Social media in undergraduate medical education: A systematic review

Jonathan Guckian1,2 | Mrudula Utukuri3 | Aqua Asif4 | Oliver Burton5 | Joshua Adeyoju6 | Adam Oumeziane7 | Timothy Chu2 | Eliot L. Rees8,9

1Dermatology Department, Leeds Teaching Hospitals NHS Trust, Yorkshire, UK
2School of Medical Education, Newcastle University, Newcastle Upon Tyne, UK
3School of Clinical Medicine, University of Cambridge, Cambridge, UK
4Leicester Medical School, University of Leicester, Leicester, UK
5Warwick Medical School, University of Warwick, Coventry, UK
6Faculty of Medicine, University of Southampton, Southampton, UK
7School of Medicine, Anglia Ruskin University, Chelmsford, UK
8School of Medicine, Keele University, Newcastle-under-Lyme, UK
9Research Department of Primary Care and Population Health, University College London, London, UK

Correspondence
Eliot L. Rees, School of Medicine, Keele University, KH77, Keele Hall, Newcastle-under-Lyme ST5 5BG, UK.
Email: e.rees@keele.ac.uk

Funding information
National Institute for Health Research Academic Clinical Fellowship, Grant/Award Number: ACF-2020-18-021

Abstract

Introduction: There are over 3.81 billion worldwide active social media (SoMe) users. SoMe are ubiquitous in medical education, with roles across undergraduate programmes, including professionalism, blended learning, well being and mentoring. Previous systematic reviews took place before recent explosions in SoMe popularity and revealed a paucity of high-quality empirical studies assessing its effectiveness in medical education. This review aimed to synthesise evidence regarding SoMe interventions in undergraduate medical education, to identify features associated with positive and negative outcomes.

Methods: Authors searched 31 key terms through seven databases, in addition to references, citation and hand searching, between 16 June and 16 July 2020. Studies describing SoMe interventions and research on exposure to existing SoMe were included. Title, abstract and full paper screening were undertaken independently by two reviewers. Included papers were assessed for methodological quality using the Medical Education Research Study Quality Instrument (MERSQI) and/or the Standards for Reporting Qualitative Research (SRQR) instrument. Extracted data were synthesised using narrative synthesis.

Results: 112 studies from 26 countries met inclusion criteria. Methodological quality of included studies had not significantly improved since 2013. Engagement and satisfaction with SoMe platforms in medical education are described. Students felt SoMe flattened hierarchies and improved communication with educators. SoMe use was associated with improvement in objective knowledge assessment scores and self-reported clinical and professional performance, however evidence for long term knowledge retention was limited. SoMe use was occasionally linked to adverse impacts upon mental and physical health. Professionalism was heavily investigated and considered important, though generally negative correlations between SoMe use and medical professionalism may exist.
Conclusions: Social media is enjoyable for students who may improve short term knowledge retention and can aid communication between learners and educators. However, higher-quality study is required to identify longer-term impact upon knowledge and skills, provide clarification on professionalism standards and protect against harms.

1 | INTRODUCTION

The explosion of Social Media (SoMe) has infiltrated all aspects of modern society. The scale of the phenomenon can be daunting to conceptualise. There are over 3.81 billion worldwide active users, while individuals have an average of seven accounts each. Facebook, the most popular platform, boasts over a billion more users than the population of China, the world’s most populous country. It is perhaps inevitable that this technology would become ubiquitous within medical education delivery. Our field craves innovation and strives to adapt its future workforce to changing environments. There are few arenas as chaotic as SoMe: since its conception, hundreds of platforms have risen in prominence before falling into obscurity. Various platforms have been utilised within undergraduate medical education curricula, with roles in professionalism, blended learning, student well-being and mentoring representing a fraction of their application.

The argument for the integration of SoMe into educational practice is supported by multiple theoretical standpoints, including the development of rapid, accessible communities of practice, based upon constructivist principles, generated during classes, conferences or conversations. Moreover, connectivism is a ‘new’ educational theory, mirroring constructivism, born out of such technology-enhanced learning.

Research to date examining the impact of SoMe use on medical students’ academic attainment has found no relationship between daily usage, platform preferences and performance on summative assessments. The majority of medical students have SoMe accounts. This has resulted in a powerful tool that can reach virtually every student on platforms already integrated into their daily lives.

As educators, it is no longer a question of whether SoMe has educational applications, or whether undergraduate populations are actively using these platforms for their learning. Rather, we should be asking how best to utilise SoMe, and whether such platforms can facilitate specific outcomes.

Previous systematic reviews of SoMe in medical education have investigated outcomes and efficacy. However, these took place before the SoMe ‘boom’ of recent years, revealing a paucity of high-quality empirical studies assessing the effectiveness of SoMe in medical education. An updated review, synthesising contemporary SoMe scholarship, is required for the current landscape.

