




OPEN ACCESS

# The value of Facebook in nationwide hospital quality assessment: a national mixed-methods study in Norway

Oyvind Bjertnaes , Hilde Hestad Iversen, Katrine Damgaard Skyrud, Kirsten Danielsen

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2019-009456>).

Nasjonalt folkehelseinstitutt, Oslo, Norway

## Correspondence to

Dr Oyvind Bjertnaes, Nasjonalt folkehelseinstitutt, Oslo 0403, Norway; [oyvindandresen.bjertnaes@fhi.no](mailto:oyvindandresen.bjertnaes@fhi.no)

Received 13 February 2019

Revised 14 August 2019

Accepted 17 August 2019

Published Online First

24 August 2019

## ABSTRACT

**Objectives** The objective was to assess the possibility of using a combination of official and unofficial Facebook ratings and comments as a basis for nation-wide hospital quality assessments in Norway.

**Methods** All hospitals from a national cross-sectional patient experience survey in 2015 were matched with corresponding Facebook ratings. Facebook ratings were correlated with both case-mix adjusted and unadjusted patient-reported experience scores, with separate analysis for hospitals with official site ratings and hospitals with unofficial site ratings. Facebook ratings were also correlated with patient-reported incident scores, hospital size, 30-day mortality and 30-day readmission. Facebook comments from 20 randomly selected hospitals were analysed, contrasting the content and sentiments of official versus unofficial Facebook pages.

**Results** Facebook ratings were significantly correlated with most patient-reported indicators, with the highest correlations relating to unadjusted scores for organisation (0.60,  $p<0.000$ ) and nursing services (0.57,  $p<0.000$ ). Facebook ratings were significantly correlated with hospital size ( $-0.40$ ,  $p=0.003$ ) and 30-day mortality (0.31,  $p=0.040$ ). Sentiment analysis showed that 84.7% of the comments from unofficial Facebook sites included neutral comments that did not give any specific description of experiences of the quality of care at the hospital. Content analysis identified common themes on official and unofficial Facebook pages.

**Conclusions** Facebook ratings were associated with patient-reported indicators, hospital size, and 30-day mortality. Qualitative comments from official Facebook are more relevant for hospital evaluation than unofficial sites. More research is needed on using Facebook ratings as a standalone indicator of patient experiences in national quality measurement, and such ratings should be reported together with research-based patient experience indicators and with explicit criteria for the inclusion of unofficial sites.

## BACKGROUND

The internet offers a large number of websites where healthcare services might be rated, ranging from commercial sites like Yelp.com to public sites like NHS Choices. Research on web rating sites has

grown rapidly in recent years, covering topics such as the content of comments on rating sites,<sup>1–3</sup> efforts to produce quantitative scores from qualitative comments,<sup>4</sup> and correlation studies between ratings on websites and established quality indicators such as patient experiences, mortality and readmission.<sup>5–9</sup> Previous research shows that rating sites include comments on similar topics to research-based patient experience questionnaires but also that various additional topics are described.<sup>1–3</sup> Furthermore, previous research shows a low to moderate association between scores on rating sites and indicators from systematic patient experience surveys.<sup>5–9</sup>

Facebook ratings have been included in several previous studies, mostly for hospital services<sup>4 10–14</sup> but also for nursing homes.<sup>15</sup> As for other web rating sites, Facebook studies in the hospital setting have shown low to moderate association between Facebook ratings and indicators from systematic patient experience surveys. However, previous correlation studies mostly compare Facebook ratings with a narrow set of systematic patient experience indicators, some only including overall satisfaction, hospital recommendation or a composite score.<sup>4 12</sup> Furthermore, most previous studies are from the USA and some have unclear coverage or very low coverage of hospitals.<sup>4 11</sup> For instance, the study by Huppertz *et al*<sup>4</sup> analysed less than 20% of the included hospitals. Thus, the generalisability of findings from previous hospital correlation studies is highly unclear.

To obtain official ratings on Facebook, hospitals must first establish an official Facebook page, and then actively turn on or allow public ratings on this site. Apart



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Bjertnaes O, Iversen HH, Skyrud KD, *et al*. *BMJ Qual Saf* 2020;**29**:217–224.

