D, Brooks P, Blyth F, *et al*. The Epidemiology of low back pain. *Best Pract Res Clin Rheumatol.* 2010;24:769–781.
5. van Oostrom SH, Driessen MT, de Vet HCW, *et al*. Workplace interventions for preventing work disability. *Cochrane database Syst Rev.* 2009;15: CD006955.
6. Henschke N, Maher CG, Refshauge KM, *et al*. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. *BMJ.* 2008;337:a171.
7. Steffens D, Ferreira ML, Latimer J, *et al*. What triggers an episode of acute low back pain? A case-crossover study. *Arthritis Care Res.* 2015;67:403–410.
8. Smith C. *By The Numbers: 150+ Amazing Twitter Statistics*. 2015. Available at: URL: http://expandedramblings.com/index.php/march-2013-by-the-numbers-a-few-amazing-twitter-stats/. Accessed July 29, 2015.
9. Milinovich GJ, Williams GM, Clements AC, *et al*. Internet-based surveillance systems for monitoring emerging infectious diseases. *Lancet Infect Dis.* 2014;14:160–8.
10. Ahlwardt K, Heaivilin N, Gibbs J, *et al*. Tweeting about pain: comparing self-reported toothache experiences with those of backaches, earaches and headaches. *J Am Dent Assoc.* 2014;145:737–743.
11. Pagoto S, Schneider KL, Evans M, *et al*. Tweeting it off: characteristics of adults who tweet about a weight loss attempt. *JAMIA.* 2014;21:1032–1037.
12. Williams SA, Terras M, Warwick C. How twitter is studied in the medical professions: a classification of twitter papers indexed in PubMed. *Med 2.0.* 2013;2:e2.
13. Sullivan SJ, Schneiders AG, Cheang C-W, *et al*. 'What's happening?' A content analysis of concussion-related traffic on Twitter. *Br J Sports Med.* 2012;46:258–263.
14. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. *PLoS One.* 2013;8:e83672.
15. Tremayne M. Anatomy of protest in the digital era: a network analysis of twitter and occupy wall street. *Soc Mov Stud.* 2014;13:110–126.
16. de Vet HCW, Heymans MW, Dunn KM, *et al*. Episodes of low back pain: a proposal for uniform definitions to be used in research. *Spine (Phila Pa 1976)* 2002;27:2409–2416.
17. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol.* 1991;133:144–153.
18. Mittleman M, Maclure M, Tofler G, *et al*. Triggering of acute myocardial infarction by heavy physical exertion. *N Engl J Med.* 1993;329:1677–1683.
19. Bhaskaran K, Armstrong B, Hajat S, *et al*. Heat and risk of myocardial infarction: hourly level case-crossover analysis of MINAP database. *BMJ.* 2012;345:e8050.
20. Sorock GS. A case-crossover study of transient risk factors for occupational acute hand injury. *Occup Environ Med.* 2004;61:305–311.
21. Engel GL. The need for a new medical model: a challenge for biomedicine. *Science.* 1977;196:129–136.
22. Moseley GL and Butler DS. 15 years of explaining pain - the past, present and future. *J Pain* 2015;16:807–813.
23. Tighe JP, Goldsmith CR, Gravenstein M, *et al*. The painful Tweet: text, sentiment, and community structure analyses of tweets pertaining to pain. *J Med Internet Res.* 2015;17:e84.
24. Heaivilin N, Gerbert B, Page JE, *et al*. Public health surveillance of dental pain via Twitter. *J Dent Res.* 2011;90:1047–1051.
25. Maclure M, Mittleman M. Should we use a case-crossover design? *Annu Rev Public Health. 2000*;21:193–221.
26. Young SD, Rivers C, Lewis B. Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes. *Prev Med.* 2014;63:112–115.
27. Wehner MR, Chren MM, Shive ML, *et al*. Twitter: an opportunity for public health campaigns. *Lancet.* 2014;384:131–132.
28. Milinovich GJ, Magalhães RJS, Hu W. Role of big data in the early detection of Ebola and other emerging infectious diseases. *Lancet Glob Heal.* 2015;3:e20–e21.
29. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection - Harnessing the Web for public health surveillance. *N Engl J Med.* 2009;360:2153–2155.
30. Berenbaum F. The social (media) side to rheumatology. *Nat Rev Rheumatol.* 2014;10:314–318.
31. Laranjo L, Arguel A, Neves AL, *et al*. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *JAMIA.* 2015;22(1):243–256.
32. Hendrick PA, Ahmed OH, Bankier SS, *et al*. Acute low back pain information online: an evaluation of quality, content accuracy and readability of related websites. *Man Ther.* 2012;17:318–324.
33. Koes BW, van Tulder M, Lin CWC, *et al*. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J.* 2010;19:2075–2094.
34. Traeger AC, Hübscher M, Henschke N, *et al*. Effect of primary care–based education on reassurance in patients with acute low back pain systematic review and meta-analysis. *JAMA Intern Med.* 2015;175:733–743.
35. Carpenter KM, Stoner SA, Mundt JM, *et al*. An online self-help cbt intervention for chronic lower back pain. *Clin J Pain.* 2012;28:14–22.
36. Dear BF, Titov N, Perry KN, *et al*. The Pain Course: a randomised controlled trial of a clinician-guided Internet-delivered cognitive behaviour therapy program for managing chronic pain and emotional well-being. *Pain.* 2013;154:942–950.
37. Cavallo DN, Tate DF, Ries AV, *et al*. A social media-based physical activity intervention: a randomized controlled trial. *Am J Prev Med.* 2012;43: 527–532.
38. Bull SS, Levine DK, Black SR, *et al*. Social media-delivered sexual health intervention: a cluster randomized controlled trial. *Am J Prev Med.* 2012;43:467–474.
39. Yoon S, Elhadad N, Bakken S. A practical approach for content mining of tweets. *Am J Prev Med.* 2013;45:122–129.
40. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One.* 2011;6:e19467.

AUTHOR AFFILIATIONS

[1]Neuroscience Research Australia (NeuRA), Sydney, NSW, Australia

[2]Prince of Wales Clinical School, University of New South Wales, Sydney, NSW, Australia

[3]EMGO+ Institute, VU University Medical Centre, Amsterdam, Netherlands

[4]The George Institute for Global Health, University of Sydney, Sydney, NSW, Australia

[5]Sansom Institute for Health Research, University of South Australia, Adelaide, SA, Australia