Twitter metrics complement traditional conference evaluations to evaluate knowledge translation at a National Emergency Medicine Conference

Stella Yiu, MD[*]; Sebastian Dewhirst, MD[*]; Ali Jalali, MD[†]; A. Curtis Lee, PhD[‡]; Jason R Frank, MD, MA(Ed)[*§]

CLINICIAN'S CAPSULE

What is known about the topic?
Traditional conference evaluations might be limited in evaluating conference knowledge translation.
What did this study ask?
How do Twitter metrics complement traditional conference evaluation.
What did this study find?
There is no correlation between Twitter metrics and traditional evaluation, but Twitter metrics could be a complementary tool to evaluate conference knowledge dissemination.
Why does this study matter to clinicians?
Clinicians might wish to adopt Twitter to receive disseminated conference content.

ABSTRACT

Objectives: Conferences are designed for knowledge translation, but traditional conference evaluations are inadequate. We lack studies that explore alternative metrics to traditional evaluation metrics. We sought to determine how traditional evaluation metrics and Twitter metrics performed using data from a conference of the Canadian Association of Emergency Physicians (CAEP).
Methods: This study used a retrospective design to compare social media posts and tradition evaluations related to an annual specialty conference. A post ("tweet") on the social media platform Twitter was included if it associated with a session. We differentiated original and discussion tweets from retweets. We weighted the numbers of tweets and retweets to comprise a novel Twitter Discussion Index. We extracted the speaker score from the conference evaluation. We performed descriptive statistics and correlation analyses.
Results: Of a total of 3,804 tweets, 2,218 (58.3%) were session-specific. Forty-eight percent (48%) of all sessions received tweets (mean = 11.7 tweets; 95% CI of 0 to 57.5; range, 0–401), with a median Twitter Discussion Index score of 8 (interquartile range, 0 to 27). In the 111 standard presentations, 85 had traditional evaluation metrics and 71 received tweets ($p > 0.05$), while 57 received both. Twenty (20 of 71; 28%) moderated posters and 44% (40 of 92) posters or oral abstracts received tweets without traditional evaluation metrics. We found no significant correlation between Twitter Discussion Index and traditional evaluation metrics (R = 0.087).
Conclusions: We found no correlation between traditional evaluation metrics and Twitter metrics. However, in many sessions with and without traditional evaluation metrics, audience created real-time tweets to disseminate knowledge. Future conference organizers could use Twitter metrics as a complement to traditional evaluation metrics to evaluate knowledge translation and dissemination.

RÉSUMÉ

Objectif: Les congrès sont conçus pour favoriser l'application des connaissances, mais les méthodes classiques d'évaluation ne conviennent pas vraiment à l'objet visé, et il existe peu d'études sur la recherche d'autres instruments de mesure. La présente étude avait donc comme objectif de comparer la performance des méthodes classiques d'évaluation avec celle d'indicateurs Twitter, à l'aide de données recueillies au cours d'un congrès de l'Association canadienne des médecins d'urgence.
Méthode: Il s'agit d'une étude rétrospective dans laquelle ont été comparées des publications dans les réseaux sociaux et des méthodes classiques d'évaluation en lien avec un congrès annuel de médecine de spécialité. Les publications (« gazouillis ») faites sur la plateforme de réseau social Twitter étaient retenues si elles se rapportaient à une séance. L'équipe a fait la distinction entre les gazouillis originaux et les échanges, et

From the *Department of Emergency Medicine, University of Ottawa Faculty of Medicine, Ottawa, ON; †Department of Innovation in Medical Education and Faculty of Medicine, University of Ottawa, Ottawa, ON; ‡School of Medicine and Public Health, University of Newcastle, Callaghan, NSW, Australia; and the §Royal College of Physicians and Surgeons of Canada, Ottawa, ON.