from the study by Perez and Freedman,<sup>12</sup> previous studies indicate large coverage challenges, both in the hospital setting<sup>4 11</sup> and in other settings.<sup>15</sup> Furthermore, the study by Campbell and Li<sup>11</sup> showed that almost 30% of Facebook-registered hospitals did not allow public ratings on their site. There are a number of possible reasons for hospitals not having a Facebook site or not allowing public ratings. However, an association between these reasons and the outcome of interest (hospital quality) would highly question the validity of using Facebook as a source for hospital quality. One particular problem would be if poor performing hospitals were over-represented in the group of non-Facebook users. In addition to the coverage challenge, hospital organisation might also challenge a useful role for Facebook as a source for information on hospital quality. Hospitals might be formally connected to hospital systems<sup>11</sup> or hospital trusts,<sup>6</sup> thereby allowing for the opportunity for system-level or trust-level Facebook sites with common ratings for several underlying hospitals. Such aggregated ratings are less relevant for hospitals, the public and patients, since they mask differences between hospitals within the system or trust. To reduce the coverage challenge mentioned above, the current study used unofficial Facebook pages for hospitals lacking an official page. The unit of analysis was the hospital, and providers without hospital level Facebook and patient experience data were excluded.

One important source for nation-wide hospital quality assessments in Norway is the national quality indicator system. The national quality indicator system consists of systematically derived quality indicators, and is established to support quality improvement, management of health services, decision-making by patients and public accountability. To be useful in this context, the Facebook ratings should demonstrate an association with other quality indicators, especially indicators from systematic patient experience surveys. The main objective of our study was to assess the possibility of using a combination of official and unofficial Facebook ratings as a basis for nation-wide hospital quality assessments. We assessed the association between Facebook ratings and patient-reported experience measures (PREMs), using the latter as a gold standard, and with separate analysis for hospitals with official Facebook ratings and hospitals with unofficial ratings. Based on previous research, we hypothesised a low to moderate association with PREMs at the hospital level. We supplemented the case-mix adjusted PREMs with unadjusted indicators, to be able to assess the potentially different correlations between Facebook and adjusted/unadjusted indicators. Facebook ratings are unadjusted, and we therefore hypothesised a higher correlation with unadjusted than with adjusted indicators. A secondary objective of our study was to make an in-depth assessment of the content and quality of unofficial Facebook ratings, since no previous research

has referred to such sites. We conducted a qualitative analysis of a random sample of Facebook comments, contrasting the content of official versus unofficial Facebook pages in respect of both patient experience topics mentioned and the magnitude of real patient experiences described. To our knowledge, this is the first study analysing quantitative and qualitative data from official and unofficial Facebook rating sites, in the context of national quality measurement.

## METHODS

The study was a secondary analysis of hospital-level quality indicators and Facebook reviews. All hospitals in Norway with patient-reported experience scores from a national survey in the autumn of 2015 were matched with corresponding Facebook ratings. 30-day mortality and 30-day readmission data for 2015 were collected from the national quality indicator system and merged with the other indicators at hospital level.

### Facebook data

The main criterion for matching Facebook and patient experience survey data was having Facebook reviews at the same level as the hospital survey, either from an official or unofficial Facebook page. We collected Facebook ratings (average score on a scale from 1 to 5) and the number of ratings from the included sites. First, two researchers (OB and KD) conducted an independent collection of Facebook reviews and ratings in the weeks 47–48 in 2018, based on the list of hospitals from the 2015 patient experience survey. The average rating of all ratings ever were included. Second, the two researchers met and compared the collected Facebook ratings and numbers for each hospital. Differences were resolved by conducting searches together in Facebook to verify the correct site, rating and number.

For analysis of qualitative comments, we randomly selected four small hospitals, three medium hospitals and three large hospitals for both official and unofficial Facebook sites. A minimum number of qualitative comments is necessary to establish a measure of sentiment, but we did not find any guidance in the literature on the statistical properties for the minimum text size needed to perform the analysis. To establish a criterion, we relied on previous research,<sup>16</sup> and up to 50 posts were randomly selected from each hospital. We only considered hospitals with more than 50 registered comments as relevant for inclusion. This process yielded 876 codable statements from 20 hospitals. Official versus unofficial sites were compared by exploring both sentiments and topics in the Facebook comments. Two researchers (HI and KD) independently coded the responses through an inductive approach and selected by consensus the most common themes for presentation in this study. The comments were classified as positive, negative, both positive and negative, or neutral. We used content analysis to examine the main themes. Neutral comments that did not address the

hospital or give any specific evaluation of the health-care quality or experience were not included in the content analyses. Three typical examples of these are as follows: (1) 'At the hospital again', (2) 'Good night' and (3) 'Wondering if I can go home today'.

### Patient-reported indicators

The patient experience survey has been described elsewhere.<sup>17</sup> In short, the questionnaire was sent to 400 randomly selected adult inpatients who were discharged from each hospital in Norway in a 3-month period in the autumn of 2015. Patients were sent a postal invitation with both a pen-and-paper questionnaire and an electronic response option. Non-respondents were sent up to two reminders. The response rate was 58.7% (n=12 844). Non-response weights were computed to compensate for non-response, accounting for patient age, gender, the number of diagnoses, admission type (emergency, elective) and length of stay.

