

UNIVERSITEIT TWENTE.

M. Sc. Thesis

**Comparing textual characteristics of Covid-19 patients after hospitalization with
high and low PTSD symptomatology**

Anne Kieft S2466880

Faculty Behavioural Management and Social Sciences, University of Twente

Supervisors: Anneke Sools and Sandra Oude Wesselink

Date: 08-04-2022

Abstract

The aim of this study is to identify differences in characteristics in texts of open-ended questions from COVID-19 patients after hospitalization to indicate PTSD symptoms. This was done by comparing the textual data of two groups, namely a group scoring high on PTSD symptoms and a group scoring low on PTSD symptoms. In this study text-mining techniques were used to examine the different characteristics of word count, topics, and sentiment. Word count was analyzed by counting words in WORD and analyzing them in SPSS. Analyzing the topics was done in two ways: word frequency and topic modelling. With word frequency, frequently used words were compared between the groups. Topic modelling was done by using machine learning to identify different topics in the text. The program of ORANGE was used to analyze texts on topics and sentiment. In conclusion, word count and sentiment do not indicate PTSD symptoms. For word frequency, it can be concluded that indicators for PTSD symptoms are the usage of words regarding procedures, not physically doing better, loneliness and sickness. Especially, the personal surrounding of the patients regarding the absence of visitors and the feeling of loneliness could indicate PTSD symptomatology. In addition, frequent use of superlatives in a negative context seem to indicate PTSD symptoms. In contrary, discharged COVID-19 patients writing about fast recovery and doing physically better resulting in going back to work seem not to indicate PTSD symptoms. For topic modelling, in the low scoring group three topics were found, and in the high scoring group four topics. It seems that writing about repetitive procedures and long-term persuasive implications; indispensable and important professions during and after COVID-19 infection; Impactful, confrontational, unexpected experiences and outcomes; and experiencing loneliness but receiving good care during hospitalization indicate PTSD. The findings of this study could be used by the hospital to predict and prevent PTSD symptoms in COVID-19 patients after hospitalization.

Index

Introduction	3
Methodology	9
<i>Design</i>	9
<i>Participants</i>	9
<i>Measurements</i>	12
<i>Analyses</i>	12
Results	19
<i>Word count</i>	19
<i>Topic modeling</i>	20
<i>Sentiment</i>	32
Discussion	35
Literature	44

Introduction

The coronavirus disease (COVID-19) was first reported in Wuhan, China and spread across the world rapidly. The World Health Organization declared COVID-19 as a world pandemic a few months later (WHO, March 12, 2019). COVID-19 could cause respiratory distress, which in most severe cases could lead to death (Fu et al., 2020). People severely affected by the COVID-19 disease were hospitalized due to the symptoms. In the beginning of the pandemic the amount of hospitalized people arose, leading at the peak to 3284 hospital admissions in the Netherlands on the 28th of March, of which 992 people were confined in the Intensive Care Unit (Ritchie et al., 2020).

Mental health disturbances due to COVID-19 have been reported by health care workers and the general population (Kang et al., 2020; Qui et al., 2020). A global event that causes harm and death, has lasting impacts on mental and physical health of people (Goldmann, & Galea, 2014). Due to hospitalization and the treatments in isolation COVID-19 patients might experience fear, anger, boredom, and loneliness (Xiang et al., 2020). In addition, along with COVID-19 symptoms the patient may experience a series of stressors, such as physical isolation, death of family members and insecurities about the illness, which causes anxiety, depression, and insomnia (Liu et al., 2021; Zhou et al., 2020). In addition, according to Zhang and colleagues (2020) persons who experienced COVID-19 infection reported higher prevalence of depression symptoms in comparison to non-infected participants in quarantine. They are also more likely to express depressed moods and somatic symptoms. Some people even experience persistent chronic mental health problems due to the traumatic experience, such as anxiety disorders, depression, and PTSD (Adams, & Boscarino, 2006).

Many COVID-19 patients have a long way to go in recovering from their hospital stay, because they need to adjust to the physical and psychological implications. The

experience has a negative effect on the person's psychological wellbeing (Tingey et al., 2020). According to Xiao and colleagues (2020), a life-threatening physical illness can lead to symptoms associated with post-traumatic stress disorder (PTSD) after recovery. Prevalence rates for PTSD symptoms for a life-threatening situation is respectively 14-59% in an intensive care unit (Tedstone, & Tarrier, 2003, as cited in Wu et al., 2005). Post-traumatic stress disorder is a psychological disorder that is caused by a traumatic event (Schaap et al., 2000). Symptoms of PTSD are re-experiencing the traumatic event, either by flashbacks, nightmares, disturbing thoughts related to the traumatic event and high arousal in similar events. That COVID-19 patients experience PTSD symptoms was confirmed by a study of Bo and colleagues (2021), in which 96,2% clinically stable COVID-19 patients reported significant post-traumatic stress disorder symptoms. Besides, Chamberlain and colleagues (2021) identified significant elevated rates of PTSD symptoms in COVID-19 patients that required hospital admission with or without mechanical ventilation, in which intrusive images were the most prominent symptom. In comparison, the SARS-Cov-1 outbreak in 2003 was less trembling than the current outbreak, but still had indirect effects that resulted in psychiatric complications (Vindegaard & Benros, 2020). Back then the occurrence of PTSD symptoms was between 28-31% (Hawryluck et al., 2004). Therefore, it is expected that the prevalence rates of PTSD will be higher with the current outbreak, which leads to a major concern for general healthcare (Vindegaard & Benros, 2020).