Correspondence to: Dr. Stella Yiu, The Ottawa Hospital, 1053 Carling Avenue, Room M206, Ottawa, Ontario K1Y 4E9 Canada; Email: syiu@toh.ca

 CJEM 2020;22(3):379–385 DOI 10.1017/cem.2020.15

les gazouillis partagés, puis a pondéré le nombre de gazouillis originaux et de gazouillis partagés afin de constituer un nouvel indice d'échanges sur Twitter. Le résultat de l'évaluation des conférenciers a été tiré de l'évaluation du congrès. L'équipe a finalement procédé au calcul de statistiques descriptives et à des analyses de corrélation.

Résultats: Sur un total de 3804 gazouillis, 2218 (58,3%) se rapportaient à des séances en particulier. Quarante-huit pour cent (48%) des séances ont fait l'objet de gazouillis (moyenne : 11,7; IC à 95% : 0- 57,5; plage : 0-401) et le résultat médian de l'indice d'échanges sur Twitter s'élevait à 8 (IIQ : 0-27). Sur les 111 présentations de type usuel, 85 ont été soumises à des évaluations classiques; 71 ont suscité des gazouillis ($p > 0,05$) et 57 ont fait l'objet et d'évaluations classiques et de gazouillis. Vingt séances d'affichage animées (20 sur 71; 28%) et 44% (40 sur 92) des présentations par affiches ou des présentations orales de résumés ont fait l'objet de gazouillis seuls. Il ne s'est dégagé aucune corrélation significative de la comparaison entre l'indice d'échanges sur Twitter et les instruments classiques d'évaluation ($\rho = 0,087$).

Conclusion: Aucune corrélation n'a été établie entre les instruments classiques d'évaluation et les indicateurs Twitter. Toutefois, l'équipe a constaté que, durant bon nombre de séances soumises ou non à des mesures classiques d'évaluation, l'assistance envoyait des gazouillis en temps réel pour diffuser des connaissances. Les organisateurs de congrès futurs pourraient donc utiliser des indicateurs Twitter comme instruments complémentaires des méthodes classiques d'évaluation au regard de l'application et de la diffusion des connaissances.

Keywords: Education, knowledge translation, social media

INTRODUCTION

Medical conferences are the largest venues for continuing medical education. In addition to networking, they aim to disseminate and advance research, and enhance practice through education.[1,2] In 2018, 1,303 health and medical conferences were held in the United States, with another 118 in Canada.[3] However, they change physician knowledge or patient outcomes minimally.[3–5] Conferences are evaluated by traditional evaluation metrics, but they are inadequate to assess knowledge translation for several reasons. Traditional evaluation metrics mostly focus on social experience, overall satisfaction, and processes instead of learning outcomes or commitment to change.[6] They aggregate individual responses and diminish potentially diverse input.[7,8] With 30% to 40% response rates, there are also questions about their quality and utility.[9] Overall, they provide little insight into how much conference research knowledge is disseminated into practice.[10]

Given the emerging prevalence of social media in this digital age, perhaps there are alternative metrics to complement traditional evaluation metrics and provide greater insight into the how conference knowledge disseminates. Twitter, an online micro-blogging service launched in 2006, enables people to communicate in real-time by means of limited character entries (initially 140, now expanded to 280), called "tweets." These tweets can be repeated by anyone reading them, known as "retweets." Twitter aligns with social-constructivist pedagogy of seeking, sharing, and collaborating.[8,10] It has transformed communication in journalism, public health research, and emergency response. Academics use Twitter to obtain and share real-time information, expand professional networks, and contribute to wider conversations.[11] Educators use Twitter as virtual communities of practice, sharing innovations and feedback informally despite brief interactions.[12–14] Furthermore, Twitter metrics have been compared with other traditional metrics in knowledge translation for published articles. They can predict highly cited articles within the first 3 days of publication. They correlate significantly with traditional bibliometric indicators (readership, citations) in some journals.[7,9,15] They also correlate with citations at Google Scholar™ (a free online search engine for scholarly literature including articles, theses, books, abstracts). Twitter metrics predict top-cited articles with 93% specificity and 75% sensitivity.[9,16]