The questionnaire comprised 62 closed-ended items,<sup>17</sup> including patient experience questions, patient incident questions<sup>18</sup> and background questions. Patient experience questions relate to aspects of structure and process, whereas patient incident questions concern the occurrence of specific safety incidents during the hospital stay. Most items had a 5-point response format ranging from 1 ('not at all') to 5 ('to a very large extent'). A total of 34 items related to patient experiences were basis for nine quality indicators: waiting time (one item), standard (six items), contact with next-of-kin (two items), organisation (four items), doctor services (seven items), nursing services (seven items), information (three items), discharge information (two items) and cooperation with other health services (two items). The patient experience questions are described in online supplementary appendix A. The nine PREMs were computed for each hospital and delivered to the national quality indicator system. Except one indicator on waiting time, the PREMs were adjusted for self-perceived health, patient age, admission type, Charlson Comorbidity Index, respondent type (patient alone or patient with help) and the number of admissions in the last 2 years. In addition, a patient-incident index consisting of 12 safety items (patient-reported incident in hospital instrument (PRIH-I)) was computed at the hospital level, and reported in the hospital reports. In the current study, we used the 10 indicators mentioned above, all scored 0–100 where 100 is the best possible score. Furthermore, we created a composite patient experience indicator by calculating the arithmetic mean of eight of the patient experience indicators. The waiting time indicator is different from the other PREMs, and is not part of the composite score and not adjusted. The indicator consists of one item about a concrete aspect prior to the hospital stay, and only includes elective patients. Furthermore, the available case-mix variables for PREMs perform poorly for this indicator (low explained variance), indicating less need

to adjust for these variables. The incident indicator (PRIH-I) is conceptually different from the PREMs, measuring patient safety from the patient perspective.<sup>18</sup> Currently, the indicator is not compared across hospitals in the reports from the surveys, and thus lack a validated case-mix model.

### Mortality and readmission rates

The 30-day mortality and readmission rates used in this study are hospital-wide measures based on administrative data from the year 2015.<sup>19 20</sup> We only used case-mix adjusted indicators on mortality and readmission, which were collected from the national quality indicator system. The 30-day mortality rate is case-mix adjusted for age, gender, pre-admissions (number of previous admissions during the last 2 years), Charlson Comorbidity Index, admission type (elective or emergency) and Clinical Classifications Software category. The 30-day readmission rate is case-mix adjusted for age, gender and type of diagnosis. Readmission was defined as an emergency admission, regardless of reason and hospital for readmission, occurring between 8 hours and 30 days following discharge from a previous hospital stay (primary stay). The number of hospitals in correlation analysis with 30-day mortality and 30-day readmission was 44.

### Hospital size

Patient experiences are known to correlate with hospital size, with smaller hospitals having better experience scores.<sup>21 22</sup> Hospital size was not easily available from official statistics for all hospitals, so we used data from the survey sample construction to calculate hospital size. Each eligible hospital patient had a known probability of being selected in the sample, a fixed property at the hospital level calculated as one divided by the number of eligible patients. Hospital size was calculated as the actual sample size for the hospital divided by the hospital level probability of being selected, using the last inclusion month as a basis for the estimations. We used hospital size as a continuous variable in the correlation analysis, and as a categorical variable in the descriptive analysis: small (<301 discharges per month), medium (301–900 discharges per month) and large (>900 discharges per month).

### Statistical analysis

Facebook ratings at the hospital level were correlated with both case-mix adjusted and unadjusted PREMs, with separate analysis for hospitals with official site ratings and hospitals with unofficial site ratings. The Spearman's rank correlation coefficient was used in the correlation analysis. We conducted sensitivity analysis by re-running analysis without outliers and without hospitals with low n in Facebook ratings (n<6). Statistical outliers were identified in a bivariate regression with Facebook rating as dependent and patient experience scores as independent variables, using Cook's

