Regular Article

International Journal of Qualitative Methods
Volume 21: 1–13

DOI: 10.1177/16094069221118993
journals.sagepub.com/home/ijq
SAGE

A Comparison of Leximancer Semi-automated Content Analysis to Manual Content Analysis: A Healthcare Exemplar Using Emotive Transcripts of COVID-19 Hospital Staff Interactive Webcasts

Teyl Engstrom, BMath, BBus, MEpi[1], Jenny Strong, BOccThy, MOccThy, PhD, FOTORA[2], Clair Sullivan, MMBS (Hons), MD, FRACP, FRACP, FACHI, CHIA[3], and Jason D. Pole, BSc (Hons), MSc (Epi), PhD[4]

Abstract

Effective consumer centred healthcare incorporates consumer and clinician perspectives into decision making, in addition to traditional quantitative measures. This information is usually captured in qualitative data that requires manual analysis. Healthcare systems often lack resources to systematically incorporate qualitative feedback into decision making. Semi-automated content analysis tools, such as Leximancer, provide an efficient and objective alternative to time consuming manual content analysis (MCA). Literature on the validity of Leximancer in healthcare is sparse. This study seeks to validate Leximancer against MCA on a broad emotive conversational dataset gathered in a healthcare setting. At the outset of the COVID-19 pandemic, a large Australian hospital and health service conducted interactive webcasts with staff to provide updates and answer questions. A manual thematic analysis and a Leximancer content analysis were conducted independently on 20 webcast transcripts. The findings were compared, along with the time required to the complete each analysis. The Leximancer analysis identified nine concepts, while the manual analysis identified 12 concepts. The Leximancer concepts mapped to five of the concepts identified in the manual analysis, which accounted for 74% of mentions tagged in the text through the manual analysis. Leximancer missed concepts which required an emotional or contextual interpretation. The Leximancer analysis took 21 hours (excluding time to learn the program), compared to 73 hours for the manual analysis. Semi-automated content analysis provides an efficient alternative to manual qualitative data analysis, shifting it from a small-scale research activity to a more routine operational activity, albeit with some limitations. This is critical to be able to utilise at scale the rich narratives from consumers and clinicians in healthcare decision making.

Keywords

qualitative data, content analysis, COVID-19, routinely collected health data

Introduction

Healthcare has traditionally been monitored using quantitative data. This has been effective to date, however this bias to quantitative data did result in the opinions, feelings and perspectives of consumers and clinicians, often recorded as

[1]Centre for Health Services Research, The University of Queensland, Herston, QLD, Australia
[2]Metro North Hospital and Health Services, Brisbane Australia, The University of Queensland, School of Health and Rehabilitation Sciences, Professor Occupational Therapy, University of Southern Queensland
[3]Centre for Health Services Research, The University of Queensland, Herston, QLD, Australia; Metro North Hospital and Health Services, Brisbane Australia
[4]Centre for Health Services Research, The University of Queensland, Herston, QLD, Australia; The University of Toronto, Dalla Lana School of Public Health, Toronto, ON, Canada

Corresponding Author:
Jason D. Pole, Centre for Health Services Research, The University of Queensland, Level 6, 288 Herston Rd, Herston, QLD 4006, Australia.
Email: j.pole@uq.edu.au

rich qualitative data, being excluded from healthcare decision making (Rusinová et al., 2009). As a result, healthcare decision making has traditionally been centred around easy to measure quantitative data such as activity and finance.

This approach is no longer fit for purpose, and the rise of consumer centred healthcare, has forced the collection and analysis of qualitative data in order to capture the consumer and clinician perspective (Pope et al., 2000; Al-Busaidi, 2008). There has been a rapid rise in enterprise applications for healthcare systems to record narrative comments from consumers, with the expectation being that staff at the healthcare provider will manually read all the narrative, synthesise these data manually and respond in a timely manner (Rozenblum et al., 2013). Although collection of qualitative data is certainly occurring at scale, this is a recent development, and analysis has been limited to individual research projects. Few healthcare providers are currently resourced to promptly manually read and code the qualitative data as an operational activity at scale to develop actionable insights (Gleeson et al., 2016).

This rise in the volume of qualitative data being collected and the resultant increased popularity of qualitative data analysis methods, combined with advances in artificial intelligence have resulted in increased utilisation of computational tools in qualitative analysis that automatically or semi-automatically analyse and categorise large bodies of text (Krippendorff, 2019; Nelson, 2020). These semi-automated content analysis (SACA) tools can analyse larger volumes of data more efficiently than manual content analysis (MCA), reduce the human bias and subjectivity inherent in MCA and improve reproducibility (Grimmer & Stewart, 2013). SACA tools generally follow three steps, similar to MCA, the first of which is identifying initial terms which will be used to classify the text. Next, SACA tools define concepts, by creating a thesaurus for each initial term. Finally, the text is classified according to the concepts defined (Nunez-Mir et al., 2016).

Leximancer is one SACA tool which has been increasingly used in research. Leximancer uses machine learning to quickly analyse large bodies of text data and identify the salient points. Leximancer uses takes a completely unsupervised approach, meaning that it does not utilise any pre-defined dictionaries or datasets to understand the meaning of words and which words are similar (Smith & Humphreys, 2006). Rather it infers similarity of meaning and context by detecting patterns of how words travel together throughout each piece of text and provided by the user (Smith & Humphreys, 2006). This approach of being completely grounded in the text means that Leximancer can be used to analyse text in any language (Smith & Humphreys, 2006). Leximancer provides a graphical user interface which allows users to upload a file containing a large amount of text. Leximancer's machine learning algorithms will analyse the data, and identify concepts from words that travel together throughout a body of text and group them into related themes (Smith & Humphreys, 2006). The software also allows users the option to specify concepts of interest to perform a deductive analysis (Smith & Humphreys, 2006). The output can be refined by tweaking other user settings such as combining concepts, removing filler words, etc. (Smith & Humphreys, 2006). The combination of objective approach, speed and customisability offered by Leximancer makes it a strong SACA tool.