In medical conferences, Twitter has been shown to disseminate research knowledge.[17–26] Attendees tweet presented results.[22] Those absent from the conference read these tweets, and some choose to retweet to their followers on Twitter, creating a second tier of knowledge dissemination. Some might add their own comment in the retweets. This retweeting can continue for many tiers to diffuse knowledge.[21] With increasing Twitter activity in conference in recent years, some authors encouraged conference organizers to encourage Twitter use for maximum effectiveness.[27] Aside from content broadcasting, Twitter might be useful for evaluation and community discussion. With this in mind, authors

https://doi.org/10.1017/cem.2020.15 Published online by Cambridge University Press

Table 1. Properties of traditional evaluation metrics v. Twitter

Traditional evaluation metrics	Twitter
• Static • Focus on participant satisfaction and processes (8) • Aggregate individual responses (7) • Questionable utility and quality (43) • Little insight on research dissemination (10)	• Contemporaneous • Disseminate breaking results (22) • Could disseminate tiers of knowledge (21) • Available to public and those absent • Could involve active debate (31,32)

Table 2. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
• Contain #CAEP14 AND • Attributable to a specific session: • Mentions speaker name, or, • Matches session content and timing	• Not about session content (e.g.: logistics, social) • Advertises a session beforehand • Not attributable to a specific session

Table 3. Tweet classification system

Class	Description
Disseminating (1)	Tweet including direct quotes from speakers' talks
1R	Retweet of Class 1 tweets
Engagement (2)	Tweet that discuss points stemming from speakers' talks
2R	Retweet of Class 2 tweets

have suggested that Twitter could be a novel real-time speaker impact evaluation tool[23] (see Table 1). Perhaps Twitter metrics can complement what is frequently missed by traditional evaluation metrics. To our knowledge, there has not been a study that compares Twitter metrics to traditional evaluation metrics as a speaker impact evaluation tool. We, therefore, asked:

1. Do Twitter metrics correlate with traditional evaluation metrics?
2. How do these two metrics measure speaker impact differently?

METHODS

This study used a retrospective design. The hashtag (a metadata tag in Twitter) #CAEP14 was prospectively registered with Symplur, an online Twitter management tool, so that all tweets bearing the hashtag #CAEP14 (Canadian Association of Emergency Physicians conference 2014) were archived. Attendees were encouraged to tweet using this hashtag. All tweets that date from the start date of the conference to 30 days afterward were collected.

Two authors (S.Y. and S.D.) independently assessed each tweet for inclusion. Table 2 described the inclusion and exclusion criteria of each tweet.

We developed a classification system (see Table 3) to differentiate original tweets from retweets, and tweets that generated further discussion. All researchers discussed and agreed upon a coding scheme. Two authors assessed and coded the first 200 tweets together to ensure a uniform approach to coding, and independently coded the remaining tweets. All tweets were revisited and coding discrepancies were resolved by consensus among the study team. Manual coding was chosen based on previous literature, because content analysis is hampered by brevity and unconventional forms of written expression.[28]

With the consent of CAEP, all available conference speaker evaluations from the CAEP 2014 conference were collected confidentially. One author (S.Y.) reviewed conference speaker evaluations, and abstracted the value corresponding to the rating item "The speaker was an effective communicator." We chose this because it was the only one item specific to a speaker rather than the whole session (in which multiple speakers would speak). We believed this was the most appropriate unit as the tweets were aimed at specific speaker and not the whole session.

We calculated descriptive statistics using proportions, means, or medians with standard deviations of interquartile ranges (IQRs) as appropriate. Means and proportions were compared using t-tests with reported *p*-values. Linear correlation analyses, using Pearson's R, were calculated to compare conference evaluation scores and Twitter metrics.