In the advent of the COVID-19 pandemic, a once-in-a-century event, the medical education community has transformed. There have been calls for a new paradigm of educational delivery through technology-enhanced learning, with SoMe at the forefront of this revolution. It is imperative that, before committing to placing SoMe as a foundation of any new approach, we critically examine the evidence for its efficacy.

This review aims to synthesise the evidence regarding SoMe interventions in undergraduate medical education, in order to identify features associated with positive and negative outcomes.

2 | METHODS

This is a systematic review reported in accordance with the STructured appRoach to the Reporting In healthcare education of Evidence Synthesis (STORIES) statement.

2.1 | Search strategy

We performed an electronic search of 31 terms and their Boolean combinations (illustrative full terms for one database are provided in Appendix S1) through seven databases: Medline, Cumulative Index of Nursing and Allied Health Literature (CINAHL), British Education Index (BEI), Education Resources Information Center (ERIC), Embase, PsycINFO, Applied Social Sciences Index and Abstracts (ASSIA), and the Australian Education Index. Search terms were derived through reviewing keywords of papers identified through a pilot search, and a list of social media platforms. No limits were imposed. Searches were performed between 16th June and 16th July 2020 by one reviewer (ER).

We scrutinised reference sections and performed forwards citation tracking of all included papers using PubMed’s ‘cited in’ feature, and previous pertinent reviews in order to identify further relevant papers. One reviewer (JG) hand searched the full contents of one journal (MedEdPublish) that was considered relevant but not indexed in the above electronic databases. This journal was chosen as studies identified in the reference section of identified papers were published here, and we were aware that it was not indexed in any of the databases we had searched.

All citations were downloaded and imported into web-based systematic review software (DistillerSR, Evidence Partners,
2.2 | Inclusion and exclusion criteria

For inclusion, articles had to provide primary data on the use of SoMe amongst undergraduate medical students. Studies describing SoMe interventions and research on exposure to existing SoMe were both included. Only English language papers were included.

Defining SoMe has become increasingly challenging as platforms diversify. We accept in principle the definition of SoMe articulated by Cheston et al, later adopted by Sutherland and Jalali, describing SoMe as “Web-based technologies that facilitate multi-user interaction that goes beyond fact sharing”.17,19

Given the evolution of SoMe since the development of this definition in 2013, more recent examinations should be considered. Chan et al24 highlight that platforms “facilitate creation and distribution of content”, which may be “user-generated or user-curated” to “virtual communities of practice”. We would therefore add to Cheston et al’s definition that ‘such interactions may take place in either public or private domains’, stipulate that ‘the primary purpose of the platform must be multi-directional interaction’ and ‘be centred around content creation, curation and community’. Using our adapted definition, we considered true SoMe platforms to place user interaction at their heart. We, therefore, exclude websites or blogs which happen to feature a comments page, as “multi-directional interaction” is a secondary purpose in these instances. We also exclude podcasts on the basis that they are generally unidirectional in nature. YouTube was included as user creation is often driven by community comments, users may curate content and video replies are common.

We defined undergraduate medical education as any educational aspect of the period between the commencement of medical school and graduation.

2.3 | Screening of studies

All titles were independently screened by two reviewers (from MU, AA, OB, JA, AO, TC). Reviewers prioritised sensitivity over specificity at this point, so any titles that could potentially have been relevant to our review objective were included. Disagreements progressed to abstract screening in order to enable a more informed decision to be made.

Abstracts of all papers included from title screening, and all papers identified through reference and hand searching were screened against inclusion criteria by two reviewers (from MU, AA, OB, JA, AO, TC). If it was not clear from the abstract whether the paper met the inclusion criteria, it was included for full-text screening. Disagreements at the outcome or criterion level (ie both reviewers elect to exclude but based on different exclusion criteria) were resolved by a third reviewer (ER or JG).

Papers that met inclusion criteria in abstract screening, or for which insufficient information was presented to enable a judgement, were read in full and reviewed again against the review’s inclusion and exclusion criteria. Full-paper screening was conducted independently by two reviewers (ER & JG). Conflicts were resolved through discussion.

2.4 | Data extraction

A pilot data extraction form was developed. All reviewers independently reviewed two papers and met to ensure we were extracting consistently and to revise the fields in the coding form. Once the form was finalised, all reviewers independently reviewed a further two papers and met to ensure consistency. Following this pilot phase, all papers were reviewed by two reviewers (one from MU, AA, OB, JA, AO, TC and either ER or JG).

The following data were extracted:

- Participants (country, profession, stage of training, number)
- Intervention (focus, aim, brief overview, duration, SoMe platforms used) or Exposure (research question, overview of exposure, SoMe platforms investigated)
- Evaluation methods (impact of intervention using modified Kirkpatrick’s hierarchy,25 study design, data collection methods)
- Results (summary, results for each Kirkpatrick level investigated, key conclusions)

We calculated kappa values for full data extraction and quality assessment.