PTSD is caused by a traumatic event and has many consequences. A traumatic event could increase the chances of chronic mental diseases, such as major depression, substance abuse, panic disorder and generalized anxiety disorder (Kessler et al., 1995). In addition, trauma could cause physical complications, such as chronic pain syndromes and particularly hypertension (Zatzick et al., 1997). The symptoms of PTSD have been described earlier and could have negative consequences on someone's life such as lower quality of life and

influences individuals social and occupational functioning (Monson et al., 2017). PTSD could cause avoidance of stimuli that relate to the trauma or numbness in responsivity (Schaap et al., 2000). Avoidance reduces the possibilities to diminish and extinguish the fear (Yehuda, 2002). Next to that, PTSD patients suffer from sleep disturbances, difficulty concentrating, distressing dreams, irritability, and headaches (Otis et al., 2011; McNally et al., 2015). Undiagnosed PTSD could also lead to a higher suicide risk (Stanley et al., 2017). Next to the consequences of PTSD, the symptoms of PTSD could also be very disturbing for the patient and hinder the daily functioning of the person.

Considering the negative impacts of PTSD, early identification of PTSD is critical to initiate an intervention quickly (He et al., 2017). Early identification is important to prevent personal suffering and to initiate an appropriate intervention for COVID-19 patients (Bo et al., 2021). According to the guidelines, the first step in diagnosing PTSD is using a questionnaire to screen the symptoms of PTSD. Multiple questionnaires about feelings and emotions need to be filled in to reach a diagnosis (National center for PTSD, n.d.). If the questionnaire is positive, the psychologist will conduct an anamnesis with the patient in order to subtract more information and give a proper diagnosis (GGZ standaarden, 2021). However, according to an article of Formplus (2018) survey response fatigue is a recurring problem for patients when taking surveys. According to Porter and colleagues (2004), some people feel 'over surveyed', because of the number of questionnaires they receive. They found a negative correlation between previous survey contacts and the participation in a later survey. Next to that, the respondent could be overwhelmed by the number of questions, which could cause demotivation.

A possible other way for indicating PTSD is examining narratives of patients. Language is the most common and reliable way for people to translate their internal thoughts and emotions into a form (Tausczik & Pennebaker, 2010). Therefore, a lot of valuable data

can be found in texts. Words and expressions in text could detect mental health disorders early (Franklin & Thompson, 2005). By examining the narratives of people with high and low PTSD symptomatology differences in narrative characteristics can be found that could be a predictor in identifying PTSD symptomatology. In addition, letting the patients write texts could counteract the feeling of over surveying and prevent demotivation in the diagnosis process.

The characteristics of the narratives can be examined by using text-mining. The text-mining technique uses unstructured text documents to codify this data in structured data and with this data clinical information can be identified (Trusko et al., 2010). "Text-mining seeks to extract useful information from document collection through the identification and exploration of patterns among unstructured textual data" (He, Veldkamp, & De Vries, 2012). An example of a text-mining study that examined characteristics in narratives is the study of Howes and colleagues (2014). In which they studied the topics and sentiment of the texts written by patients and predicted the depressive symptom severity.

There are different text-mining methods to identify characteristics in texts. One of the more advanced methods uses text classification. Text classification is the process of assigning categories to unstructured text (Monkeylearn, n.d.-a). One of the methods of text classification is sentiment analysis, which is the classifications of sentiment from narratives (Hussein, 2018). Another text classification method is topic analysis, which helps to understand the topics of the text, what is talked about. Next to that, there are more descriptive methods, such as word count, which is the number of words used.

In this study qualitative data of discharged COVID-19 hospital patients will be examined and compared to see the differences in characteristics in narratives between patients with scoring high on PTSD symptomatology and scoring low on PTSD symptomatology. The goal is to identify differences in characteristics in written text between

the two groups to indicate PTSD symptoms. Those characteristics could help indicate the risk of high PTSD symptomatology after hospitalization in the future. The research question is: To what extent are there differences in characteristics in narratives of hospitalized COVID-19 patients with high and low PTSD symptomatology?

Sub questions are: To what extent is there a difference in topics between hospitalized COVID-19 patients scoring high and low on PTSD symptomatology? To what extent is there a difference in word count between COVID-19 patients scoring high and low PTSD symptomatology? To what extent is there a difference in sentiment COVID-19 patients with high and low on PTSD symptomatology?

As described above, the characteristics that will be examined are word count, topic modeling and sentiment analysis. This study will be an explorative study investigating those characteristics. According to the literature investigating these three characteristics are relevant. It is relevant to examine sentiment analysis, because according to a study of Sawalha and colleagues (2022) individuals with PTSD use more frequently neutral and negative words. Whilst persons without PTSD use more positive words. Topic modelling will be examined to see underlying topic structures and examine what is talked about. According to a study of Kleim and colleagues (2018) people with high PTSD symptomatology use fewer cognitive processing words (talk about insight, causation, certainty), more death related words, more emotional response words (fear, trauma, helplessness, horror, guilt and shame) and more frequent use of first-person singular pronouns (I, me, and my). Word count is another characteristic that will be examined. In a study of He and colleagues (2012), respondents with PTSD use more words in the self-narratives in comparison to respondents without PTSD.

In the next section the methods used in this study will be described. The text-mining techniques will be examined by using the text-mining application ORANGE and quantitative

data will be analyzed using SPSS. After the method section the results will be elaborated on to lastly formulate a conclusion to the research questions.