To quantify the impact of tweets, we proposed a novel theory-based Twitter Discussion Index. This index included all original tweets (Class 1 – Disseminating tweets and Class 2 – Engagement tweets) and retweets

https://doi.org/10.1017/cem.2020.15 Published online by Cambridge University Press

(Class 1R and Class 2R). We have defined the Twitter Discussion Index to be as follows:

$$\text{Twitter Discussion Index} = 2\,x\,\textit{Class 1 (Disseminating)} + 3\,x\,\textit{Class 2 (Engagement)} + \textit{Class 1R} + \textit{Class 2R}$$

We weighted the components of this equation based on our lens of Twitter as aligned with social-constructivist pedagogy.[8,29] Through Class 1 tweets, users construct and disseminate their own understanding. Class 2 tweets create important networks of interactivity and virtual communities of practice, and are vital to further layers of discussion. Nonoriginal retweets (Classes 1R and 2R) are endorsements and amplify the message spread.

RESULTS

Description of Twitter data

In total, 3,804 tweets contained the hashtag #CAEP14, and fell within the prescribed date range. Of these, 2,419 (63.59%) were included. The others were excluded as they had no relation to content. (Examples of excluded tweets: personal communications, logistics about room locations, reminders for upcoming sessions.) Forty-eight percent (48%) of sessions received at least one tweet (mean = 11.7 tweets; 95% CI of 0 to 57.5; range, 0–401). Included tweets were classified as follows: 634 (26.21%) were Class 1; 1,276 (52.75%) were Class 1R; 190 (7.85%) were Class 2; and 319 (13.19%) were Class 2R (Table 4).

Plenary sessions and sessions that encouraged audience input received a much higher number of tweets (mean = 219.8 tweets, overall conference mean per session = 11.7 tweets).

Comparison between Twitter metrics and traditional evaluation metrics

In this conference, there were 274 sessions total including standard presentations (111), posters (71) and abstracts (92). Only the 111 standard presentations have traditional evaluation metrics for attendees to fill. Within these 111 standard presentations, 85 (76.58%) received traditional evaluation metrics, and 71 (63.96%) received tweets.

Fifty-seven (57 of 111; 51.35%) standard presentations received both traditional evaluation metrics and tweets. Of the 26 standard presentations with no traditional evaluation metrics, 14 received tweets. Of the 40 standard presentations that received no tweets, 14 received traditional evaluation metrics.

Table 4. Classification of tweets from conference

Description	Number of tweets
Total	3,804
Included tweets (n)	2,419 (63.6% of total)
Disseminating (1)	634 (26.2% of n)
1R	1,276 (52.8% of n)
Engagement (2)	190 (7.9% of n)
2R	319 (13.2% of n)

In all sessions (including standard presentations, posters, and abstracts), 48% (131 of 274) received tweets (mean = 11.7 per session). For the posters and abstracts when no traditional evaluation metrics were available, 28% (20 of 71) moderated posters and 44% (40 of 92) posters or oral abstracts received tweets (see Figure 1).

In sessions with both tweets and evaluation scores (n = 57) there was no significant correlation between the number of tweets (any Class), Twitter Discussion Index, and the evaluation scores. The median traditional evaluation metrics score was 3.61 of 5, IQR of 3.4 to 3.7 (see Figure 2).

DISCUSSION

Interpretation of findings

Medical conferences are venues designed to bridge the gap between research and practice,[1] but static traditional evaluation metrics are not designed to assess knowledge. Given the emerging prevalence of social media in this digital age, we sought to study whether Twitter metrics can complement traditional evaluation metrics and provide greater insight how medical conferences translate knowledge.