There is a paucity of studies examining the validity of SACA, especially in healthcare. We identified only six studies that investigated the validity of Leximancer findings by comparing its output to findings from a MCA on the same dataset (Laura & Jameson, 2020; Harwood et al., 2015; Wilk et al., 2019; Sotiriadou et al., 2014; Penn-Edwards, 2010; Maccarthy & Shan, 2021). Four of those studies analysed non-emotive topics such as risk management (Harwood et al., 2015), brand advocacy (Wilk et al., 2019), sports management (Sotiriadou et al., 2014) and literacy (Penn-Edwards, 2010) and one study analysed the semi-emotive topic of employee engagement (Laura & Jameson, 2020). Of these five studies, one study found that all concepts identified in the manual analysis had an equivalent in the Leximancer analysis and vice-versa (Penn-Edwards, 2010). Four studies found more than half of the MCA themes were mirrored through Leximancer; two of which reported Leximancer found additional concepts or details that added to the MCA (Laura & Jameson, 2020; Sotiriadou et al., 2014). One manual phenomenography study also verified the relationships between concepts presented in Leximancer's visual output (Penn-Edwards, 2010). Two studies highlighted that Leximancer is not able to incorporate the researchers' insight to identify concepts which require more contextual knowledge (Harwood et al., 2015; Wilk et al., 2019) and two suggested that Leximancer and MCA should be used together to complement each other (Wilk et al., 2019; Sotiriadou et al., 2014). The remaining study analysed emotive data, using Trip Advisor reviews from visitors to Australia and New Zealand Army Corps (ANZAC) commemorative sites, which are of significant sentimental value (Maccarthy & Shan, 2021). This study found that Leximancer was not able to identify the underlying emotional concepts that were found in the MCA, such as community and gravitas (Maccarthy & Shan, 2021). Further, the existing validation studies have been limited in the types of data they have been applied to, using transcripts from one-on-one structured interviews (Laura & Jameson, 2020; Harwood et al., 2015; Sotiriadou et al., 2014), written survey responses (Penn-Edwards, 2010) or social media data (Maccarthy & Shan, 2021; Wilk et al., 2019). Language is used differently across mediums – for example formal written text is very different to online chats, and interviews vary from language used conversationally – so it is important that the SACA validation research spans the breadth of how we communicate.

This study will explore whether qualitative analysis using Leximancer produces the same findings as a MCA when applied to a broad conversational dataset which includes both narrative and emotional communication in an interactive webcast healthcare setting. The primary aim of this study is to

Table 1. Example of Verbatim Content From Third VIDCAST.

Vidcast Section	Position of Quote from Vidcast (mm:ss)	Executive Staff Member Quote
General information	0:00	Those who don't recognise me. I'm the chief execitive of the HHS. I see were at – a little over 100 have joined in so far
General information	3:38	So we've had three significant concerns that we've actually been getting from staff, and I will say what they are and then I am going to deal with them in a reverse order. So, number one has been PPE. The second is around workforce and redeployment and … then the third area has been around environmental cleaning
Restating staff question	32:31	I'm approaching the at-risk group, is the redeployment mandatory?
Answering question	32:45	So, the at-risk group, and there is a national standard set by national cabinet last night, absolutely says that we actually have to have a discussion with them. Ultimately, it is for us to actually agree with the individuals around, how do we actually want to manage what is the risk for them?..
Restating staff question	41:58	Is there any planning for the possible increase of traffic of those that would drive to the hospital?
Answering question	42:11	Quite frankly, until it was raised earlier, no. You've got to remember that a huge amount of our actual traffic has already been diverted with the reduction in outpatients, a reduction in planned surgery
Restating staff question	51:29	Has there been consultation with the ED department regarding remote learning? How are we able to fit in with their expectations to keep kids home if we have to work?
Answering question	52:03	So there's a couple of things. Obviously, this week essential workers are exempt from having to, and that actually includes all employees and contractors of the health service, are able to send that. …. But essential services will actually be exempt from that

validate (1) the concepts identified and (2) the relative concept prevalence through Leximancer as compared to the MCA findings (a novel quantitative method to compare findings). Second, we will compare the time required, usability and reproducibility of each method of analysis. This study will add to the literature examining strengths and limitations of SACA, specifically using Leximancer, identifying issues that researchers should be aware of when considering SACA for datasets which include both factual and emotional concepts.

Methods

Setting

A large Australian hospital and health service (HHS) conducted daily interactive webcasts at the outset of the COVID-19 pandemic in March and April 2020. The healthcare service has 21,000 staff and a budget of approximately $3.5 billion AUD. The webcasts served as an information conduit between the executive team and staff in a rapidly changing environment and provided opportunity for staff to ask questions directly of the HHS executives. Uncertainty and fear were high in the early stages of the pandemic and staff sought information in these sessions on all aspects of COVID-19 and the potential impact on the hospital, themselves, and their families.

Data Collection

This study analysed professional transcriptions of 20 interactive one-hour webcasts, digitally recorded sessions conducted from 30-March to 24-April-2020, yielding a corpus of 149,863 words. The conversations usually included two executive staff members sharing audio and video via a digital platform (Microsoft Teams). Each webcast attracted between 300-400 staff members. All HHS staff were invited to join the session and could ask questions via a chat function, either anonymously or not. The executive members would provide a brief topical update about the HHS response to the pandemic at the beginning of the sessions. The exact topics discussed varied by day, depending on who was hosting and what was happening in relation to COVID-19 in the country and HHS on that particular day. However, they usually included topics such as state-wide and HHS COVID-19 case numbers, capacity planning and personal protective equipment stock. They would then read aloud a question asked by a staff member (through the chat function) and provide an answer; this was repeated with the next question until the end of the session. Staff could ask questions about anything that was on their mind about the impact of COVID-19, which often included infection control in the hospital, staffing levels and caring for vulnerable family members. An example of the content from a webcast is Table 1 (reproduced from MCA of manuscripts (Strong et al., 2021)).

Leximancer Analysis

Leximancer version 4.5 was used to analyse the transcripts (Leximancer Pty Ltd, 2018). Leximancer creates concepts from words that often occur together throughout the text (Leximancer Pty Ltd, 2018). Leximancer then groups related concepts together into themes (Leximancer Pty Ltd, 2018). The terms concepts and themes will be used here after.

One researcher (TE) loaded the transcripts, which included the executive team members introduction and update, staff questions read aloud and answered, as text files into Leximancer. An initial 'concept map' was created without altering any settings. The concept map is a visual representation of the concepts and themes identified. Each concept is displayed as a small dot, with the size of the dot representing the degree of connectivity to other concepts – larger dots are more connected. Lines join concepts which are highly connected. Themes are represented by larger bubbles. The bubbles are heat mapped to represent their relative importance/prevalent – warmer colours (red, orange, yellow) represent more prevalent themes and cooler colours (green, blue, purple) represent less prevalent themes.