Methodology

Design

This study used text-mining techniques to identify differences in word count, topics, and sentiment on self-reported open-ended questions written by COVID-19 patients 12 months after hospitalization. This study is part of a longitudinal study of the MST hospital into the wellbeing of COVID-19 patients after hospitalization discharge. The data was collected in the hospitals MST in Enschede and ZGT in Almelo. 12-months after the hospitalization period the patients filled in a questionnaire about their clinical and psychosomatic symptoms. This survey consisted of multiple parts. Two of those parts were the textual writing in response to open-ended questions and filling in the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5) questionnaire (Boeschoten et al., 2020). The PCL-5 was filled in to indicate PTSD symptoms. The open-ended questions were filled in and analyzed with text-mining to examine the differences. The questionnaire was sent to the participants by email.

The scores on the PCL-5 were used to divide the respondents into two groups to examine their differences. One group scoring high on PTSD symptoms and one group scoring low in PTSD symptoms. The textual data of these two groups was compared using word count, topic modelling and sentiment analysis.

Participants

The data derived from 73 persons in the 12-month follow-up. The persons filled in the questionnaire in November 2021, which means that their hospitalization period was in November 2020. By looking at the data from the Dutch government, in November 2020 there was a second peak of hospital admissions during COVID-19 (Rijksoverheid, 2022). In the period of November 2020 there was a lockdown due to the rising infections. Because of the rising infections there were measures taken by the government, such as only two visitors a

day and the closing of public places. This affected the hospital as well, in which patients could only receive one visitor per day (personal communication, October 25, 2021). The patients were in a lock chamber and the nurse would monitor the patient 4 times a day, dependent on the hospital ward. The 73 participants have all been admitted to the hospital. However, the data gives no insight in which ward the patients were in. Therefore, it is not possible to differentiate the patients on types of wards they were admitted to.

From the original 73 participants, the baseline demographical data of five participants was missing. In addition, one person did not fill in the PCL-5 at the 12-month follow-up. These participants were excluded from the data analysis. In the textual data three persons were excluded from further analysis. One participant was deleted, because he only gave an ‘X’ as response. Another case was deleted because the person wrote ‘nee’ [no] for both questions, which means the person did not want to fill in the question. The last person was deleted because that person referred to his website and said nothing about his experiences. The 64 remaining respondent cases included 21 females (32.8%) and 43 males (67.2%), with a mean age of 62 (SD=10.84, range 32 to 82). The mean age of the group is quite high, which means that mostly older people were admitted to the hospital. In addition, 34 participants (53.1%) were hospitalized in MST and 29 persons (45.3%) in ZGT, and one person’s hospital location data was missing.

To split up the data in two groups the 64 participants were analyzed on PTSD symptomatology by the PCL-5 questionnaire. The total 64 participants had a mean score of 9.47 (SD=10.21). The standard deviation is high, which means there was much variability in the scores (range 0 to 60). Most people scored low on PTSD symptoms, with the most frequent score of one. Only one person had a score that indicated clinical PTSD.

Using the data above, the two groups were separated into a high and low total score group to compare the textual data. The data was split up by the highest scores and lowest

scores on the PCL-5 test. The 10 highest scored group scored 19 points or more on the PCL-5 questionnaire, which indicates that some traumatic symptoms were present. The group of 10 lowest scored participants had total scores that were either 0 or 1, which means that barely any traumatic symptoms were present. In which 6 persons scored 0 and 4 persons scored 1. However, there were 7 respondents in the group that scored 1. Therefore, the low scoring group was expanded to 13 participants. Resulting in the comparison of 10 participants scoring high on PTSD symptomatology respondents and 13 participants scoring low on PTSD symptomatology.

The high scoring PTSD group consisted of five males (50%) and five females (50%) with an average age of 62 years (SD=11.00) and a minimum age of 46 and a maximum of 77. In addition, five participants (50%) were hospitalized at ZGT and five participants (50%) at MST. Next to that, the respondents were asked how they perceived their general health on this moment, in which 7 participants (70%) perceived their general health as moderate and 3 participants (30%) as good. 7 respondents (70%) had another condition next to the COVID-19 infection, and 3 persons (30%) choose the option of ‘‘otherwise, namely’’, and then specified their other conditions or described that they did not have any other condition. The low scoring PTSD group consisted of 8 males (61.5%) and 5 females (38.5%), with an average age of 60 (SD=10.88) and a minimum age of 42 and a maximum age of 73. Eight respondents (61.5%) were hospitalized in MST and five respondents (38.5%) in the ZGT hospital. Next to that, the respondents were asked how they perceived their general health on this moment, in which 3 participants (23.1%) perceived their general health as moderate, 7 participants (53.8%) as good, and 3 participants (23.1%) as very good. 7 respondents (53.8%) had another condition next to the COVID-19 infection, and 6 persons (46.2%) choose the option of otherwise, namely. The two groups were compared on word counting, topic modelling and sentiment.

Measurements

Next to the PCL-5 the patients filled in two open-ended questions about the hospitalization period and the post-hospitalization period. These two questions were: ‘‘Kunt u het verhaal vertellen over uw opname in het ziekenhuis vanwege het Corona-virus?’’ [Could you tell the story about your hospital admission because of the corona-virus?] and ‘‘Kunt u het verhaal vertellen over hoe uw leven na uw opname is verlopen?’’ [Could you tell the story about your life after the hospital admission due to corona]. The participants were invited to write what comes to mind without a maximum number of words. The textual data of the two questions will be combined in the analysis.