We found no correlation between traditional evaluation metrics and Twitter. This is not due to discordant results between the two (such as highly tweeted session receiving low scoring traditional evaluation metrics). Rather, we were unable to correlate the measures due to a lack of traditional evaluation metrics and a narrow range of scores in those available (median score of 3.61 out of 5 with IQR of 3.4 to 3.7).

https://doi.org/10.1017/cem.2020.15 Published online by Cambridge University Press

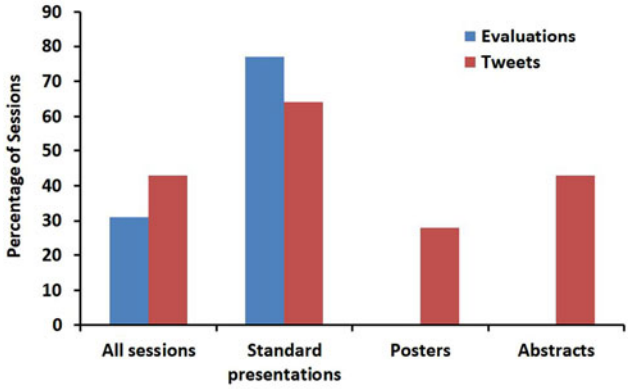

Figure 1. Sessions receiving conference evaluations and tweets separated by type.

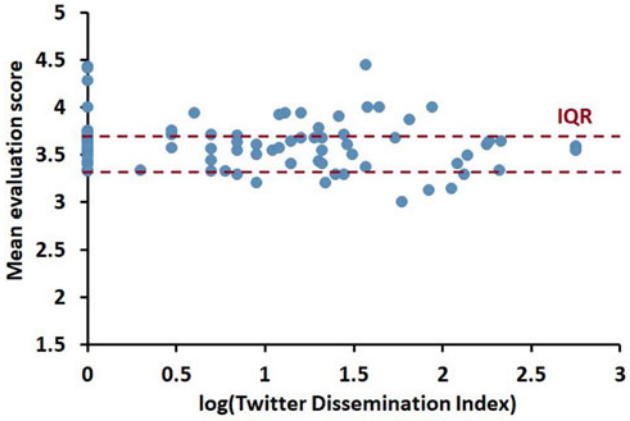

Figure 2. Mean evaluation score v. log of twitter discussion index per session. Note the much higher variability of the Twitter Discussion Index compared with that of the mean evaluation score.

While tweets were generated in a similar rate in sessions with traditional evaluation metrics, a substantial percentage of sessions without traditional evaluation metrics (by design before conference) generated tweets. Tweets displayed a wide variation of total number among sessions. Most tweets focused on knowledge content rather than logistics or processes, similar to previous studies.[19,25] The majority of tweets (78.95%) disseminated content (Class 1 and Class 1R), and the rest (21.04%) sparked further debate as "discussion" tweets (7.9% were "discussion" tweets and 13.2% were retweets of those).

Comparison to previous studies

We found that our Twitter metrics are similar to previous conferences.[19,22,24,25] These metrics might provide information about knowledge translation and dissemination that traditional evaluation metrics lack. These qualities are derived from the nature of the social media platform: real-time, accessible, searchable, and focused on knowledge acquisition and transfer. We will discuss these qualities below.

Twitter metrics are real-time. Didactic conference sessions suggest a single focus of attention, and restrict individuals to the role of either speaker or listener.[20] Feedback, collaboration, and interaction are often missing.[30] By contrast, Twitter engages. Discussion typically involves active debate, despite character limit.[31,32] Desai advocated that real-time feedback might be less subjected to bias than traditional evaluation metrics.[33] Twitter feedback may improve presentation quality, particularly if speakers were informed of the need for clear key messages. Even though higher Twitter metrics might reflect more knowledge translation and dissemination, there are caveats. Elements such as audience size and content that is extreme, shocking, humorous, or controversial would change tweets.[14] Tweets can also be an echo chamber of comments representing shared opinion rather than knowledge translation.[14] Others cautioned the risk of sensationalism inflating the tweet numbers, or presenters "sterilizing" down their content for risk of being misquoted or misinterpretated.[23] As there is little anonymity on Twitter, it might deter those writing negative feedback. Retractions and errata from authors are rare and might be unnoticed on Twitter.[14] Content analysis of tweets might mitigate these risks in future studies, but no formal process of fact checking is in place as of yet.