**Table 1** Facebook site status by hospital size, Facebook ratings and patient experience indicators

	Official Facebook site (n=16)	Unofficial Facebook site (n=38)	Total (n=54)
Hospital size			
Small hospitals, n (%)	6 (27.3)	16 (72.7)	22 (100)
Medium hospitals, n (%)	3 (17.6)	14 (82.4)	17 (100)
Large hospitals, n (%)	7 (46.7)	8 (53.3)	15 (100)
Facebook variables			
Number of ratings, mean (min-max)	156.8 (22–476)	103.7 (1–338)	119.4 (1–476)
Ratings, mean (min-max)*	4.3 (3.4–4.8)	4.0 (2.2–5)	4.1 (2.2–5)
<i>Patient experience indicators (adjusted):†</i>			
Composite score, mean (min-max)	72.5 (65.9–78.7)	71.6 (64.5–79.6)	71.9 (64.5–79.6)
Nursing services, mean (min-max)	78.9 (70.6–85.0)	77.8 (72.3–85.0)	78.1 (70.6–85.0)
Doctor services, mean (min-max)	77.6 (73.0–82.0)	76.6 (69.6–82.6)	76.9 (69.6–82.6)
Information, mean (min-max)	74.1 (69.2–80.1)	73.3 (66.3–79.6)	73.5 (66.3–80.1)
Organisation, mean (min-max)	70.6 (61.2–78.8)	69.5 (62.9–81.6)	69.8 (61.2–81.6)
Contact with next-of-kin, mean (min-max)	78.7 (70.9–89.0)	78.4 (68.9–86.7)	78.5 (68.9–89.0)
Standard, mean (min-max)	75.2 (65.8–83.6)	73.3 (61.6–89.2)	73.9 (61.6–89.2)
Discharge information, mean (min-max)	61.7 (54.4–69.1)	58.7 (46.9–70.4)	59.6 (46.9–70.4)
Cooperation with other health services	63.6 (56.6–70.8)	65.3 (57.1–79.5)	64.8 (56.6–79.5)
Waiting time (elective patients), unadjusted	65.5 (57.9–76.5)	65.2 (52.5–76.9)	65.3 (52.5–76.9)
Patient-reported incidents (PRIH-I), unadjusted	88.2 (80.5–92.9)	87.3 (81.0–93.8)	87.6 (80.5–93.8)

\*Rated 1–5, where 5 represents the best possible rating.

†Scored 0–100, where 100 represents the best possible score. All indicators were adjusted for the following case-mix variables, except waiting time and PRIH-I: self-reported health, admission type (emergency, elective), Charlson Comorbidity Index, type of respondent (patient or proxy) and number of admissions in last 2 years.

PRIH-I, patient-reported incident in hospital instrument.

$D > 4/n$  as criteria for detecting outliers.<sup>23</sup> Facebook ratings were also correlated with patient-reported incident scores (PRIH-I), hospital size, 30-day mortality and 30-day readmission.

## RESULTS

### Characteristics of the sample

In total, 54 of 61 hospitals (88.5%) had Facebook ratings, of which 16 had an official Facebook site and 38 an unofficial site. For large hospitals, 46.7% had an official Facebook site; the corresponding figure was 17.6% for medium hospitals and 27.3% for small hospitals (table 1). The average number of Facebook reviews was 156.8 for hospitals with official pages (min: 22, max: 476) and 103.7 for hospitals with unofficial sites (min: 1, max: 338), with somewhat higher ratings for official sites (4.3) versus unofficial sites (4.0).

### Quantitative analysis

Facebook ratings were significantly correlated with 9 of 11 unadjusted patient-reported indicators, with five correlations being at or above 0.5 (table 2): composite score patient experiences (0.54,  $p < 0.000$ ), nursing services (0.57,  $p < 0.000$ ), organisation (0.60,  $p < 0.000$ ), standard (0.51,  $p < 0.000$ ) and PRIH-I (0.50,  $p < 0.000$ ). All correlations with adjusted patient-reported experience scores were significant and two were at or above 0.5: nursing services (0.51,  $p < 0.000$ ) and organisation (0.50,  $p < 0.000$ ). A scatterplot on the association between Facebook ratings and the adjusted

patient experience composite indicator is presented in figure 1 (0.48,  $p < 0.000$ ). Apart from the indicator on cooperation with other health services, all correlations between official Facebook ratings and patient experience indicators were significant, varying from 0.51 ( $p = 0.045$ ) for the unadjusted waiting time indicator to 0.82 ( $p < 0.000$ ) for the unadjusted organisation indicator. For hospitals with unofficial Facebook pages, the majority of correlations with patient experience indicators were significant, and these correlations varied from 0.33 ( $p = 0.040$ ) for the unadjusted indicator cooperation with other health services to 0.48 ( $p = 0.002$ ) for the unadjusted organisation indicator.