Analyses

Software

ORANGE. For text-mining a program was used to distribute and analyze the textual data. Text-mining can be done in multiple programs, but in this study the program ORANGE was used. The program orange was developed by the Bioinformatics Lab in Slovenia in collaboration with the GitHub open-source community. ORANGE is a component-based data mining software program that analyses data. The software uses machine learning and creates a visualization of the data. ORANGE is the best choice for researchers that are not data scientists, because it is relatively easy to use. ORANGE uses widgets, which can be dragged on the board and easily connected. These connected widgets are called pipelines or workflows. The pipelines that were used in this study are described in the data analysis section. For this study a text-mining module was added to ORANGE, with multiple widgets to analyze text. ORANGE is free to use and can be downloaded on the website (orange, n.d.). In this study the 3.30.1 version of ORANGE was used. In order to analyze the text, the textual data first needed to be preprocessed.

IBM-SPSS. In this study the software of SPSS was used. SPSS is a software for performing statistical tests, which was used for statistically differentiating the two groups. The license of SPSS is available through the University of Twente. For this study, the SPSS version 27 was used.

WORD. WORD is a text processing program which is made by Microsoft. This program was used for word count, to count the number of words that the participants use. The license of WORD is available through the University of Twente. For this study, WORD version 16.57 is used.

Preprocessing

The text needed to be preprocessed in order to classify text to improve the efficacy and to be able to generalize unseen data. In this study two forms of preprocessing were used. The first one was to preprocess by hand. This included correcting spelling mistakes, abbreviations written in full, deleting double spacing, linking punctuation to the words, and deleting double punctuation. This was especially necessary for the method of word count because punctuation was seen as a word on itself, which is incorrect and would lead to invalid results. For example, one person wrote ‘‘maken ???’’, in which the double use of punctuation was unnecessary, and by not linking the words to the punctuation the question marks would have been analyzed as one word. Next to that, for sentiment analysis the abbreviations need to be written in full to give a certain sentimental score to the word. An example of this is the abbreviation ‘‘ivm’’, which was transferred in ‘‘in verband met’’ [in relation to].

The second preprocessing strategy was using the preprocess widget in Orange, which was used with the methods of topic modeling and sentiment analysis. This involved tokenization, normalization (stemming and lemmatization), Filtering (deleting stop-words). Tokenization includes breaking up sequences into meaningful pieces such as words. This was

done with regexp /w+, which converts the text into individual words and omit the punctuation. The text was transformed into lowercase and stop-words were filtered out. In order to do this a pre-set list formed by a NLTK package had to be used. This list contains frequently occurring pronouns, prepositions, conjunctions, and adverbs. Furthermore, n-grams were created to tokenize consecutive sequences of words. In this study bigrams were used with the topic modeling method, to examine pairs of words, these pairs give more information about what is meant with the word. Unigrams were used with sentiment analysis and the word frequency part of topic modeling, which consist of single words. In addition, a Part-Of-Speech (POS) tagger was used to distinguish between the same words that have different meanings. For example, a desert is a dry area of land, whilst used as a verb it is meant as abandoning people. In this study the Averaged Perceptron Tagger was used, which is software that is built by linguistic rules on which word could occur in a certain position and therefore has a certain meaning.

Word count

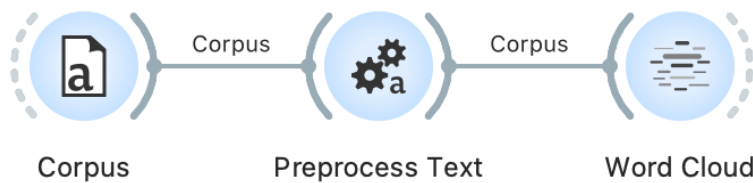
In this study the length of text from the two groups was compared, to see if there is a difference in length of responses to the questions between the groups. This was done by selecting the text in WORD and counting the number of words per respondent. These findings were transferred to SPSS for further analysis. The mean of both groups was calculated and compared to see if there is a significant difference or not by using the Mann-Whitney U test.

Topic modeling

Topic modeling consists of two parts, word frequency and topic modeling by machine learning. Within the section of word frequency, the most frequent used words were calculated to investigate what topics the groups write about. This was done by using a pipeline in ORANGE, which is shown in figure 1.

Figure 1

Word frequency pipeline in ORANGE



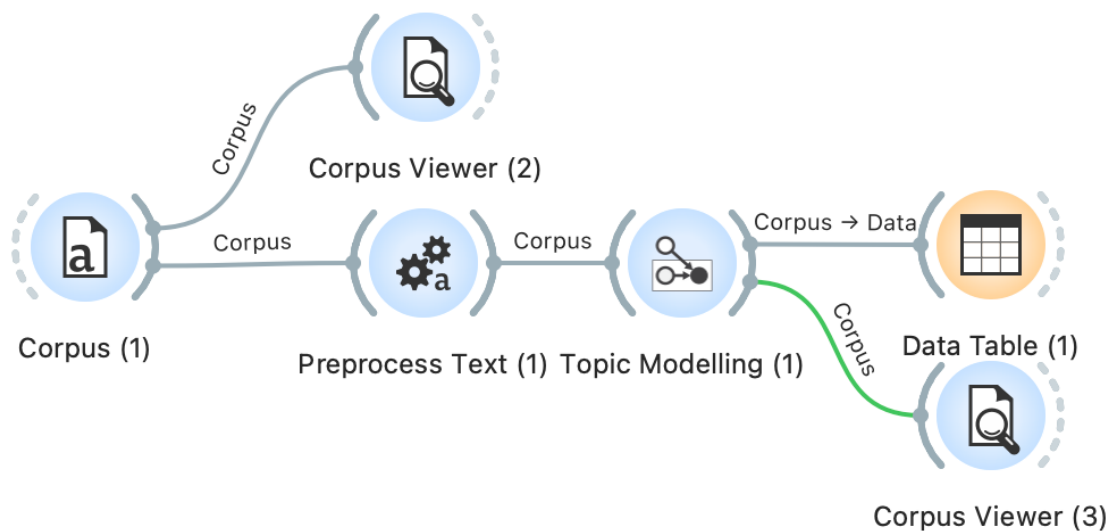
The textual data was aggregated into a word cloud, to have a clear overview of the topics per group. In addition, a list of most frequent used words was analyzed. However, when aggregating the texts, it seemed that when someone used a word 4 times, it was seen as a frequently used word by the whole group. Therefore, the data had to be deduplicated, so that frequently used words by one person would count as 1. This was done by using a deduplication program written by Dr. Erik Tjong Kim Sang (personal communication, February 4, 2022). This program was written in Python and could be applied in the Terminal feature on the MacBook. The program deduplicated the words per respondent. After this, the new data set was loaded into ORANGE and the pipeline of figure 1 was used. The results are shown in a word cloud and in a data table.