2.5 | Assessment of methodological quality

All included papers were assessed for their methodological quality using two or more tools.

Depending on whether papers employed quantitative and/or qualitative methods, they were assessed for their methodological quality using the Medical Education Research Study Quality Instrument (MERSQI) and/or the Standards for Reporting Qualitative Research (SRQR) instrument, respectively. The MERSQI is a tool containing 10 items in six domains: study design, sampling, type of data, the validity of evaluation instrument, data analysis, and outcomes.26 The SRQR comprises of 21 reporting standards for high-quality qualitative research.27

Papers describing educational interventions were assessed for risk of reporting bias using the risk of bias measure described by Gordon et al.28 This assesses the risk of bias due to incomplete reporting of educational interventions. Included papers were assessed against five sources of potential bias using a three-point scale. Papers that provided adequate description were considered low risk of bias, those that provided some but insufficient details were rated as unclear risk
of bias, and those that did not provide any details were rated as high risk of bias. The potential sources of bias considered were the theoretical underpinning of the development, the resources required, the setting, the educational methods employed, and the content.

An overall rating of the strength of the conclusions drawn by the authors was made using the BEME collaboration’s five-point scale.  

2.6 | Synthesis of evidence

Extracted data were synthesised using narrative synthesis. This involves synthesising the findings from primary studies textually, without conducting meta analyses. This approach enabled this review to synthesise findings from both qualitative and quantitative studies to provide a comprehensive synthesis of the research literature in this field. The review group met virtually every week throughout the data extraction and synthesis phases to discuss evolving findings. We constructed overview findings for the subgroups of intervention studies and exposure studies. We undertook conceptual mapping to identify themes within which to synthesise and present the findings of primary studies.

3 | RESULTS

Database searches yielded a total of 1,442 papers. A further 321 papers were identified through reference and citation searches, and 44 through hand searching. Deduplication identified 720 duplicates, leaving 1087 papers for screening. Title and abstract screening excluded 327 and 441 papers, respectively. After reviewing the full texts of 319 full papers, 112 met our inclusion criteria and were included in the final review. An illustration of record flow can be found in Figure 1. Weighted overall Kappas for data extraction and quality assessment using MERSQI and SRQR were 0.98, 0.96 and 1.00, respectively.

3.1 | Details of included papers

The studies included participants from 26 different countries across six continents: Asia (37), North America (33), Europe (28), Australia (20), Africa (1), and South America (1). There were an additional six international studies that included participants from multiple countries.

Thirty-nine papers included preclinical students, 32 included clinical students, 21 included all stages of students, and 20 did not describe the stage of included medical students. For the purposes of this review, preclinical students were in Years 1 and 2 whilst clinical students were in Years 3, 4 and 5. Fourteen studies also included postgraduate clinical trainees, 10 included non-training grade doctors (eg consultants/attendings, general practitioners), 7 studies included other health professionals, and 17 included other students. There was a total of 35,428 participants across the included studies. These consisted of a median (range) of 151 (6-4244) participants per study.

---

**FIGURE 1** PRISMA diagram of studies
3.2 | Summary of SoMe interventions/exposures

Twenty-nine different SoMe platforms were studied. The most common was Facebook (65/112, 58%), followed by Twitter (37/112, 33%) and YouTube (33/112, 29%). Forty (36%) studies investigated the use of more than one SoMe platform.

Fifty-eight out of the 112 included studies (52%) involved the development of novel SoMe teaching (intervention studies) whilst the other 54 (48%) focussed on how students currently utilise SoMe platforms in medical education (exposure studies).

3.2.1 | SoMe Interventions

Novel SoMe interventions were created with the aims of: (a) improving knowledge and skill development; (b) supporting curricular activities; and (c) assessing the acceptability of SoMe.

Types of knowledge and skill development included information acquisition and retention, reflection and professionalism. Supporting curricular activities ranged from enhancing student engagement, reducing student anxiety with the taught material, improving communication between faculty and students, and between students themselves outside of the classroom. Many interventions were evaluated to assess student opinions and attitudes towards inclusion of SoMe within their medical curriculum.

Privacy was a focus of 28/58 (48%) SoMe interventions with the use of closed groups, private communication platforms such as WhatsApp and WeChat, and even a bespoke institution-specific SoMe platform. 16/58 (28%) studies involved case-based or problem-based learning.

3.2.2 | Exposures to SoMe

Of the 54 included studies addressing exposures to SoMe-mediated medical education, the most prevalent focus was assessing usage patterns of SoMe platforms by students, seen in 38 studies (70%). Projects looked to identify SoMe platforms and services most commonly used and the duration or purpose of such use.

The second major theme was that of professionalism. In particular, student and faculty behaviour online was analysed for whether these behaviours met regulatory standards. Questions were raised as to whether SoMe makes students and doctors more likely to have lapses in professional behaviour, whether guidelines in professional SoMe use are necessary, and who should be responsible for producing and leading them.