A bivariate regression with Facebook ratings as dependent variable and the adjusted patient experience composite score as independent variable showed that three hospitals had Cook's D values above  $4/n$ . All of these had an unofficial Facebook site. Correlations between unofficial Facebook ratings and adjusted patient experience indicators without the statistical outliers were higher for all indicators (table 3): all correlations were significant and the increase in correlation coefficient varied from 0.05 (cooperation with other health services) to 0.14 (nursing services, organisation, contact with next-of-kin and discharge information). Five hospitals had five or fewer ratings, all of them having an unofficial Facebook site. With the exclusion of these five hospitals, all remaining hospitals with unofficial ratings had more than 30 ratings. Correlations between unofficial Facebook ratings and adjusted patient experience



**Table 2** Correlations between Facebook ratings and patient experience indicators,\* by Facebook site status

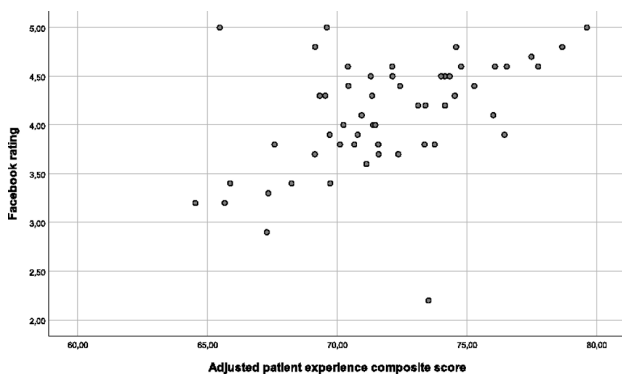
	Ratings official sites (n=16)	P value	Ratings unofficial sites (n=38)	P value	Total (n=54)	P value
Unadjusted for case-mix						
Composite score patient experiences	0.76	0.001	0.43	0.006	0.54	<0.001
Nursing services	0.8	<0.001	0.41	0.011	0.57	<0.001
Doctor services	0.64	0.008	0.36	0.026	0.47	<0.001
Information	0.69	0.003	0.35	0.032	0.46	<0.001
Organisation	0.82	<0.001	0.48	0.002	0.6	<0.001
Contact with next-of-kin	0.76	0.001	0.43	0.007	0.49	<0.001
Standard	0.64	0.008	0.45	0.005	0.51	<0.001
Discharge information	0.69	0.003	0.37	0.024	0.49	<0.001
Cooperation with other health services	0.2	0.455	0.33	0.04	0.21	0.138
Waiting time (elective patients)	0.51	0.045	-0.21	0.213	-0.1	0.489
Patient-reported incidents (PRIH-I)	0.77	0.001	0.36	0.025	0.5	<0.001
Adjusted for case-mix:†						
Composite score patient experiences	0.72	0.002	0.36	0.025	0.48	<0.001
Nursing services	0.72	0.002	0.35	0.031	0.51	<0.001
Doctor services	0.59	0.017	0.29	0.079	0.41	0.002
Information	0.65	0.007	0.24	0.155	0.36	0.008
Organisation	0.75	0.001	0.35	0.033	0.5	<0.001
Contact with next-of-kin	0.63	0.009	0.42	0.009	0.46	0.001
Standard	0.6	0.015	0.45	0.005	0.49	<0.001
Discharge information	0.63	0.008	0.28	0.093	0.4	0.003
Cooperation with other health services	0.36	0.176	0.36	0.026	0.27	0.048
Waiting time (elective patients)	–	–	–	–	–	–
Patient-reported incidents (PRIH-I)	–	–	–	–	–	–

\*Spearman's rank correlation coefficients.

†Case-mix variables: self-reported health, admission type (emergency, elective), Charlson Comorbidity Index, type of respondent (patient or proxy) and number of admissions in last 2 years.

PRIH-I, patient-reported incident in hospital instrument.

indicators without these five hospitals were higher for all indicators (table 3): all correlations were significant and the increase in correlation coefficient varied from 0.10



**Figure 1** Scatterplot on the association between Facebook ratings and the adjusted patient experience composite indicator.

(cooperation with other health services, increased from 0.36 to 0.46) to 0.26 (nursing services, increased from 0.35 to 0.61). Two of the three hospitals identified as outliers using Cook's D were among the hospitals with five or fewer ratings.

Facebook ratings were significantly correlated with hospital size ( $-0.40$ ,  $p=0.003$ ) and 30-day mortality ( $0.31$ ,  $p=0.040$ ), but not with 30-day readmission (table 4). We found no significant associations between 30-day mortality and Facebook rating in separate analysis for official ( $0.26$ ,  $p=0.436$ ) and unofficial sites ( $0.23$ ,  $p=0.189$ ).