Traditional evaluation metrics are reviewed by organizers and speakers only, while Twitter is accessible to a community. Tweets are archived and searchable, making it an attractive feature for future reference. Contrary to other social media sites (e.g., Facebook, private institution site), Twitter reaches further than a specific group. Rather, posted messages are public by default. They can be searched and tracked by hashtags. Each Twitter user can create public posts to initiate discussions and to participate in debates.[34,35] Bakshy et al. discovered that tweets tend to propagate in a power law distribution, with a small number of tweets being retweeted thousands of times.[36] These retweeters are the key to wide knowledge dissemination.[17] Because the retweeters do not need to be present in the conference, the impact is not dependent on the size of conference attendees. While traditional evaluation metrics focused on satisfaction

https://doi.org/10.1017/cem.2020.15 Published online by Cambridge University Press

and reactions instead of learning, tweets are largely about learning points, aligning with just-in-time learning and knowledge transfer. Also, despite the small character number of tweets, they were often robust and clinically relevant.[22,37,38] It is also possible that tweets can impact long-term retention by mechanisms such as retrieval practice (from the audience tweeting), feedback (from correcting others on Twitter), and spaced repetitions (from tweets and retweets).[39]

STRENGTHS AND LIMITATIONS

Our study is the first one that compares traditional evaluation metrics with Twitter metrics, and we had a novel way to differentiate between different levels of contemporaneous tweets.

We captured only tweets that bore the conference hashtag (#CAEP14). It is possible that there were related tweets without it or that some bore the wrong hashtag. As a result, these tweets might have been missed. In addition, the Twitter Discussion Index does not discriminate between positive tweets and negative tweets. In previous studies, Twitter metrics can also be a marker of strong disagreement, research error, or frank misconduct.[40] Even though we encountered no tweets with strong sentiment of disagreement in our study, this is a potential limitation of the Twitter Discussion Index.

Also, this study only has data from one conference during a single year, limiting its conclusion.

RESEARCH IMPLICATIONS

Given that Twitter could be an informative metric, we propose that our Twitter Discussion Index be treated as a measure of "disseminative impact", similar to published articles generating "buzz" with altmetrics leading to high citations.[41] It might be used to complement traditional evaluation metrics in future conference evaluation as a key performance indicator of engagement and impact.[42]

CONCLUSION

Traditional evaluation metrics are inadequate to evaluate medical conference presentations for knowledge translation. Tweets by conference attendees could amplify knowledge translation and dissemination. Tweets are real-time, accessible, searchable, and describe knowledge transfer. We found Twitter metrics a more nuanced evaluation tool that complements traditional evaluation metrics. We propose a novel index for the use of this tool. We recommend conference organizers to adopt Twitter metrics and Twitter Discussion Index as a measure of knowledge translation and dissemination.

Acknowledgements: The authors thank the Board of CAEP for the use of their conference evaluation results. Author contributions: S.Y. conceived the idea for the research. S.Y., S.D., and J.R.F. each contributed to the design of the study. The study was executed by S.D. and S.Y. AJ led the analysis along with SY, SD, and ACL. S.Y. wrote the first draft of the manuscript, and subsequent versions were reviewed, commented on, and revised critically by all authors. All authors approved the final manuscript for submission.

Competing interest: None declared.

Financial support: This research received funding from the Department of Emergency Medicine at The Ottawa Hospital.