The concepts identified in the initial concept map were reviewed in conjunction with their associated text segments through the Leximancer query interface. The query interface allows users to search for and read text segments which were tagged with a particular concept. Concepts which offered little Lexical value (e.g. filler words, patterns of speech) in this context were identified and added as stop words, meaning Leximancer ignores those words when building concepts; this was an iterative process. Concepts were merged where they were considered sufficiently similar (e.g. a word and its plural). Compound concepts were created where two or more words formed a commonly used phrase. The concept map was initially observed at a summary level through 'zooming out', by moving Leximancer's 'Theme Size' toggle to 50% which summarises the concepts into approximately 6 themes. Then individual themes and concepts were investigated in more detail by 'zooming in' to a smaller 'Theme Size' to see the concepts in more detail, as described by (Haynes et al., 2019). Note that zooming in and out doesn't change the underlying relationships but provides different visualisations which is useful in the process of understanding the concepts in more detail.

Multiple theme sizes were trialled to arrive at the final concept map. Using Leximancer's interface and query function, a sample of text segments coded with each concept across all transcripts were read until the concept was well understood and no new information emerged. The researcher then summarised the insights into main concepts and themes, regrouping where necessary based on their contextual knowledge and interpretation of the data, in conjunction with the proximity of themes and concepts in the Leximancer concept map, Leximancer counts the number of text segments coded with each concept, known as the 'number of hits'; the number of hits for a specific concept was divided by the total number of hits across all concepts in order to calculate the relative prevalence of each concept.

Manual Analysis

A MCA of this data which was completed independently was used as the gold standard for comparison; this paper has been published with the full methodological detail for the analysis described (Strong et al., 2021). To briefly summarise the manual method, one researcher (JS) immersed herself in the data, reading and re-reading the Word files of each of the 20 transcripts. An inductive approach was used, where categories were derived from the data and A manual thematic analysis was used (Braun & Clarke, 2006). This researcher used open coding, writing down headings as she read through the first transcript. The headings were then grouped into broader themes (Elo & Kyngäs, 2008). For example, the headings of PPE ("At the point where we trigger everybody is wearing PPE inside ED, any clinicians visiting or consulting into the ED will be provided and expected to use the same standard of PPE"); Infection control ("One of the things we are doing is providing or improving capacity and access so that people who do want to shower and change at work, then they do that"); and Workforce issues ("Some of those things we've had time to prepare around: accommodation for frontline workers; accommodation for patients; all our processes of how we do redeployment") were grouped into the theme of Accurate Information. A second researcher independently coded the first transcript, and then the two compared their coding framework. The first researcher then analysed the remaining 19 transcripts using the coding framework. A further three transcripts were double coded by another researcher to ensure inter-coder reliability (O'Connor & Joffe, 2020). The broader research team then read through the findings. The number of times each concept appeared in the text was counted manually and used to calculate the relative prevalence of each concept.

Comparison

After each researcher (TE and JS) independently completed their analysis using Leximancer and MCA respectively, they met to compare findings. They agreed on which concepts from their analysis were similar and which were different. The proportion of concepts which overlapped were calculated, using the relative prevalence percentages from each method. The time require for each method of analysis was estimated in hours and compared.

Results

Leximancer Analysis

The transcripts from the 20 interactive webcasts were analysed in Leximancer, with the default English list of stop words,

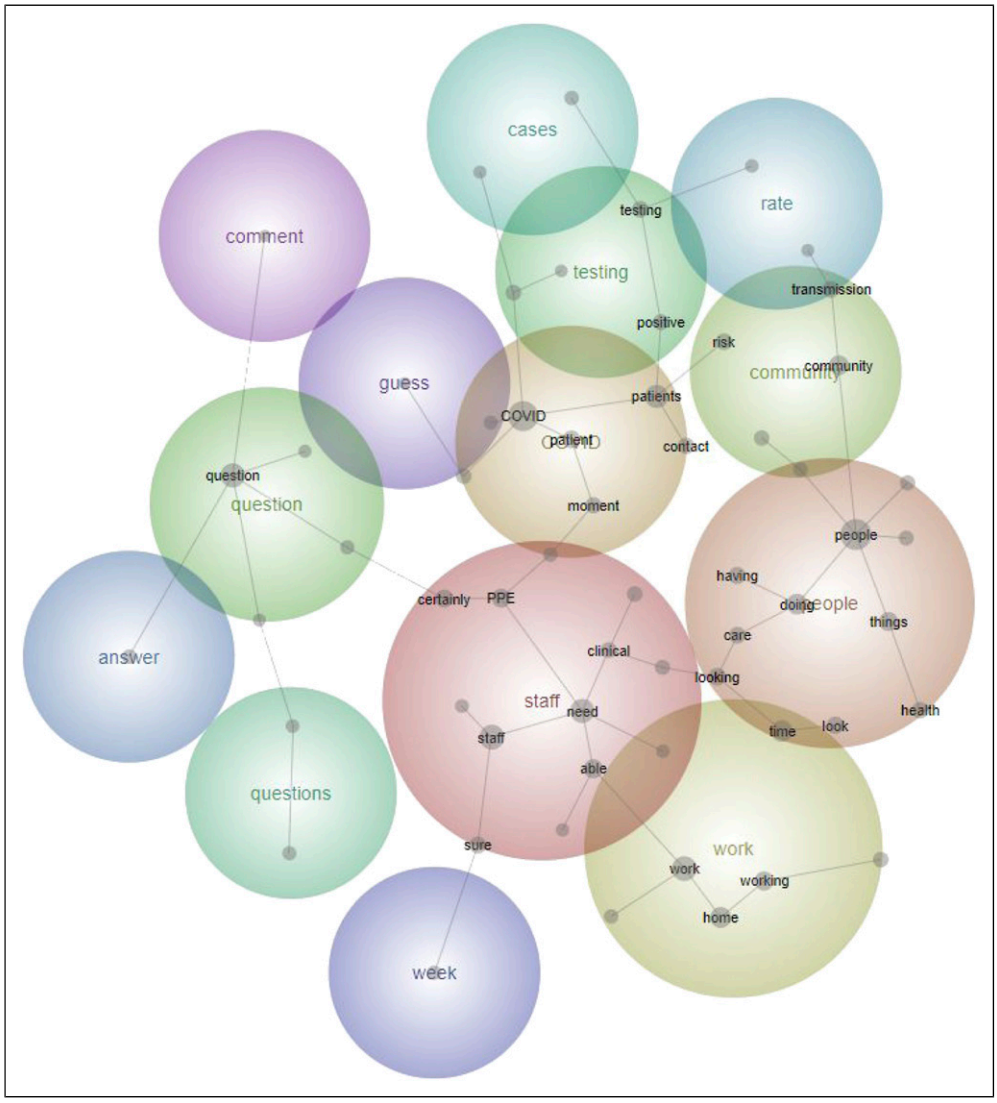

Figure 1. Leximancer initial concept map.

providing the initial default concept map in Figure 1. The concepts identified by Leximancer were reviewed and 33 words with little lexical meaning were added as stop words (Supplementary Materials 1). This included speech patterns of the hosts (e.g. starting sentences with "in fact") and frequently used words in the context of the webcast (e.g. question, answer),. The concepts mask and masks were merged (referring to the Personal Protective Equipment [PPE]). The concept map was re-created based on these settings, shown in Figure 2. The text associated with each concept was reviewed by one researcher and was ultimately formed into nine concepts that were placed into three main themes: (1) Information about COVID-19, (2) Keeping staff safe, and (3) Work policies (Description in Table 2; Illustrative quotes in Table 3).