#### Qualitative analysis

We classified 84.7% of the 392 comments from official Facebook sites as positive, 9.3% as negative, 4.1% as both positive and negative, and 2.1% as neutral. Half of the neutral comments ( $n=4$ ) did not address

**Table 3** Correlations between unofficial Facebook ratings and adjusted patient experience indicators, without statistical outliers and hospitals with five or fewer ratings‡

	All except statistical outliers (n=35)*	P value	All except hospitals with fewer than six ratings (n=33)†	P value	All unofficial hospital ratings (n=38)	P value
Composite score patient experiences	0.49	0.001	0.59	<0.001	0.36	0.025
Nursing services	0.49	0.001	0.61	<0.001	0.35	0.031
Doctor services	0.39	0.01	0.48	0.005	0.29	0.079
Information	0.35	0.021	0.45	0.009	0.24	0.155
Organisation	0.49	0.002	0.55	0.001	0.35	0.033
Contact with next-of-kin	0.56	<0.001	0.56	0.001	0.42	0.009
Standard	0.53	<0.001	0.62	<0.001	0.45	0.005
Discharge information	0.42	0.006	0.51	0.002	0.28	0.093
Cooperation with other health services	0.41	0.008	0.46	0.008	0.36	0.026

\*Three hospitals were statistical outliers, all of them with unofficial Facebook pages and one from each hospital size category.

†Five hospitals had fewer than six ratings, all of them with unofficial Facebook pages, and with two small-sized and three medium-sized hospitals.

‡Spearman's rank correlation coefficients.

the hospital, and were excluded from content analyses. 11.4% of the 484 posts from unofficial sites were categorised as positive, 3.7% as negative, 0.2% as both and 84.7% as neutral. 73.9% of the neutral comments addressed the hospital, but did not give any specific evaluation of healthcare quality. The remaining neutral comments either were unrelated to a healthcare issue or to the hospital or did not contain a meaningful unit of information. Accordingly, all neutral comments were excluded from content analyses. The content analysis identified the following common themes or domains related to hospital quality assessment on official and unofficial Facebook pages: global remarks, clinicians and staff, care/treatment, hospital standard or facilities, organisation and information/communication.

## DISCUSSION

Facebook ratings at the hospital level were associated with patient-reported indicators, hospital size and 30-day mortality. Ratings from official Facebook sites were significantly associated with all except one patient experience indicator, while ratings from unofficial Facebook sites were significantly associated with the majority of patient experience indicators. Qualitative comments from official Facebook sites were much more related to hospital evaluation than unofficial sites.

The main objective of our study was to assess the possibility of using a combination of official and unofficial

Facebook ratings as a basis for nation-wide hospital quality assessments in Norway. The study showed that only 16 of 61 hospitals had an official Facebook site with ratings, making such sites alone insufficient as a basis for national indicators. However, the strong associations between ratings from official Facebook sites and systematic patient-reported indicators show that these have the potential to be used to inform about patient experiences. With the inclusion of unofficial Facebook sites, the coverage of hospitals increased to 54 out of 61. Ratings from unofficial Facebook sites were significantly associated with the majority of patient experience indicators. Furthermore, omitting the statistically weakest unofficial sites from the analysis (<6 ratings) resulted in six of nine associations being above 0.5 for unofficial sites and all associations were significant. Thus, the level of associations for unofficial sites in our study was in line with the associations found in a previous US hospital study using only official Facebook sites.<sup>11</sup> However, except for one previous study,<sup>14</sup> we included much less hospitals than other studies using Facebook data,<sup>10–13</sup> with large statistical uncertainty also for our total hospital sample. Consequently, the p values in our study only give confidence that there are associations, rather than giving confidence about the strength of the associations.

Related to the main objective of our study, a secondary objective was to make an in-depth assessment of the content and quality of unofficial Facebook ratings.

**Table 4** Correlations between Facebook ratings, hospital size and register-based quality indicators, by Facebook site status\*

	Ratings official sites	P value	Ratings unofficial sites	P value	Total	P value
Hospital size	−0.60	0.014	−0.40	0.014	−0.40	0.003
30-day mortality†	0.26	0.436	0.23	0.189	0.31	0.040
30-day readmission†	0.28	0.397	0.01	0.939	0.12	0.455

\*Spearman's rank correlation coefficients.

†The number of hospitals in correlation analysis with 30-day mortality and 30-day readmission was 44. Both indicators were case-mix adjusted.

Hospitals with official Facebook sites had somewhat better ratings than hospitals with unofficial sites, but the difference could be due to the sample of hospitals in each group. Furthermore, the inclusion of unofficial sites reduced correlations with other quality indicators, but still all nine associations with adjusted patient-reported indicators in the total hospital sample were significant and most were medium strength correlations. Content analysis identified common themes on official and unofficial Facebook pages, but the majority of comments from unofficial Facebook sites included neutral comments that did not give any specific description of experiences and accordingly seem to have limited potential to reflect the quality of care in its current form. A majority of comments from official Facebook sites included descriptions of experiences, and most were categorised as positive. Thus, even though more caution is required for unofficial Facebook sites, the ratings and parts of the comments seem to tap into the construct of quality at the hospital level.