REFERENCES

1. Davis D. The case for knowledge translation: shortening the journey from evidence to effect. *BMJ* 2003;327(7405):33–5.
2. Ioannidis JP. Are medical conferences useful? And for whom? *JAMA* 2012;307(12):1257–8.
3. CLocate. Health and medical conferences in Canada 2018. www.clocate.com. 2019. [cited June 14, 2019]. pp. 1–33. Available at: https://www.clocate.com/ (accessed February 4, 2020).
4. Mansouri M, Lockyer J. A meta-analysis of continuing medical education effectiveness. *J Contin Educ Health Prof* 2007;27(1):6–15.
5. Forsetlund L, Bjørndal A, Rashidian A, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2009;20(2):CD003030.
6. Domino FJ, Chopra S, Seligman M, Sullivan K, Quirk ME. The impact on medical practice of commitments to change following CME lectures: a randomized controlled trial. *Med Teach* 2011;33(9):e495-e500.
7. Haustein S, Peters I, Sugimoto CR, Thelwall M, Larivière V. Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *J Assoc Inf Sci Technol* 2013;65(4):656–69.
8. Chapman DD, Aalsburg Wiessner C, Storberg-Walker J, Hatcher T. New learning: the next generation of evaluation? Academia. 2006. Available at: https://www.academia.edu/5119525/New_Learning_The_Next_Generation_of_Evaluation (accessed February 18, 2020).
9. Neves J, Lavis JN, Ranson MK. A scoping review about conference objectives and evaluative practices: how do we get more out of them? *Health Res Policy Syst* 2012;10(26):26.

https://doi.org/10.1017/cem.2020.15 Published online by Cambridge University Press