Another form is using machine learning for topic modelling, which was also used in this study. As cited in Imel and colleagues (2015). ‘‘Topic models are data-driven, machine learning procedures that seek to identify semantic similarity among groups of words.’’ It is similar to the term of factor analysis, in which underlying dimensions can be found. It is a method of analyzing content, what is talked about. Topic modeling discovers topics in a corpus by analyzing clusters of expressions in the text that it thinks are related. There are three topic modelling algorithms: the Latent Semantic Indexing, Latent Dirichlet Allocation and Hierarchical Dirichlet Process. In this study Latent Dirichlet Allocation (LDA) was used.

LDA is an unsupervised model that learns topics from the data. It can connect similar words and distinguish between words with multiple meanings. The result is an underlying structure with different topics, each topic is a set of terms that have a shared theme. This shared theme needs to be found, interpreted, and labeled by the researcher. By comparing the topic keywords and trying out different amounts of generated topics the clearest and fitting underlying structure can be found. By comparing the topic keywords by looking at the context in which the keywords occur a possible underlying theme could be interpreted and eventually labeled. In figure 2 the pipeline of topic modeling is shown in ORANGE. The widget of topic modelling also displays the topic weight per document in a data table. These topic weights show the weights that the respondents' texts have on a certain topic, which could indicate the most important topic in the texts. In this study the topics were analyzed and interpreted for both groups. The results are shown in a data table.

Figure 2

Pipeline of unsupervised topic modeling



Sentiment analysis

A sentiment analysis was conducted on the open-ended questions texts in order to compare the sentiment of the two groups. Sentiment addresses the positivity or negativity of

the message transferred in the text. Sentiment analysis is a natural language processing technique to determine if the texts are positive, negative, or neutral (Monkeylearn, n.d.-b). It is a general method to extract subjectivity and polarity from text (Taboada et al., 2011). There are two main approaches in sentiment analysis to extract sentiment from text automatically, the dictionary and machine learning based approach (Pang, & Lee, 2008; Boldenthusiast, 2019). The machine learning approach involves the use of machine learning algorithms and predicts the sentiment of text. It uses examples of positive, negative, and neutral classes and trains the model based on those examples (Itech, 2021). Eventually the outcome of the text is a polarized sentiment, so positive, neutral, or negative. The dictionary approach uses dictionaries with words that have a predefined valence to denote their polarity and sentiment strength (Itech, 2021). These valence scores per word together calculate the sentiment score. The dictionary approach could be split up in a rule-based and a lexicon-based approach. The rule-based approach uses predefined grammatical structures to calculate the valence scores, whilst the lexicon-based approach calculates the sentiment by using the valence of words from predefined dictionaries.

In this study the lexicon-based approach was used to identify sentiments in the text. To make this classification in written texts Multilingual sentiment was used. It is a lexicon-based sentiment analysis tool for multiple languages and can be used in the both the Python and ORANGE software (Hutto, & Gilbert, 2014). It is a pre-existing language data set available in Dutch which recognizes the words and gives a valence to it. It is sensitive to the subjectivity of expressed sentiments in text. The valence scores indicate both the sentiment polarity (positive/negative) of the text and the sentiment intensity, in which a score below 0 indicates a negative sentiment and vice versa. The discrepancy from the 0 indicates the intensity of sentiment of the text. The sentiment intensity could be increased or decreased by a word that could change the semantic intensity of a lexicon. For example, ‘‘very’’ could

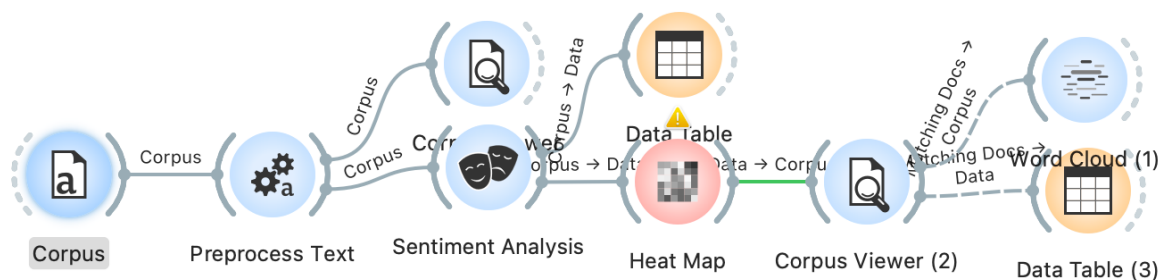
increase the valence of the lexicon item ‘‘good’’ (Taboada et al., 2011). The outcome is a score (positive/negative), which indicates the overall valence of the text written and the research interprets the polarity or overall sentiment of the texts. These outcomes per respondent were compared to the other group scoring high or low on PTSD symptomatology.