'Information about COVID-19' was the most prevalent theme with 49% of mentions (2066/4225). This included concepts 'Current HHS situation and future capacity' (23% of mentions), 'COVID-19 epidemiology, transmission, risk and treatment' (13%), 'COVID-19 measures' (9%), and 'COVID-19 testing' (4%). Participants were particularly interested in how COVID-19 cases were being managed at this HHS and the capacity for future cases. Further, questions about the epidemiology and testing of COVID-19, as well as how the disease was being measured in the community and the country were of interest.

The second most prevalent theme concerned 'Work policies' with 29% of mentions (1236/4225). This included concepts 'Staff ways of working' (19% of mentions) and 'Working from home' (11%) Participants asked questions on working from home and how policies were being updated to enable this during the pandemic, especially for staff considered vulnerable or living with vulnerable people. Staff were also interested in how other work practices would change through the pandemic.

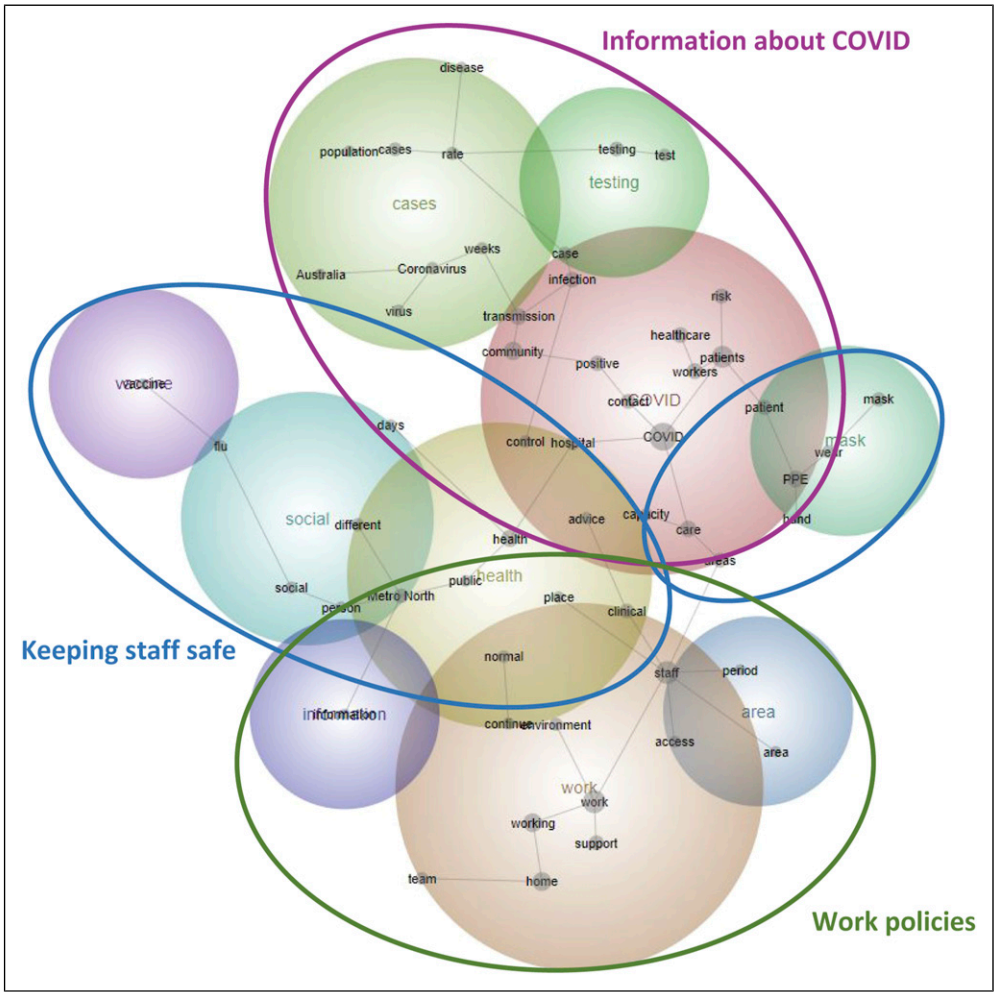

Figure 2. Leximancer final concept map after removing stop words and merging concepts.

The final theme concerned 'Keeping staff safe' with 22% of mentions (923/4225). This included concepts of 'Infection control practices' (15% of mentions), 'Personal protective equipment' (4%) and 'Flu vaccine' (2%). These safety concerns included conversations about the availability of PPE and how and when it should be worn. There was also discussion of new infection control practices that could or should be implemented in the hospitals.

Manual Analysis

The MCA identified 12 concepts which were grouped into three main themes: Accurate information, Reassurance and support, and Innovation. The full results have been published, but the results are briefly summarised here to enable comparisons and contrasts (Strong et al., 2021).

'Accurate information' was the most prevalent theme with 72% of mentions (1216/1683). The concepts were 'COVID-19' (29% of mentions), 'Personal protective equipment' (19%), 'Workforce issues' (10%) 'Infection Control' (9%), and 'Hospital business' (5%). Most conversations related to information on COVID-19, specifically case numbers, care plans, transmission, epidemiology and testing criteria. Information about PPE supply was equally prevalent. Staff were also interested in infection control practices, hospital business and workforce issues such as job security and leave arrangements.

'Reassurance and support' was the second most prevalent theme identified in the MCA with 18% of mentions (306/1683). The concepts were 'Promoting staff well-being' (7% of mentions), 'Adapting to fast changing situations' (5%) 'Managing emotions of fear and anxiety' (4%) and 'Building connectedness' (3%). This theme described how the executive staff were able to manage emotions of fear, anxiety and uncertainty and build connectedness in the way they responded to questions. Staff well-being was promoted, and the executive highlighted how the HHS was adapting to the fast-changing situation.

'Innovation' was the final theme of the MCA with 10% of mentions (161/1683). The included concepts were 'New ways

Table 2. Summary of Concepts From Leximancer Analysis.