10. Maloney S, Tunnecliff J, Morgan P, et al. Translating evidence into practice via social media: a mixed-methods study. *J Med Internet Res* 2015;17(10):e242.
11. Mohammadi E, Thelwall M, Kwasny M, Holmes KL. Academic information on Twitter: a user survey. *PLoS One* 2018;13(5):e0197265.
12. Rosell-Aguilar F. Twitter: a professional development and community of practice tool for teachers. *J Interact Media Educ* 2018;(1):6.
13. Coleman JM, Rice ML, Wright VH. Educator communities of practice on Twitter. *J Interact Online Learn* 2018;16(1):1–179.
14. Choo EK, Ranney ML, Chan TM, et al. Twitter as a tool for communication and knowledge exchange in academic medicine: a guide for skeptics and novices. *Med Teach* 2015;37 (5):411–6.
15. Amath A, Ambacher K, Leddy JJ, Wood TJ, Ramnanan C. Comparing alternative and traditional dissemination metrics in medical education. *Med Educ* 2017:51(9):935–41.
16. Eysenbach G. Can Tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *J Med Internet Res* 2011;13(4): e123.
17. Mishori R, Levy B, Donvan B. Twitter use at a family medicine conference: analyzing #STFM13. *Fam Med* 2014;46 (8):608–14.
18. Anderson G, Gleeson S, Rissel C, Wen L, Bedford K. Twitter tweets and twaddle: twittering at AHPA's national health promotion conference. *Health Promot J Austr* 2014;25(3):143–6.
19. Nomura JT, Genes N, Bollinger HR, Bollinger M, Reed III. JF Twitter use during emergency medicine conferences. *Am J Emerg Med* 2012;30(5):819–20.
20. Ross C, Terras M, Warwick C, Welsh A. Enabled backchannel: conference Twitter use by digital humanists. *J Doc* 2011;67(2):214–37.
21. McKendrick DR, Cumming GP, Lee AJ. Increased use of Twitter at a medical conference: a report and a review of the educational opportunities. *J Med Internet Res* 2012;14 (6):e176.
22. Chaudhry A, Glode LM, Gillman M, Miller RS. Trends in Twitter use by physicians at the American Society of Clinical Oncology Annual Meeting, 2010 and 2011. *J Oncol Pract* 2012;8(3):173–8.
23. Roland D, May N, Body R, Carley S. Are you a SCEPTIC? SoCial mEdia Precision & uTility In Conferences. *Emerg Med J*. 2015;32(5):412–3.
24. Jalali A, Wood T. Analyzing online impact of Canadian Conference of Medical Education through Tweets. *EIMJ* 2013;5(3). doi: 10.5959/eimj.v5i3.162.
25. Neill A, Cronin JJ, Brannigan D, O'Sullivan R, Cadogan M. The impact of social media on a major international emergency medicine conference. *Emerg Med J* 2014;31(5):401–4.
26. Scott KR, Hsu CH, Johnson NJ, Mamtani M, Conlon LW, DeRoos FJ. Integration of social media in emergency medicine residency curriculum. *Ann Emerg Med* 2014;64 (4):396–404.
27. Gricks B, Woo HH, Lai CSL. Tweeting the meeting: an analysis of Twitter activity at the Royal Australasian College of Surgeons Annual Scientific Congress from 2015 to 2018. *ANZ J Surg* 2019;89(11):1485–9.
28. Kim AE, Hansen HM, Murphy J, Richards AK, Duke J, Allen JA. Methodological considerations in analyzing Twitter data. *J Natl Cancer Inst Monogr* 2013;2013(47):140–6.
29. Maloney S, Moss A, Ilic D. Social media in health professional education: a student perspective on user levels and prospective applications. *Adv Health Sci Educ Theory Pract* 2014;19(5):687–97.
30. Reinhardt W, Ebner M, Beham G, Costa C. How people are using Twitter during conferences. In: Hornung-Prahauser V, Luckmann M, eds. *Creativity and Innovation Competencies on the Web*. Salzburg: Salzburg, Austria; 2009:145–6.
31. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: Twitter and antibiotics. *Am J Infect Control* 2010;38(3):182–8.
32. Jalali A, Sherbino J, Frank J, Sutherland S. Social media and medical education: exploring the potential of Twitter as a learning tool. *Int Rev Psychiatry* 2015;27(2):140–6.
33. Desai B. A novel use of Twitter to provide feedback and evaluations. *Clin Teach* 2014;11(2):141–5.
34. Bruns A, Stieglitz S. Quantitative approaches to comparing communication patterns on Twitter. *J Technol Hum Serv* 2012;30(3-4):160–85.
35. Boyd D, Golder S, Lotan G. Tweet, tweet, retweet: conversational aspects of retweeting on twitter. In: 43rd Hawaii International Conference on System Sciences, Honolulu, HI, 2010.
36. Bakshy E, Hofman JM, Mason WA, Watts DJ. Everyone's an influencer: quantifying influence on twitter. 2011. Available at: http://snap.stanford.edu/class/cs224w-readings/bakshy11 influencers.pdf (accessed February 4, 2020).
37. Brown A, Riddell J, Jauregui J, Yang J, Nauman R, Robins L. What's all the chatter? A mixed-methods analysis of emergency physician's tweets. *West J Emerg Med* 2017;18 (Suppl):S32–3.
38. Riddell J, Brown A, Robins L, Nauman R, Yang J, Jauregui J. What's all the chatter? A mixed-methods analysis of emergency physician's tweets. *West J Emerg Med* 2020;21(1).
39. Butler AC, Raley ND. The Future of Medical Education: Assessing the Impact of Interventions on Long-Term Retention and Clinical Care. *J Grad Med Educ* 2015;7 (3):483–5.
40. Carpenter CR, Cone DC. Using publication metrics to highlight academic productivity and research impact. *Acad Emerg Med* 2014;21:1160–72.
41. Trueger NS, Thoma B, Hsu CH, Sullivan D, Peters L, Lin M. The Altmetric Score: a new measure for article-level dissemination and impact. *Ann Emerg Med* 2015;66(5):549–53.
42. Neiger BL, Thackeray R, Van Wagenen SA, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. *Health Promot Pract* 2012;13(2):159–64.
43. Neves J, Lavis JN, Ranson MK. A scoping review about conference objectives and evaluative practices: how do we get more out of them? *Health Res Policy Syst* 2012;10(1):274.

https://doi.org/10.1017/cem.2020.15 Published online by Cambridge University Press