It is expected that positive phrasing will occur more in the texts of people that have little traumatic symptoms. According to a study of Sawalha and colleagues (2022) sentiment analysis could identify people with high PTSD symptoms. Individuals with PTSD use more frequently neutral and negative words. Whilst persons without PTSD use more positive words.

In this study ORANGE is used to perform the sentiment analysis, in figure 3 the pipeline of the sentiment analysis is shown with the widgets that were used. The results are shown in a heatmap, and SPSS is used to investigate the differences of the two groups. The groups were compared by the Mann-Whitney U test.

Figure 3

Pipeline of sentiment analysis in ORANGE



Results

Word count

In table 1 the differences in word count between the groups are shown. The total amount of words used by the high scoring PTSD group is 3227, with an average amount of words of 322.70 (SD=572.04). However, there are major discrepancies in the data as seen in the min and max data in the table. This is because of one outlier in the data. One person wrote 1875 words, whilst the follow up person wrote only 572 words.

The low scoring PTSD group used 1318 words in total. The minimum number of words used is 13 and the maximum number of words used is 229, with an average of 101.38 words (SD = 68.57).

Differences in mean are shown in table 1. The data is not normally distributed (Kurtosis = 17.896, skewness = 4.073). There are multiple outliers in the data that could affect normality. However, because of the small sample size outliers cannot be removed. Therefore, a non-parametric test, the Mann-Whitney U test, is used. According to table 2, there do not seem to be a significant difference between the two groups (Mann Whitney U = 69, $p = .832$). Although there seems to be a difference in average amount of words used, these findings do not support the conclusion that there is a significant difference in number of words used between the two groups.

Table 1

Comparing word counts between the 2-groups by using two-sample t-test and Independent-Samples Median Test

						Mann Whitney-U	
Group	Observations	Total	Min/max	Mean	SD	U	p
Highest PTSD scores	10	3227	5/1875	323	572	69	.832

Lowest	13	1318	13/229	101	69
PTSD					
scores					

Note. With the independent samples t-test the equal variances were not assumed.

Topic modeling

Word frequency

In figure 4 the word cloud of the low scoring PTSD group is shown, which displays the most frequent used words per group. In figure 5 the word cloud of the high scoring PTSD group is shown. Comparing the word clouds of the high and low scoring PTSD group the word ‘‘Hospital’’ (ziekenhuis) is used frequently by both groups, which is expected because one of the questions was about the experiences in the hospital and both groups were hospitalized. In addition, the word ‘‘Admission’’ (opname) was frequently used, this can also be explained by the subject of the other open-ended question, which is about their hospital stay. These two words are mentioned, but not included for further analysis.

The most frequently used word by the low scoring PTSD group is ‘‘again’’ (weer), which is named 10 times by the respondents. Other frequently used words are: ‘‘Hospital’’ (ziekenhuis), ‘‘Admission’’ (opname), ‘‘went’’ (ging), ‘‘good’’ (goed), ‘‘fast’’ (snel), ‘‘days’’ (dagen), ‘‘dag’’ (day), ‘‘very’’ (heel). In which ‘‘Hospital’’ is used 8 times and ‘‘Admission’’ 7 times, the other words are used 6 times. An example of a sentence in which the word ‘‘again’’ is mentioned is: ‘‘So far, everything is going back to normal’’ (Tot nu toe gaat alles weer normal). Other examples relate to going back to work, which is mentioned 7 times. The word ‘‘went’’ is used in the context of doing better, such as ‘‘Went well, so I was allowed to go to a regular ward’’ (ging goed waardoor ik naar een gewone afdeling mocht), ‘‘went a little better’’ (ging een beetje beter) or ‘‘went so well that’’ (ging zo goed dat). In addition, this word is frequently used next to the word ‘‘good’’. ‘‘good’’ is used in the

context of a fast recovery and in the context of how well the respondent's physical condition is now. "Fast" is used in the context of a fast recovery, such as feeling better fast. "Day" and "days" are used frequently, however in different contexts. "days" is used to represent the number of days spend in the hospital, whilst "day" represents the time of a certain event. An example of this is "was admitted to the hospital the same day" (dezelfde dag opgenomen) or "admission to the hospital at the end of the day" (opname aan het einde van de dag). The word "very" is used as a superlative in the texts, mostly indicating very positive experiences regarding the physical condition.

Overall, the low scoring PTSD group write about fast recovery and feeling better. In which feeling better applies to an increase in physical condition and going back to work. In addition, superlatives are used in a positive context, indicating positive experiences regarding the physical condition.