Several studies explored professionalism in the context of acceptable patient interaction.

The findings from included studies have been grouped thematically into six groups: usage, acceptability, educational activity, academic performance, professionalism, and health risks.

3.3 | Usage

SoMe use by undergraduate medical students is widespread. Facebook was identified as the most popularly used platform amongst this group, although YouTube and WhatsApp were the most favoured platforms for educational content. Between a third and a half of students used social media for educational purposes on a daily basis. Frequency of social media usage was not associated with gender or academic year. Stopping the use of Facebook was associated with a sense of loss.

3.4 | Acceptability

SoMe were generally perceived favourably by medical students. Enjoyment and engagement with various SoMe platforms were described. Only two studies differed: one found students’ opinion on SoMe efficacy to be ‘divisive’, whilst another described less than half of their student population finding a SoMe intervention useful in their studies.

Usability was key to effective SoMe use in medical education. Familiarity with SoMe platforms was a mediator of success. Indeed, unfavourable student reviews were reported when unfamiliarity was highlighted. One study with a bespoke SoMe platform reported that students required greater accessibility and a more user-friendly interface.

Technological issues and poor digital literacy may contribute to inconsistencies in the effectiveness of SoMe interventions.

The amount of information presented via a SoMe platform contributed to its effectiveness. Students preferred short content with the “time-consuming” nature of searching SoMe for relevant information contributing to ineffectiveness. It is suggested that the need to regularly check SoMe to keep up to date with content contributes to such inefficiencies; however, push notifications may improve this.

Type of educational information was also important, with many studies reporting a preference for SoMe posts that inform and test, for example, quizzes and images with missing labels. Visual posts with images and videos were suggested to be effective methods of disseminating information on SoMe, with one study reporting that highly visual platforms such as Instagram and Snapchat are the most popular amongst students for medical education.

One study argued that SoMe platforms may be perceived as more user friendly and less academic than institutionally designed solutions. This study used Facebook and YouTube to deliver peer developed resources to alleviate stress and depression amongst recent entrants to medical school. Contributions were monitored by a trained mental health professional.
One concern with SoMe in medical education was the lack of critical appraisal performed by students on the information presented.\textsuperscript{57,95,108,116} Content creation by faculty members\textsuperscript{61,76,108} may be a solution to this but adds to the workload of the educators.\textsuperscript{114} Many studies argue that training is required for both staff and students to capitalise on the benefits of SoMe interventions in medical education.\textsuperscript{8,49,62,79,101,117,118}

### 3.5 Educational activity

Social media platforms were used in myriad ways. Educators used them to share resources, establish a dialogue with students, and facilitate classroom activities. Students used them for informal conversation, for accessing and sharing educational resources, to arrange educational and social events, to discuss opinions, and to participate in surveys, quizzes and educational games.\textsuperscript{80,109,110,113} Several studies delivered structured teaching using social media platforms.\textsuperscript{6,36,48,65,85} Use of social media within structured learning activities was found to improve communication and participation,\textsuperscript{32,88} and facilitate teamwork.\textsuperscript{62,82}

Several studies reported that SoMe initiatives made students feel more able to ask questions, in comparison to asking questions in a clinical setting,\textsuperscript{87} in part driven by an expectation that they would generally receive answers to questions more quickly from a multitude of voices, especially international professionals.\textsuperscript{74,119} Students reported a change in the student-educator relationship.\textsuperscript{14,37,120} They perceived SoMe as having “flattened hierarchies”, allowing students to feel more comfortable interacting with educators.\textsuperscript{14,121} Similar improvements furthered peer communication and working,\textsuperscript{62,88} with facilitators in one study noticing greater student collaboration with SoMe.\textsuperscript{62} Students reported newfound appreciation and interest for the subject material when presented with novel SoMe options for learning for example videos on YouTube\textsuperscript{58} and discussion groups on Facebook.\textsuperscript{54,88}

SoMe were also commonly used to supplement classroom teaching.\textsuperscript{62,67,72,73,74} One study utilised SoMe to create a ‘flipped classroom’ in which a case was reviewed by participants on Twitter before classroom sessions.\textsuperscript{67} Classroom discussions were described as more efficient, and Twitter lessened the educator burden. SoMe was found to enhance communication and collaboration between students and educators\textsuperscript{69,72,74,106,112} as well as students and their peers.\textsuperscript{62,69,72,83,106,112,122}

Students reported SoMe facilitated them being more productive in their studies,\textsuperscript{16} and enabling them to work faster.\textsuperscript{120} Video-based SoMe platforms were considered to be particularly useful.\textsuperscript{95,114,116} SoMe were also considered useful in supporting reflective practice.\textsuperscript{46}

Finally, SoMe were considered useful in learners’ professional identity formation.\textsuperscript{23,118}