The ability to compensate for non-response and adjust for case-mix in quality comparisons are strengths of the survey-based approach, and the main reasons for treating this approach as the gold standard. A survey without these corrections is more similar to Facebook ratings, which is why we generally would expect stronger associations between Facebook and unadjusted indicators. However, both previous research<sup>24 25</sup> and our national quality measurements show rather modest effects of case-mix adjustment. The results of our study showed small changes in associations with Facebook ratings when using adjusted instead of unadjusted indicators, but the small number of hospitals gives large statistical uncertainty. Given the knowledge base on the modest effects of case-mix,<sup>24 25</sup> and the indications from our study, we believe that the case-mix argument is not decisive for the possibility of using Facebook ratings in quality measurement. However, future research should assess methods for correcting Facebook ratings for non-response and case-mix, for instance by using effect estimates from previous surveys. We also argue that quality measurements using Facebook ratings should integrate these with research-based patient experience indicators when reporting results, giving an open and balanced representation of systematically and unsystematically derived indicators. Reporting efforts using Facebook data should also include information like the number of ratings, the minimum number of ratings to be included, rating scores, status of Facebook site (official, unofficial) and the exact Facebook addresses. These implementation issues are important because Facebook ratings still have not been used in national quality measurement systems but also because much more reliability and validity testing is needed to rely on Facebook ratings as a standalone indicator of patient experiences. The research agenda for Facebook as a standalone indicator includes non-investigated topics like stability in hospital-level scores over time, responsiveness or sensitivity to change and potential effects of gaming.

The use of Facebook and social media in healthcare quality assessments is still a rather new and undeveloped research field. In Norway and in other countries,<sup>4 11 15</sup> healthcare providers seem rather slow to establish and use official Facebook pages. Given the quality of official Facebook data indicated in the current study, one interesting line of research would be on barriers and facilitators for the establishment and use of official hospital Facebook sites. This should include hospital employees' and leaders' perspectives and views regarding the establishment of such sites, the choice of allowing or not allowing reviews, and how this kind of information is or might be used as part of quality assessment and improvement work. Another line of research should assess how Facebook ratings and comments might be used in external measurement systems, for instance how Facebook ratings might be integrated with patient experience scores, how uncertainty could be presented, how changes might be captured and presented and how qualitative data from Facebook might be analysed and used. The sentiment analysis approach applied in another study is interesting,<sup>4</sup> but further research should also find ways of using more of the qualitative data than simply the frequency of positive comments.

#### Limitations

One limitation of the current study is the time lag between the survey in 2015 and the extraction of Facebook data, implying that scores from the patient survey and some of the Facebook ratings might be 3 years apart. However, Facebook is cumulative and will at a specific time include a number of ratings and comments from previous time periods. Furthermore, the number of significant correlations and the level of correlations between Facebook ratings and other indicators reduce the probability of major bias in respect of the time lag limitation. Norway has a small population, with only 61 hospitals included in the 2015 survey. The low uptake of official Facebook pages means that analysis for this group is conducted with a low *n* (*n*=16). Thus, the estimated correlation coefficients for both the official sample, the unofficial sample and the total hospital sample have large statistical uncertainty. Another limitation is the rather large number of missing hospitals in the correlation analysis with 30-day mortality and 30-day readmission. Several hospitals are not included in 30-day mortality and 30-day readmission analysis, and others did not have corresponding indicators at the same level as the patient survey. Thus, the generalisability of these correlation analysis is more uncertain than the rest of the analysis.

#### CONCLUSIONS

Facebook ratings at the hospital level were associated with patient-reported indicators, hospital size and 30-day mortality. Qualitative comments from official Facebook sites were much more closely related to hospital evaluation than unofficial sites. More research

is needed on using Facebook ratings as a standalone indicator of patient experiences in national quality measurement, and such ratings should be reported together with research-based patient experience indicators and with explicit criteria for the inclusion of unofficial sites.