Figure 4

Word Cloud low scoring PTSD group



Frequently used words from the high scoring PTSD group are displayed in figure 5. Words that are frequently used are: ‘terribly’ (erg), ‘hospital (ziekenhuis)’, ‘very’ (heel), ‘well’ (wel), ‘good’ (goed), ‘have had’ (gehad), ‘days’ (dagen), ‘after that’ (daarna), ‘admission’ (opname), ‘visitation’ (bezoek), ‘week’ (week), ‘later’ (later), ‘sick’ (ziek), ‘fast’ (snel), ‘man’ (man), ‘day’ (dag), ‘corona’ (corona), ‘alone’ (alleen), ‘tired’ (moe) and ‘lonely’ (eenzaam). In which ‘terribly’ is used 6 times, ‘hospital’ and ‘very’ are used 4 times and the other words are used three times. The words ‘terribly’ and ‘very’ are used as a superlative. These words are used in combination with another word, such as, a superlative amount in shortness of breath, a lot of sweat, very emotional, very hard or very lonely. This indicates that the respondents feel like they have had an extreme negative experience. The word ‘well’ is used in the context of someone’s condition, although they experienced a lot, they are doing a little better. An example of this is: ‘Doing a little better’ (wel iets beter met mij). The word ‘good’ is used in the context of not good, writing about their physical condition. However, this word is also used in the context of the quality of care in the hospital, in which one person was very satisfied with the quality of care and someone else was not satisfied with the hospital admission quality. Respondents that use the word ‘have had’ frequently write about certain procedures in the hospitalization, such as ‘received medical pills’ (pillen gehad) and ‘had a follow-up’ (nacontrole gehad). The words of days/day or week is in the context of the time spend in the hospital for recovery. ‘After that’ is used by indicating the process and the order of events. The words ‘visitation’, ‘alone’ and ‘lonely’ seem to appear in the same context. In which the group writes about the absence of visitors in the hospital and the feeling of loneliness that came with that. In addition, ‘visitation’ and ‘man’ is about the patients surrounding and the absence of them. The word ‘fast’ is used in a negative way, so not as fast recovery, but in the context of something needs to happen fast or there will be negative consequences. An

example of this is the sentence of ‘‘would fast be over’’ (snel gebeurd met mij) or ‘‘Had to take action fast’’ (snel aan worden gepakt). ‘‘sick’’, ‘‘corona’’ and ‘‘tired’’ is about the physical condition of the patient in a negative context.

Overall, the high scoring PTSD group writes about implications regarding the physical condition, procedures, and not doing well physically. In addition, they use superlatives when writing about experiences, mostly negatively formulated. After the period of sickness, they seem to do a little better, however they write about a bad physical condition and implications. Lastly, the group writes about the absence of persons around them during the hospital admission, and the feeling of loneliness as a result.

Figure 5

Word cloud of high scoring PTSD group



In table 2 the most frequent words per group are compared to indicate differences between the groups. A positive difference indicates that the word is more frequently used by

the low scoring PTSD group, whilst a negative difference indicates that the high scoring PTSD group more frequently use the word.

Table 2

Comparing the most used words of the two groups

Word	Frequency low PTSD group	Frequency High PTSD group	Difference
Again	10	2	8
Day(s)	11	6	5
Good	6	3	3
Fast	6	3	3
Home	5	2	3
Very	5	4	1
Well	4	3	1
After that	2	3	-1
Week	2	3	-1
Later	2	3	-1
Tired	2	3	-1
Terribly	4	6	-2
Visitation	0	3	-3
Sick	0	3	-3
Alone	0	3	-3
Lonely	0	3	-3

Note. Negative differences mean the words are more frequently used in the high PTSD group. Positive differences mean the words are more frequently used in the low PTSD group.

The low scoring PTSD group write more about ‘‘again’’ (weer) and ‘‘day(s)’’ (dag(en)), but also ‘‘good’’ (goed), ‘‘fast’’ (snel) and ‘‘huis’’ (home). An example of a sentence in which the word ‘‘again’’ is mentioned is: ‘‘So far, everything is going back to normal’’ (Tot nu toe gaat alles weer normaal). Other examples relate to going back to work, which is mentioned 7 times. ‘‘day(s)’’ indicates process and the number of days spent for recovery or hospitalization. For example, ‘‘was discharged on the 5th day’’ (ben de 5e dag ontslagen) and ‘‘It got a little better every day’’ (Elke dag ging het een beetje beter). An example of the word ‘‘good’’ used in a sentence is ‘‘recovery went well beyond

expectations'' (herstel wel boven verwachting goed verlopen) and ''Took good care of me'' (heeft me goed verzorgd). The word ''fast'' is used three times more, than with the high PTSD group, by using for example the words ''fast recovery'' (snel weer bovenop). In addition, the word ''home'' is used in the context of going back home.

Remarkable of the high PTSD group is that a frequently mentioned topic is ''alone'' (alleen), ''lonely'' (eenzaam), ''sick'' (ziek) and ''visitation'' (bezoek), but also ''terribly'' (erg). The word ''alone'' is used as being alone, but also as only having/only experiencing this symptom. The word ''lonely'' seems to be about experiencing loneliness during hospitalization. An example of the word ''lonely'' used in a sentence is: ''Being alone in the hospital for 10 days without visitors is very lonely and not nice'' (alleen 10 dagen in het ziekenhuis zonder bezoek is wel erg eenzaam en ongezellig). The word ''sick'' is represented by the sentence of ''Very sick'' (erg ziek), but also ''often sick'' (Vaak ziek). ''Visitation'' is mostly used in the context of the absence of visitors. Lastly, ''terribly'' is used as a superlative of the experiences, such as experiences regarding the physical condition of the person ''terribly sick'' (erg ziek), ''terribly tired'' (erg moe) and ''terribly stuffy'' (erg benauwd).

Comparing the two groups it seems like the low scoring PTSD group uses more positive words, than the high scoring PTSD group. In addition, the high PTSD group uses the superlative of terribly more often. The low scoring PTSD group experienced the hospitalization and recovery period as good and fast and seem to elaborate in a neutral way on the process by using words neutral words as ''again'', ''day(s)'' and ''home''. Whilst the high scoring PTSD group seem to write about their physical complaints and being lonely without visitors.