### 3.6 Academic performance

Subjective improvements in students’ self-reported performance\textsuperscript{56,60,63,123,124} and objective increases in assessment outcomes\textsuperscript{33,35,38,40,69,82,125} were reported, including in two studies with controlled designs.\textsuperscript{32,80} However, it is unclear whether the differences found in the latter two studies result from the use of SoMe or other confounding variables such as weekly assessment\textsuperscript{32} or earlier exposure to educational material.\textsuperscript{60} Some studies showed that the rate of utilisation of SoMe positively correlated with test score.\textsuperscript{37,39,63} However, others suggested time spent on SoMe had no impact on knowledge improvement.\textsuperscript{126} Furthermore, most interventions tested participants’ knowledge and skills acquisition in the short term, with limited findings on long-term effectiveness.\textsuperscript{46} Knowledge retention in the long term may be poorer with SoMe learning compared to lecture-based learning.\textsuperscript{39}

Three studies demonstrated no effect from SoMe on academic performance.\textsuperscript{34,36,41} However, one studied ‘technology enhanced learning’ (including SoMe), making it difficult to identify any SoMe specific results\textsuperscript{61} and another did not assess baseline knowledge between the experimental and control group prior to intervention.\textsuperscript{34}

In one study, students linked social media use to subjectively worse test performance.\textsuperscript{127}

The studies reporting higher academic performance tended to be dialogue-focussed, using push notifications, poll quizzes and closed communities of practice. They generally used WhatsApp, Facebook Groups and WeChat.

### 3.7 Professionalism

One concern with the use of SoMe was maintaining professionalism. One study assessed students’ Facebook activity levels and perceptions of guidelines for professional behaviour,\textsuperscript{128} while another tested their response to simulated medical professionalism scenarios.\textsuperscript{129} The latter demonstrated an inverse correlation between having a personal board on PTT, Taiwan’s largest SoMe platform and humanism, as well as a similar correlation between SoMe use and medical professionalism scores.\textsuperscript{128} This was despite the fact that SoMe use was actually associated with increased awareness of the need for professional behaviour on SoMe.\textsuperscript{128} Additionally, despite this negative correlation, some students highlighted that SoMe, specifically Twitter, allowed them to develop empathy and understand previously unknown elements of patient experience.\textsuperscript{119}

One study indicated that simply by surveying students on contentious behaviours (in this case patient-targeted googling, PTG) the act of surveying may reduce the incidence of such behaviours.\textsuperscript{130} Students involved favoured more explicit teaching around PTG and such surveys may represent an opportunity to improve levels of professionalism.

Multiple studies found that students’ behaviours were subject to change with the knowledge that their peers would be able to view the content they posted.\textsuperscript{45,52} This may be linked to increasing awareness of professionalism, and one study found that approximately 11% of the SoMe profiles assessed were deemed to have committed some form of violation.\textsuperscript{131}
<table>
<thead>
<tr>
<th>Domain</th>
<th>MERSQI Item</th>
<th>Studies</th>
<th>Score</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No. (%)</td>
<td>Maximum</td>
<td>Item</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Domain</td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. Study design</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Single group cross-sectional or single group post-test only</td>
<td>63</td>
<td>75%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Single group pre-test and post-test</td>
<td>10</td>
<td>12%</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Nonrandomised, 2 group</td>
<td>8</td>
<td>10%</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Randomised control trial</td>
<td>3</td>
<td>4%</td>
<td>3</td>
</tr>
<tr>
<td>Sampling</td>
<td>2. No. of institutions studied</td>
<td></td>
<td>3</td>
<td>0.63 0.33 1.52 0.55</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>73</td>
<td>87%</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>1%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&gt;2</td>
<td>10</td>
<td>12%</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>3. Response rate, %</td>
<td></td>
<td>0.99 0.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>8</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;50 or not reported</td>
<td>32</td>
<td>38%</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>50-74</td>
<td>14</td>
<td>17%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>≥75</td>
<td>30</td>
<td>36%</td>
<td>1.5</td>
</tr>
<tr>
<td>Type of data</td>
<td>4. Type of data</td>
<td>3</td>
<td>1.57 0.91 1.57 0.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assessment by study participant</td>
<td>60</td>
<td>71%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Objective measurement</td>
<td>24</td>
<td>29%</td>
<td>3</td>
</tr>
<tr>
<td>Validity of evaluation instrument</td>
<td>5. Internal structure</td>
<td>3</td>
<td>0.29 0.46 1.06 1.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>4</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>57</td>
<td>68%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Reported</td>
<td>23</td>
<td>27%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6. Content</td>
<td></td>
<td>0.56 0.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>2</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>36</td>
<td>43%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Reported</td>
<td>46</td>
<td>55%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>7. Relationship to other variables</td>
<td></td>
<td>0.24 0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>1</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>63</td>
<td>75%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Reported</td>
<td>20</td>
<td>24%</td>
<td>1</td>
</tr>
<tr>
<td>Data analysis</td>
<td>8. Appropriateness of analysis</td>
<td></td>
<td>0.88 0.33 2.43 0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data analysis inappropriate for study design or type of data</td>
<td>10</td>
<td>12%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Data analysis appropriate for study design or type of data</td>
<td>74</td>
<td>88%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>9. Complexity of analysis</td>
<td></td>
<td>1.55 0.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Descriptive analysis only</td>
<td>38</td>
<td>45%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Beyond descriptive analysis</td>
<td>46</td>
<td>55%</td>
<td>2</td>
</tr>
<tr>
<td>Outcomes</td>
<td>10. Outcomes</td>
<td></td>
<td>1.31 0.39 1.31 0.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Satisfaction, attitudes, perceptions, opinions, general facts</td>
<td>47</td>
<td>56%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Knowledge, skills</td>
<td>22</td>
<td>26%</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Behaviours</td>
<td>15</td>
<td>18%</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Patient/health care outcome</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Total Score</td>
<td></td>
<td></td>
<td>18</td>
<td>9.11 2.30</td>
</tr>
</tbody>
</table>
Two articles indicated that in response to participating in the study, students would actively change their SoMe profiles to be less publicly accessible.52,54