Theme	Concept	Concept Detail
Information about COVID-19	Current HHS situation and future capacity	- Current COVID diagnoses - Number of presentations, testetc. - ICU capacity - Proportion of telehealth - Differences to other HHS, states
	COVID-19 epidemiology, transmission, risk and treatment	- Transmission channels - Future immunity - Treatment/care options
	Testing	- Rate of testing - Capacity for more testing - Broadening testing criteria - Need for more antibody, sentinel and serology testing
	COVID-19 measures	- Death rates - Test positivity rates - Proportion of asymptomatic cases - Proportion of cases from overseas
Work policies	Working from home	- Guidelines for working from home - Vulnerable staff policy for working from home - Hours of work when working from home
	Staff ways of working	- Different hours/shifts/reallocation - Capacity - Redeployment to other areas - New wellness initiatives
Keeping staff safe	PPE	- What PPE/masks should be worn - When should PPE/masks be work - Qualities of N95 versus P2 masks
	Infection control practices	- Practices when dealing with patients/entering patient rooms - Use of screens at front desk - Precautions for in-home workers - Alternative accommodation for healthcare workers with who live with vulnerable people - Wearing scrubs at all times inside hospital - Showering before going home
	Flu vaccine	- If and when to get the flu vaccine - Staff reimbursement for flu vaccine

of working' (5% of mentions), 'Communication' (2%) and 'HHS digital agency' (2%). This theme focused on how the Digital Agency at this HHS was introducing new means of communication, apps and virtual models of care to enable working in this changing environment.

Comparison

A visual comparison of the concepts identified in the two analysis methods is displayed in Figure 3. Of the 12 concepts found in the MCA, five had corresponding concepts in Leximancer; these five concepts were the most prevalent concepts derived from the MCA, representing 74% of all mentions tagged in the text. Leximancer analysis identified one additional concept that wasn't explicitly identified in the MCA, flu vaccines, which was the least prevalent theme of all Leximancer concepts, accounting for only 2% of all mentions tagged in the text by Leximancer. However, the MCA coded flu vaccines into multiple different concepts – staff well-being, infection control and hospital business depending on the context.

The manual analysis identified seven concepts that the Leximancer analysis did not. Only one of the five concepts identified in the Accurate Information theme was not found in the Leximancer analysis, being 'Hospital Business', which represents 5% of all mentions tagged manually. The Leximancer analysis did not identify three of the concepts from the 'Reassurance and support' theme, with these being 'Managing emotions of fear and anxiety', 'Building connectedness' and 'Adapting to fast changing situations', which account for 12% of all mentions tagged manually.

The 'Innovation' theme concepts, 'HHS Digital Agency', 'New ways of working' and 'Communication' were identified by the MCA (9% of all mentions tagged) but not through Leximancer.

The Leximancer analysis took approximately 42 hours to complete. This included the time for one researcher (TE), proficient with using data and analytical software programs, to

Table 3. Illustrative Quotes for Concepts From Leximancer Analysis.

Theme	Concept	Illustrative Quotes
Information about COVID-19	Current HHS situation and future capacity	"To be sure though, we were expecting to be seeing more cases in hospital by now and we haven't and we're not really sure what the reasons for that are. So it's just bought us a bit of time."
		"How many community-acquired cases in the region?"
		"I do think it's a good idea around looking at where we do have closed patient capacity and the opportunities around off-site training."
	COVID-19 epidemiology, transmission, risk and treatment	"So my understanding is that you're likely to be most infectious after that point and less so before one is symptomatic. So yes, there's asymptomatic transmission."
		"So sooner or later we're actually going to start getting community transmission and start building a herd immunity"
		"What is your opinion on the use of convalescent serum to treat acute infection?"
	Testing	"Can you discuss the accuracy of serology testing for COVID-19 in relation to antibody response?"
		"Australia, this little island country, with some 26 million people, has one of the highest per capita testing rates in the world."
		"We've broadened criteria, but fever clinic numbers remain low; do you think people might be scared to present due to social stigma and financial consequences?"
	COVID-19 measures	"I think in China it was 8%, now it's 4%. Italy has a staggering death rate, I read recently, of 15%."
		"Yes, the nasopharyngeal swab has a sensitivity about 20% up to 80%. So there is a false negative rate I want to say something about the curve that people have observed that might be flattening in Australia, in Queensland."
		"So until about 4 weeks ago, the rate of positivity was 0.5%. So of all the 20,000-odd tests that were done, 0.5% were positive."
Work policies	Working from home	"So if you feel that working at home is going to help with that situation have a conversation with your manager."
		"If we live with parents with comorbidities, are we considered to be in the vulnerable category?"
		"Are the hours of work when working from home negotiable outside the awards?"
	Staff ways of working	"There is a document that goes into about redeployment and working with COVID-19 patients"
		"Next week we're launching the well-being framework to actually support that"
		"What should we be telling our casual staff that are not getting any shifts?"
Keeping staff safe	PPE	"Would you recommend for nurses to wear non-disposable gown to protect our uniforms from touching potentially contaminated surfaces and patient linens?"
		"We recommend P2 because of this theoretical risk of aerosolisation and infection as a result of it, but I believe that's adequate protection for routine cares."
		"What are the PPE recommendations for those attending child birth for the purpose of neo natal resuscitation for asymptotic?"
	Infection control practices	"When doffing our PPE in a single room, is it okay to doff everything except the mask and goggles in the room?"
		"We are making sure that there are shower facilities for all of those staff to shower on arrival and departure"
		"You should wear the same level of PPE when you're doing a home visit as you would if you were providing care on a ward."
	Flu vaccine	"I think if you get the vaccine now, you're pretty much covered through winter, so I wouldn't delay getting your flu shot."
		"Why is pre-operative staff not considered frontline staff for the flu vaccine?"
		"I will come back to you on claiming reimbursement for my flu vac."