**Contributors** OB planned the paper together with HHI, KD and KDS, extracted Facebook data, carried out the statistical analysis and drafted the paper. HHI participated in the planning process, conducted the qualitative analysis, revised the drafts critically and approved the final version. KD participated in the planning process, extracted Facebook data, conducted the qualitative analysis, revised the drafts critically and approved the final version. KDS participated in the planning process, revised the drafts critically and approved the final version.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Oyvind Bjertnaes <http://orcid.org/0000-0002-4261-5455>

## REFERENCES

- Kilaru AS, Meisel ZF, Paciotti B, *et al.* What do patients say about emergency departments in online reviews? A qualitative study. *BMJ Qual Saf* 2016;25:14–24.
- López A, Detz A, Ratanawongsa N, *et al.* What patients say about their doctors online: a qualitative content analysis. *J Gen Intern Med* 2012;27:685–92.
- Ranard BL, Werner RM, Antanavicius T, *et al.* Yelp reviews of hospital care can supplement and inform traditional surveys of the patient experience of care. *Health Aff* 2016;35:697–705.
- Huppertz JW, Otto P. Predicting HCAHPS scores from hospitals' social media Pages: a sentiment analysis. *Health Care Manage Rev* 2018;43:359–67.
- Greaves F, Pape UJ, King D, *et al.* Associations between web-based patient ratings and objective measures of hospital quality. *Arch Intern Med* 2012;172:435–6.
- Greaves F, Pape UJ, King D, *et al.* Associations between Internet-based patient ratings and conventional surveys of patient experience in the English NHS: an observational study. *BMJ Qual Saf* 2012;21:600–5.
- Greaves F, Pape UJ, Lee H, *et al.* Patients' ratings of family physician practices on the Internet: usage and associations with conventional measures of quality in the English National health service. *J Med Internet Res* 2012;14:e146.
- Gao GG, McCullough JS, Agarwal R, *et al.* A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *J Med Internet Res* 2012;14:e38.
- Bardach NS, Asteria-Peñaloza R, Boscardin WJ, *et al.* The relationship between commercial website ratings and traditional Hospital performance measures in the USA. *BMJ Qual Saf* 2013;22:194–202.
- Griffiths A, Leaver MP. Wisdom of patients: predicting the quality of care using aggregated patient feedback. *BMJ Qual Saf* 2018;27:110–8.
- Campbell L, Li Y. Are Facebook user ratings associated with hospital cost, quality and patient satisfaction? A cross-sectional analysis of hospitals in New York state. *BMJ Qual Saf* 2018;27:119–29.
- Perez V, Freedman S. Do crowdsourced Hospital ratings coincide with Hospital compare measures of clinical and nonclinical quality? *Health Serv Res* 2018;53:4491–506.
- Glover M, Khalilzadeh O, Choy G, *et al.* Hospital evaluations by social media: a comparative analysis of Facebook ratings among performance outliers. *J Gen Intern Med* 2015;30:1440–6.
- Timian A, Rupcic S, Kachnowski S, *et al.* Do patients "like" good care? measuring hospital quality via Facebook. *Am J Med Qual* 2013;28:374–82.
- Gaudet Hefeje J, Li Y, Campbell L, *et al.* Nursing home Facebook reviews: who has them, and how do they relate to other measures of quality and experience? *BMJ Qual Saf* 2018;27:130–9.
- Iversen HH, Bjertnaes OA, Skudal KE. Patient evaluation of hospital outcomes: an analysis of open-ended comments from extreme clusters in a national survey. *BMJ Open* 2014;4:e004848.
- Holmboe Bjertnaes ØA. *Inpatients' Experiences with Somatic Hospitals in 2015. National results.* Oslo: Norwegian Institute of Public Health, 2016.
- Bjertnaes O, Skudal KE, Iversen HH, *et al.* The patient-reported incident in hospital instrument (PRIH-I): assessments of data quality, test-retest reliability and hospital-level reliability. *BMJ Qual Saf* 2013;22:743–51.
- Hansen TM, Kristoffersen DT, Tomic O, *et al.* *The Quality Indicator 30-Day Mortality Following Hospital Stay. Results for 2015.* Oslo: Norwegian Institute of Public Health, 2016.
- Hansen TM, Kristoffersen DT, Tomic O. *The Quality Indicator 30-Day Readmission Following Hospital Stay. Results for Hospitals and Municipalities 2015.* Oslo: Norwegian Institute of Public Health, 2016.
- Sjetne IS, Veenstra M, Stavem K. The effect of hospital size and teaching status on patient experiences with hospital care: a multilevel analysis. *Med Care* 2007;45:252–8.
- Holte TO, Bjertnaes Øyvind Andresen, Stavem K. [Is there a relationship between hospital size and patient experience?]. *Tidsskr Nor Laegeforen* 2005;125:1685–8.
- Hamilton LC. *Regression with Graphics. A Second Course in Applied Statistics.* Belmont, California: Duxbury Press, 1992.
- O'Malley AJ, Zaslavsky AM, Elliott MN, *et al.* Case-Mix adjustment of the CAHPS Hospital survey. *Health Serv Res* 2005;40:2162–81.
- Hargraves JL, Wilson IB, Zaslavsky A, *et al.* Adjusting for patient characteristics when analyzing reports from patients about hospital care. *Med Care* 2001;39:635–41.