Even though students understood the need for professionalism when using SoMe for medical education,49,93,107,125,132,133,134 they appreciated educational interventions with regards to this.40,56,135,136 Studies suggested privacy concerns were a barrier to effective SoMe use47,64,69 but closed SoMe groups or bespoke networks may alleviate such concerns.53,74 This does, however, limit global interactions53 and only three studies identified accessing expertise from around the world as an advantage.74,116,119

3.8 | Health risks

Medical students generally reported using SoMe for at least six hours per week,137 reporting poor health behaviours when using Facebook such as holding urine, skipping meals, and midnight logins, leading to disturbed sleep,103,138 headaches, back and shoulder pain and eye irritation.139 One study linked SoMe use to increased risk of anxiety & depression,140 whilst two suggested SoMe may contribute to social isolation.139,141

3.9 | Methodological quality of included papers

Eighty-four (75%) papers utilised quantitative (including mixed methods) designs. The mean (SD, range) MERSQI score for these papers was 9.1 (2.3, 5-14) out of 18 (Table 1). These studies predominantly employed a single group design with a single data collection point (63, 75%), at a single institution (73, 87%), using subjective data (60, 71%). Almost half (38, 45%) of these studies reported descriptive statistics only, although 88% were considered to be appropriate for the design and types of data collected. The domain with the poorest scores was the validity of the evaluation instruments used.

Thirty-one papers utilised qualitative designs. The mean (SD, Range) SRQR score was 9.03 (3.99, 2-16). Included studies were strongest at reporting the context (30/31, 97%), data collection methods (28/31, 90%), units of study (25/31, 81%), sampling strategies (23/31, 74%), data analysis methods (22/31, 71%), and ethical issues (22/31, 71%). They were weakest at providing rationales for techniques to enhance trustworthiness (2/31, 6%), data analysis methods (4/31, 13%), sampling strategies (7/31, 23%), data collection methods (12/31, 39%), and describing researcher characteristics and reflexivity (13/31, 42%).

Overall, the risk of bias in reporting educational interventions was reasonable (Figure 2). There were no sources of bias where over 50% of papers were considered to be at low risk of bias. The highest sources of potential bias were in reporting details regarding the settings in which interventions took place and in the educational methods used.

Finally, the modal (mean, range) strength of conclusions was 3 (2.95, 1-4) indicating that these conclusions could probably be based on the results.

4 | DISCUSSION

This review identified 112 articles studying SoMe use in undergraduate medical education. Students generally have a favourable view of using SoMe for their learning, particularly when engaging with content hosted on familiar platforms and provided in small chunks. Acceptability is driven by the perceived flattening of hierarchies, improved accessibility to faculty and platform novelty. There is some evidence of SoMe interventions resulting in improved learner performance, though this is mainly limited to self-report, non-controlled studies or short-term changes. SoMe interventions which encourage dialogue between educator and student or between peers are most effective. Whilst students have reservations about professional conduct on SoMe, educational interventions aimed at developing professional SoMe behaviours are appreciated and appear to be effective. The methodological quality of studies investigating SoMe in medical education remains poor.

4.1 | Methodological quality of primary studies

It is telling that our review has included 98 more studies than a review in 2013 and 83 more studies than one in 2017, demonstrating the proliferation of SoMe in educational scholarship.17,142 This review considers a wider variety of platforms that focussed on multidirectional communication than earlier investigations. Previous reviews were mostly comprised of blogs and Wikis,17 which were excluded in this review.