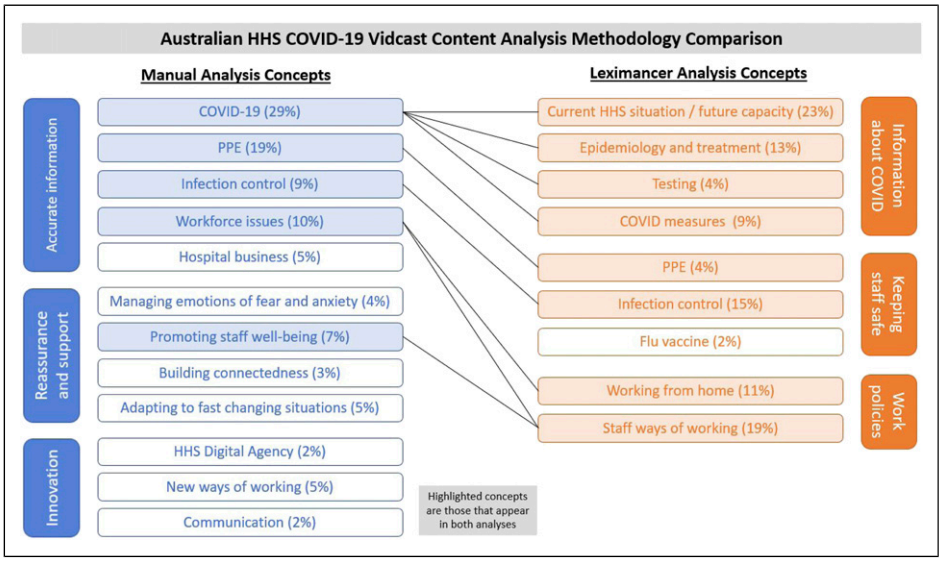

Figure 3. Comparison of concepts from manual content analysis and Leximancer content analysis.

Table 4. Strengths and Limitations of Manual and Leximancer Content Analysis.

Analysis Method	Strengths	Limitations
Manual content analysis	- Incorporates researcher contextual knowledge to identify concepts	- More time consuming, especially with larger datasets
	- Emotional concepts can be captured	- More likely to reflect human bias
Leximancer content analysis	- Faster to complete, especially for larger datasets	- Does not identify emotive concepts
		- Does not identify contextual concepts
	- Reduces human bias	- Can include some human bias in interpreting results
	- Results are reproduceable	- Less likely to detect problems for a small subgroup of people

learn how to use Leximancer. The learning and analysis occurred concurrently; however, it is estimated that to complete this analysis with existing knowledge of how to use Leximancer would take approximately 21 hours. The MCA analysis took approximately 73 hours accounting for the time from both researchers with experience in conducting content analysis. This represents a time savings for the SACA of more than 42% over the MCA, including time to learn Leximancer; or a 71% reduction in time required for the MCA if the user has experience with Leximancer.

The strengths and limitations of each method are summarised in Table 4.

Discussion

Qualitative data analysis is rapidly evolving from a research method to an operational activity for healthcare providers who are serious about including consumer and clinician perspective in their data-driven decision making (Lee et al., 2018). How to shift qualitative data analysis from a bespoke research activity to an operational activity which is achievable within constrained healthcare resourcing profiles remains unclear (Sanders et al., 2020). This paper pioneers this concept using semi-automated methods to analyse complex qualitative data from a large Australian healthcare service.

This study of hospital staff COVID-19 webcast conversation transcripts found Leximancer identified most of the concepts that were found through a (MCA) with only 29% of the time (29%–58% depending on Leximancer experience). However, the semi-automated analysis missed concepts which required an emotional or contextual interpretation. The concepts that Leximancer did identify accounted for 74% of mentions tagged manually throughout the text. Leximancer identified one additional concept, flu vaccines, which was not identified as a separate concept in the MCA but was incorporated into multiple other concepts. Leximancer's concept map provided a high-level overview to orient the user to the main topics which were being discussed in the interactive webcasts, as well as how they were related to each other. It also provided an interactive interface which made delving into the underlying text easy to do. As with MCA, the role of the researcher remains important in interpreting the Leximancer output. This study adds support to the ability for Leximancer

to be used for qualitative analysis, while highlighting the limitations that researchers should be aware of when using it.

Validity

Validity is a key criterion for evaluation of SACA (Grimmer & Stewart, 2013; Smith & Humphreys, 2006; Müller-Hansen et al., 2020); our study contributes to the shared knowledge regarding the validity and limitations of SACA tools. To our knowledge this is the first study to compare Leximancer output to a MCA which identified emotive themes from healthcare conversational data such as 'Building connectedness', and 'Managing emotions of fear and anxiety'. An example of text that was identified as 'Building connectedness' is a staff member asking: "Do you have family or children at home, if so what measures or precautions are you taking to protect them, after seeing COVID patients?" and the executive team host responding: "Really good question and something that has kept me awake a bit while we were planning. So yes, I do have a little boy who is 15 months old, and I worry about this.". Leximancer extracted the concepts 'patients', 'COVID' and 'home' from this exchange. One of the text passages labelled in the MCA as 'Managing fear and emotions' is the following: "This is a time we've got to support each other's resilience. I look at this and I think about how many stressors we have outside of work. Friends, family, parents. Our fear for them. And all of those play out and come to our work environment. We are more than a sum of just treating patients. We all have complex circumstances outside of work and so I ask you all to be generous with each other. Be kind.". In this case Leximancer identified the concepts 'support', 'patients' and 'work'.

Meaning is derived from more than just individual words; the way in which language is arranged can elicit different interpretations and feelings in the reader. These MCA concepts are based on the researcher's' interpretation of how the executives were trying to make their audience feel through how they answered the questions, rather than through words in the text. Hence it is not unexpected that analysing the same data using Leximancer – a tool grounded in the text – did not identify these emotive concepts. This is a similar finding to that of MacCarthy, who concluded that Leximancer was not able to identify emotional concepts that were found through MCA in emotive online reviews of commemorative war sites (Maccarthy & Shan, 2021). Recent research has explored how the interpretation of emotions can differ across cultures and contexts, making it even more difficult for SACA and artificial intelligence technologies to identify such concepts (Van Berkel et al., 2020). Previous studies of social media data have also reported that Leximancer was not able to capture tone of voice (Wilk et al., 2019). This finding suggests that Leximancer alone should not be used in analyses where researchers are trying to identify emotive concepts, and that researchers should be aware of this limitation in the case of trying to identify both information based and emotive concepts.

Concepts identified in the manual analysis, such as 'hospital business' and 'innovation' were not discovered through Leximancer. The 'hospital business' theme is described in the MCA as including logistics, equipment, workflows and patient care. These areas were grouped together into a single concept by the researcher based on her contextual knowledge of what constitutes hospital business; this accounted for 5% of concepts coded in the text. The 'innovation' theme included a number of activities the HHS had begun in order to continue operations in the delivery of healthcare in the pandemic; it accounted for 9%. While Leximancer goes beyond word counting and creates a thesaurus for each concept of words that occur together throughout the text and are used in similar ways, it did not bring these terms together to form a concept of sufficient size to be included in the concept map. Given the contextual knowledge required to identify these concepts and their relatively low frequency, it is not unexpected that these concepts were not matched by Leximancer. Leximancer is grounded in the text provided and hence it is unable to incorporate the contextual knowledge of the researchers; this has been reported in other studies comparing Leximancer to MCA (Harwood et al., 2015; Wilk et al., 2019).