However, this profound increase in studies has not been matched by an increase in quality. The mean (SD, range) MERSQI of previous reviews were 8.89 (3.39, 5-15.5),17 and 9.57 (2.02, 7.5-14.5).142 Our review identified a mean between those of these two previous reviews, and a lower maximum MERSQI score. While ours is the first SoMe review to use the SRQR to assess methodological quality of qualitative research, Sterling’s review142 reported reasonable quality of included qualitative research using the COREQ criteria (Consolidated Criteria for Reporting Qualitative Research).143 This
suggests that while much of the literature on SoMe remains of relatively poor quality, research on SoMe in graduate medical education appears to be of higher quality.

Several previous studies have examined the impact of SoMe upon student outcomes, with focuses on the impact upon knowledge. Whilst Cheston’s review concluded that their intervention studies led to equivalent test scores for students who did and did not use SoMe tools, another found Facebook to be beneficial to the affective aspects of learning and YouTube to be an inadequate source of information for medical education. However, the latter review describes its study outcomes as weak, poorly transferrable and relying on self-reporting. Our results conversely demonstrate subjective and objective knowledge improvement across several platforms, however, cautions against the relatively strong likelihood of confounders, suggesting that the benefits are likely limited to the short term. We echo the findings of Chan et al, where the majority of studies consisted of descriptive research, assessed lower-level Kirkpatrick hierarchies and were methodologically dominated by single-centre quantitative surveys.

Whilst SoMe platforms have grown in number across reviews and time, the subject of an investigation by studies does not appear to have changed dramatically. The dominant themes explored by Cheston et al (2013), namely professionalism and improvement in knowledge, are reflected in this review, in addition to that of Sterling et al (2017) and Sutherland & Jalali (2017).

4.2 | Strengths and limitations

While this is not the first systematic review of SoMe use in undergraduate medical education, we have identified significantly more papers than previous reviews. We believe this review has benefited from a rigorous and sensitive search strategy including multiple databases, reference, citation, and hand searching. We have benefitted from the insights and interpretations of both medical students and medical school faculty in the extraction and synthesis of original papers. We have included studies evaluating interventions and those researching exposure to SoMe, resulting in a comprehensive synthesis.

These strengths notwithstanding, this review does have some limitations. Firstly, while we believe the focus on undergraduate medical education is a strength, it also limits the generalisability of these results to other groups of learners. SoMe are used in different ways by practising physicians and therefore assumptions cannot be drawn regarding outcomes in SoMe across the spectrum of medical education. We did not contact the authors of included studies for missing data, further details or to identify other relevant or forthcoming literature. We have only included published journal articles in this review and have excluded conference abstracts. This may have resulted in the omission of innovative SoMe approaches that have not yet materialised in the peer-reviewed literature. Finally, as with all reviews, the results of this review are limited by the quality of the primary studies available for inclusion.

4.3 | Implications for practice and research

For the educator designing social media education tools, the best practice appears to be grounded in familiarity and features intensive induction for staff and learners. The most subjectively appreciated interventions appear to be highly visual, curated by faculty and blended with classroom teaching. When considering the strongest evidence for improving objective outcomes, initiatives featuring collaborative, text-based discussion seem most effective. Examples of such activities include case-based discussion, SoMe journal clubs, poll-based quizzing and smaller, near-peer communities of practice. Existing studies suggest that instant messaging services such as WeChat and WhatsApp are most likely to provide intensive dialogue to facilitate learning. Perhaps unsurprisingly, students who engage more with SoMe comments perform best in objective assessments.

This review may inform educators in navigating SoMe professionalism concerns. Despite feeling that professionalism guidelines are poorly understood, the evidence presents a clear picture of a learner population motivated to champion professional practice on SoMe. Students appreciate professionalism in teaching, quickly respond to constructive criticism on SoMe privacy and even seek out professional development opportunities on platforms. In terms of practical guidance, it is clear that fear of appearing unprofessional in front of peers or patients can hamper engagement with SoMe initiatives. Therefore, the use of closed groups may provide reassurance to both educator and learner.

This review has a number of implications for the direction of future research. As detailed above, the community should prioritise undertaking fewer, higher quality studies, rather than the current high output of methodologically weak research. The quality of the current evidence base should provide a roadmap for this work. Considering this review’s quantitative studies, the majority of the highest quality articles (MERSQI 12-14, n = 12) investigated outcomes related to blended learning (9/12, 75%), compared to one each on professionalism, reflection and the humanities. Conversely, the lowest quality studies (MERSQI 5-7, n = 20) featured a majority of these latter categories. Professionalism (7/20) was particularly poorly investigated whilst SoMe use in the humanities represented three of the four lowest scored studies in this review.

Interestingly, this pattern was inverted in the review’s qualitative studies. Half of the top 10 highest SRQR scores (12-16), including the top two studies, focus on professionalism. This is perhaps an indication that investigating professionalism should be the domain of richer, in-depth qualitative research.

Whilst understanding professionalism was a key tenet of numerous papers in this review, how this is explored should be carefully considered. A recent study of unprofessional behaviour amongst vascular surgery trainees was retracted after a surge of concerns were raised by the medical community. Such complaints focussed on the ‘shaming’ of professionals, particularly females, and the invasive, covert methods applied by the authors.
to investigate unprofessionalism. It should be highlighted that our review includes three papers that systematically searched for participant profiles, collected or reviewed personal data including photographs and made subjective judgements regarding their professionalism. Such methods are at high risk of researcher prejudice and there remains no validated tool to measure SoMe professionalism. We would suggest that future researchers maintain their own professional and ethical standards by avoiding invasive, subjective judgements and instead pursue higher quality methods of investigating such a complex phenomenon. This review serves as a reminder to educators that despite regulator guidance and much research, there is limited consensus on SoMe professionalism.

Considering the widespread perceived concerns regarding SoMe professionalism and associated guidelines, particularly given the proven hindrance such fears have upon learner engagement, we suggest that the next step for SoMe research lies in this domain. We call for a rigorous investigation to build a community consensus on SoMe professionalism.

We must also urgently reflect on how we undertake research in this field. Whilst we have examined research encompassing a diverse range of rapidly emerging platforms, we suggest a reactionary approach to SoMe research is inappropriate. Future work should focus instead on the common factors across the spectrum of social media in order to address the significant gaps raised by this review. Specifically, we recommend the study of the impact of a range of SoMe platforms upon long-term knowledge retention, largely absent from the current literature. Additionally, a robust assessment of measures to prevent the SoMe harms highlighted in this review is essential.

When considering commonalities between platforms, one must consider the underpinning theory on SoMe. SoMe platforms have their functional differences but educationally the principles are the same: these are rapid, often short-lived communities of practice built on complex socially constructed values, which themselves fluctuate across locations and generations. Brief community lifespan perhaps explains why long-term outcomes have so far proven challenging to establish. The ever-changing norms of communities and tensions between clinical environments and online spaces may contribute to a difficulty in defining SoMe professionalism.

Moreover, placing our results in the context of theoretical principles may advance deeper thinking on SoMe in health professions education. That objective performance benefit is driven by rapid peer-led dialogue highlights the primacy of the community of practice in effective SoMe interventions. Community of practice as a theoretical construct underpins the majority of the most effective interventions in this review and has been at the focus of rigorous investigations on enhancing knowledge translation in wider healthcare education. Connectivism, whilst a theory literally established for the online environment, is largely knowledge-centred rather than community-focused. Community appears to dominate social media educational practice, shaping both effective and ineffective learning cultures. This suggests that communities of practice are a more coherent and informative theoretical construct than connectivism in explaining effectiveness in social media education.

5 | CONCLUSIONS

Despite an explosion of research surrounding social media in medical education, understanding this social phenomenon has not significantly progressed in almost a decade. We have established that social media is enjoyable for students, may improve short term knowledge retention and can aid communication between learners and educators. However, students and educators alike remain wary of professionalism concerns and warnings against potential SoMe harms remain.

We suggest that rather than attempting to undertake a superficial evaluation of the latest SoMe trend, the community should instead consider longer-term, higher quality research, rooted in the underpinning educational theories which unite these diverse platforms.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

We have no declarations of interest to declare.

AUTHOR CONTRIBUTIONS

JG designed the study, extracted data, synthesised the results and contributed to writing the first draft of the manuscript and subsequent drafts. MU extracted data, synthesised results, and contributed to all drafts of the manuscript. AA extracted data, synthesised results, and contributed to all drafts of the manuscript. OB extracted data, synthesised results, and contributed to all drafts of the manuscript. AO extracted data, synthesised results, and contributed to all drafts of the manuscript. TC extracted data, synthesised results, and contributed to all drafts of the manuscript. ELR designed the study, extracted data, synthesised the results and contributed to writing the first draft of the manuscript and subsequent drafts. All authors approve of the final manuscript for submission.

ETHICAL APPROVAL

Ethical approval not sought as this is secondary research not involving human participants.

ORCID

Jonathan Guckian https://orcid.org/0000-0002-8162-1583
Mrudula Utukuri https://orcid.org/0000-0003-1510-469X
Aqua Asif https://orcid.org/0000-0003-3233-7428
Oliver Burton https://orcid.org/0000-0003-3569-069X
Joshua Adeyaju https://orcid.org/0000-0001-5003-6990
Adam Oumeziane https://orcid.org/0000-0003-2111-6996
Timothy Chu https://orcid.org/0000-0003-4579-1136
Eliot L. Rees https://orcid.org/0000-0002-6458-5808
REFERENCES

2. Mander J. Internet users have average of 7 social accounts. Globalwebindex.


95. Ben-Yakov M, Kayssi A, Bernardo JD, Hicks CM, Devon K. Do emergency physicians and medical students find it unethical to “Look up” their patients on Facebook or Google? West J

---

GUCKIAN ET AL.
126. Kitsis EA, Milan FB, Cohen HW, et al. Who’s misbehaving? Perceptions of unprofessional social media use by medical...


SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.