We have demonstrated a novel method of quantifying the validity of SACA findings relative to a gold standard MCA. The six SACA validation studies we identified all provided a narrative description of the differences compared to MCA; three of them also gave a simple count of the number of concepts or themes which were and were not matched in Leximancer (Laura & Jameson, 2020; Maccarthy & Shan, 2021; Penn-Edwards, 2010). In addition to this comparison of the number of concepts, we calculated the relative prevalence of each concept mentions tagged throughout the text and used the proportion of those mentions which were matched by Leximancer as the primary measure to validate the completeness of the SACA findings. This added level of detail is crucial to consider, as not all concepts are of equal prevalence or importance. This method of comparison has provided stakeholders comfort with SACA using Leximancer. We believe this measure is useful in addition to a qualitative narrative summary and would encourage researchers to adopt and expand on this method in future SACA validation studies.

Time Required, Usability, Subjectivity, and Reproducibility

Qualitative data analysis is a time-consuming task, and a well-documented benefit of SACA is that can incorporate large volumes of text much more quickly than MCA (Nunez-Mir et al., 2016; Penn-Edwards, 2010; Smith & Humphreys, 2006). This study confirmed this, with the Leximancer analysis taking 71% less time than required to complete the MCA (excluding the time required to learn how to use

Leximancer). The Leximancer program reads all the text provided, extracting only the portions tagged with each concept into the interface for the researcher to read. It also removes the need for text to be double-coded, which is best practice in MCA (O'Connor & Joffe, 2020).

The Leximancer software is easy to use, the interface is straight forward and producing initial results can be achieved in a matter of minutes. However, a key step to refining the results is for the researcher to adjust the settings, such as adding stop words and merging concepts. Leximancer's concept map and the ability to adjust the level of detail provides a useful overview of the results. The query window makes it easy for researchers to read text segments coded with each concept. However, the researcher's role in interpreting the results is key. There are inherent similarities in what SACA is doing to the MCA methodology which focuses on pattern recognition (Haynes et al., 2019; Janasik et al., 2009; Yu et al., 2011). Hence it is advisable that those using Leximancer have some familiarity with MCA methods such that they understand what Leximancer is doing and how to interpret the output. The utility of SACA output is likely to be increased if the user has more knowledge of the topic and context of the text (Nunez-Mir et al., 2016).

SACA offers a more objective alternative to MCA, by reducing the human bias involved (Smith & Humphreys, 2006). This is important in order to be able to analyse qualitative data at scale, such as in the healthcare context where thousands of qualitative patient responses and physician notes are captured each day. If this was to be done using MCA many people would be required to analyse the data, each introducing their own biases, and reducing consistency of the analysis. Conscious or unconscious human bias may also feed into the analysis by highlighting topics which are known to be important in the health service, or their key performance indicators. A limitation of performing SACA at scale is that issues raised by a small number of people are unlikely to be highlighted, which potentially increases theris of not including feedback from underrepresented populations.

Another important criterion in evaluating SACA is reproducibility (Smith & Humphreys, 2006; Müller-Hansen et al., 2020); this has long been a key criterion in traditional quantitative scientific research, and has been considered to have been a weakness of qualitative research (Noble & Smith, 2015). Use of SACA tools, and the ability to publish the code and settings that were used in the analysis helps to improve reproducibility in this field (Müller-Hansen et al., 2020; Thompson et al., 2014). Leximancer facilitates this by allowing stop words and other settings to be downloaded and saved for reuse. Further SACA methods minimise the impact of human bias in interpreting a text compared to MCA (Nunez-Mir et al., 2016; Penn-Edwards, 2010; Smith & Humphreys, 2006), aiding reproducibility.

SACA offers a more efficient, reproduceable and objective method of pattern recognition, albeit with some limitations around identifying emotive and contextual concepts that researchers should be aware of when applying it.

Study limitations

This study has some limitations which should be acknowledged. In any comparison of qualitative analyses, it is highly unlikely that two independent researchers would arrive at the same results, even if they were using the same method (Harwood et al., 2015), so some differences in the results are to be expected. The researchers involved in this study both participated in the live interactive webcasts sessions as part of their employment, and so may have had some preconceived ideas of their content. However, they did not discuss their findings on this data until both analyses were finalised to maintain independence. This study aimed to validate Leximancer results through comparison with MCA; however some research warns against using manual analysis as the gold standard (Song et al., 2020). The MCA being utilised in this study has followed a well-established process, including double-coding, and has undergone peer review as part of the publication process (Strong et al., 2021), however it may still reflect some author bias. Some, albeit less explicit, human bias is likely to be reflected in the Leximancer analysis as human input and interaction is required in defining the 'stop words' and interpreting the output (Nunez-Mir et al., 2016).

Conclusion

There is a growing body of literature regarding the role of SACA tools in qualitative research. This study found that a SACA tool, Leximancer, had 74% concordance with a MCA of the same data. However, this SACA tool did not identify emotive concepts and those that required more contextual knowledge. Some studies in this area have noted that the ability of AI in qualitative analysis is exaggerated (Maccarthy & Shan, 2021) or that it should only be used to complement MCA (Harwood et al., 2015; Laura & Jameson, 2020; Singleton et al., 2018). While SACA has some limitations, we believe in the value of SACA tools like Leximancer which allow us to analyse more data more frequently; the growth in data far exceeds the available time and resources for MCA, so in most cases the alternative to SACA is not MCA, it is doing nothing with the data. SACA is important in shifting qualitative data analysis from a small-scale laborious and highly skilled research activity to an operational activity for healthcare which can be delivered at scale. This is critical to be able to include rich narratives from consumers and clinicians in healthcare decision making.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Teyl Engstrom https://orcid.org/0000-0002-1778-5006

Supplemental Material

Supplemental material for this article is available online.

References

Al-Busaidi, Z. Q. (2008). Qualitative research and its uses in health care. *Sultan Qaboos Univ Med J*, *8*(1), 11–19.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Gleeson, H., Calderon, A., Swami, V., Deighton, J., Wolpert, M., & Edbrooke-Childs, J. (2016). Systematic review of approaches to using patient experience data for quality improvement in healthcare settings. *BMJ Open*, *6*(8), e011907. https:doi.org/10.1136/bmjopen-2016-011907

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Polit Anal*, *21*(3), 267–297. https://doi.org/10.1093/pan/mps028

Harwood, I., Gapp, R., & Stewart, H., Cross-Check for Completeness (2015). Exploring a novel use of leximancer in a grounded theory study. *Qualitative Report*, *20*(7), 1029.

Haynes, E., Garside, R., Green, J., Kelly, M. P., Thomas, J., & Guell, C. (2019). Semiautomated text analytics for qualitative data synthesis. *Research Synthesis Methods*, *10*(3), 452–464. https://doi.org/10.1002/jrsm.1361

Janasik, N., Honkela, T., & Bruun, H. (2009). Text mining in qualitative research: Application of an unsupervised learning method. *Organizational Research Methods*, *12*(3), 436–460. https://doi.org/10.1177/1094428108317202

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology*. Sage.

Laura, L. L., & Jameson, H. (2020). Enhancing trustworthiness of qualitative findings: Using leximancer for qualitative data analysis triangulation. *Qualitative Report*, *25*(3), 604–614.

Lee, R., Baeza, J. I., & Fulop, N. J. (2018). The use of patient feedback by hospital boards of directors: A qualitative study of two NHS hospitals in england. *BMJ Quality & Safety*, *27*(2), 103–109. https://doi.org/10.1136/bmjqs-2016-006312

Leximancer Pty Ltd (2018) Leximancer user guide (*45*).*Release*

Maccarthy, M., & Shan, H. (2021). Machine infelicity in a poignant visitor setting: Comparing human and AI's ability to analyze discourse. *Current Issues in Tourism*, *25*(8), 1289–1306. https://www.tandfonline.com/doi/full/10.1080/13683500.2021.1915252

Müller-Hansen, F., Callaghan, M. W., & Minx, J. C. (2020). Text as big data: Develop codes of practice for rigorous computational text analysis in energy social science. *Energy Research & Social Science*, *70*(2020), 101691.

Nelson, L. K. (2020). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, *49*(1), 3–42. https://doi.org/10.1177/0049124117729703

Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence Based Nursing*, *18*(2), 34–35. https://doi.org/10.1136/eb-2015-102054

Nunez-Mir, G. C., Iannone, B. V., Pijanowski, B. C., & Kong, N. (2016). Automated content analysis: Addressing the big literature challenge in ecology and evolution. *Methods in Ecology and Evolution*, *7*(11), 1262–1272. https://doi.org/10.1111/2041-210X.12602

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, *19*, 1609406919899220. https://doi.org/10.1177/1609406919899220

Penn-Edwards, S. (2010). Computer aided phenomenography: The role of leximancer computer software in phenomenographic investigation. *The Qualitative Report*, *15*(2), 252–267. https://doi.org/10.46743/2160-3715/2010.1150

Pope, C., Ziebland, S., & Mays, N. (2000). Qualitative research in health care. Analysing qualitative data. *Bmj: British Medical Journal*, *320*(7227), 114–116. https://doi.org/10.1136/bmj.320.7227.114

Rozenblum, R., Lisby, M., Hockey, P. M., Levtzion-Korach, O., Salzberg, C. A., Efrati, N., Lipsitz, S., & Bates, D. W. (2013). The patient satisfaction chasm: The gap between hospital management and frontline clinicians. *BMJ Quality & Safety*, *22*(3), 242–250. https://doi.org/10.1136/bmjqs-2012-001045

Rusinová, K., Pochard, F., Kentish-Barnes, N., & Chaize, M. (2009). Qualitative research: Adding drive and dimension to clinical research. *Critical Care Medicine*, *37*(1)–S140–S146. https://doi.org/10.1097/CCM.0b013e31819207e7

Sanders, C, Nahar, P, Small, N, Hodgson, D., Ong, B. N., Dehghan, A., Sharp, C. A., Dixon, W. G., Lewis, S., Kontopantelis, E., Daker-White, G., Bower, P., Davies, L., Kayesh, H., Spencer, R., McAvoy, A., Boaden, R., Lovell, K., Ainsworth, J., & Nenadic, G. (2020). Digital methods to enhance the usefulness of patient experience data in services for long-term conditions: The DEPEND mixed-methods study. *Health Services and Delivery Research*, *8*(28), 1–128. https://doi.org/10.3310/hsdr08280

Singleton, J. A., Lau, E. T. L., & Nissen, L. M. (2018). Waiter, there is a drug in my soup – using Leximancer® to explore antecedents to pro-environmental behaviours in the hospital pharmacy workplace. *International Journal of Pharmacy Practice*, *26*(4), 341–350. https://doi.org/10.1111/ijpp.12395

Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods*, *38*(2), 262–279. https://doi.org/10.3758/bf03192778

Song, H., Tolochko, P., Eberl, J.-M., Eisele, O., Greussing, E., Heidenreich, T., Lind, F., Galyga, S., & Boomgaarden, H. G. (2020). Validations we trust? The impact of imperfect human annotations as a gold standard on the quality of validation of

automated content analysis. *Political Communication*, *37*(4), 550–572. https://doi.org/10.1080/10584609.2020.1723752

Sotiriadou, P., Brouwers, J., & Le, T.-A. (2014). Choosing a qualitative data analysis tool: A comparison of NVivo and leximancer. *Annals of leisure research*, *17*(2), 218–234. https://doi.org/10.1080/11745398.2014.902292

Strong, J., Drummond, S., Hanson, J., Pole, J. D., Engstrom, T., Copeland, K., Lipman, B., & Sullivan, C. (2021). Outcomes of rapid digital transformation of large-scale communications during the COVID-19 pandemic. *Australian Health Review*, *45*(6), 696–703. https://doi.org/10.1071/AH21125

Thompson, J., Davis, J., & Mazerolle, L. (2014). A systematic method for search term selection in systematic reviews. *Research Synthesis Methods*, *5*(2), 87–97. https://doi.org/10.1002/jrsm.1096

Van Berkel, N., Tag, B., Goncalves, J., & Hosio, S. (2020). Human-centred artificial intelligence: A contextual morality perspective. *Behaviour & Information Technology*, *41*(3), 502–518. https://doi.org/10.1080/0144929x.2020.1818828

Wilk, V., Soutar, G. N., & Harrigan, P. (2019). Tackling social media data analysis: Comparing and contrasting QSR NVivo and Leximancer. *Qualitative Market Research*, *22*(2), 94–113. https://doi.org/10.1108/qmr-01-2017-0021

Yu, C. H., Jannasch-Pennell, A., & DiGangi, S. (2011). Compatibility between text mining and qualitative research in the perspectives of grounded theory, content analysis, and reliability. *Qualitative Report*, *16*(3), 730–744. https://doi.org/10.46743/2160-3715/2011.1085