Topic modeling

In table 4, the frequency of respondents that have high weights on a certain topic and the topic keywords for the low scoring PTSD group are shown. By comparing different amounts of topics, it seems that with a LDA of three topics the clearest pattern in topic keywords is demonstrated. Therefore, the topic keywords have an underlying structure of three topics. The topic with the highest overall weight is topic number 1, 39% of the respondents have high topic weights on this topic. One of the topic keywords is admission to the hospital, in which the respondents write about the length and quality of their admission. Another topic keyword is about the challenge regarding the physical condition, which seems to be worse than before the hospital admission, for example one person wrote: ‘‘My current physical condition and energy is lower than before’’. Another keyword is also about challenges after hospitalization, these are the keywords of ‘‘tested current’’, and ‘‘current challenge’’. In addition, ‘‘went home bother’’ (huis gegaan last) is about the physical implications the person still has. However, one keyword in this topic seems to contradict the other keywords of the topic, in which the respondent writes about not experiencing symptoms: ‘‘but did not experience symptoms, however there was a roommate...’’ (maar had nog geen last, maar er was een huisgenoot...). Although the last keyword does not seem to fit the other keywords, an appropriate overall label for the first topic that is characteristic for the low scoring PTSD group seems to be challenges after hospitalization.

The second topic identified by the topic modelling tool, was equally represented in 39% of the low scoring PTSD group. This topic seems to be as important as topic 1. The first topic keyword is about the respondents talking about their hospital admission in the past. Another topic keyword is the heavy recovery after hospitalization, meaning that the respondents found the hospitalization period shocking. This is also shown by another keyword, the ‘‘troubling period again’’, where the respondents write about their

hospitalization period as impactful. It seems that all the keywords are related to looking back at the whole sickness and hospitalization process, with all the different implications. In addition, another topic keyword is about feeling better, so it seems that after the hospitalization process the respondents feel better. The last topic keyword is “both had different” (Hadden allebei ander), which is in the context of two people having a different course of disease. The keywords seem to be mostly about the hospitalization process and the course of the illness. Overall, the topics keywords seem to be related to one overall topic label, namely evaluating the hospitalization process and illness trajectory.

Finally, 23% of the respondents of the low scoring PTSD group had a high topic weight on topic 3, in which some keywords are about the amount of days spent in the hospital after going back home, for example “Days again”, “5 days”, and “going home”. In addition, the keywords seem to be about medication, which is in the context of experiencing the positive effect of the medicine and getting well, but also throwing up the medicine due to sickness. Lastly, one keyword differentiates a lot from the rest, which is about being a diabetic patient and therefore experiencing symptoms. Looking at the different keywords, the topic seems to be about recovery by medication leading up to discharge of hospital, which lead to a label of recovery by medication.

Overall, topics 1 and 2 are equally and mostly written about. Which is about challenges after hospitalization and evaluating the hospitalization process and illness trajectory. Another topic is about recovery by medication.

Table 4

Overview of three main topics with belonging tokens of low PTSD group.

Topic	Frequency (%)	Topic keywords	Topic label
1	5 (39%)	Admission to hospital, went home bother, current challenge physical condition, lower than before admission, roommate in contact	Challenges after hospitalization

		positive, tested current, current challenge.	
2	5 (39%)	Located in hospital corona, heavy recovery, both had different, luckily feel again, troubling period again, while ago hospital.	Evaluating the hospitalization process and illness trajectory
3	3 (23%)	Days again, 5 days, going home, medication good, after home, keep in medication, also diabetic patient, medication had good effect.	Recovery by medication

In table 5 the topic weights and the topic keywords of the high scoring PTSD group are shown. Investigating the topic weight, a total of four underlying topics shows the clearest structure in topic keywords, and could show the clearest underlying patterns. Whilst when looking at the keywords of three or five topics, the topic keywords do not have clear underlying similarities. In the high scoring PTSD group, 40% had a high topic weight on topic 1. In topic 1 the respondents write about hospital procedures such as MRI-scans and operations. An example of this topic is: “During corona I endured four hip operations” (tijdens corona ook nog vier heupoperaties gehad). Which is about procedures that took place during the corona infection and hospitalization. The word of MRI-scan is used by one person and described in the context of wanting an MRI scan of the body to check if everything is ok after hospitalization. The topic keywords of “still very” (steeds heel) are written in the context of complaints the patient still has, in a superlative form. Other topic keywords indicate a certain number of times an event occurred, such as “two times” (twee keer) or “times per” (keer per). In addition, tired and emotional are relevant keywords to the topic. In which the respondents write about the period of sickness as being emotional and still feeling emotional while recovering. During the sickness period the respondents felt tired, and tiredness overall seems to be present in the high scoring PTSD group while recovering. Furthermore, there is a keyword of “days in bed” (dagen bed), which related to the period during sickness, in which respondents stayed in bed for multiple days. Lastly, there is a keyword which is about stuffy and shortness of breath. Overall, the topics of MRI-scan,

operations, days in bed, and the indication of number of times that something occurred seems to be about multiple procedures, whilst the other keywords are about implications that are still present. Therefore, the topic seems to be about repetitive procedures and long-term persuasive implications.

Furthermore, 30% of the high scoring PTSD group had a high weight on topic 2. This topic also exists of a few topic keywords. The first topic keyword is "every day" (elke dag), which relates to an event occurring every day. However, the context in which these words are used is different. For example, one person states: "I have cold legs every day" (Ik heb elke dag koude benen), whilst another person states: "Try to enjoy every day" (Probeer te genieten van elke dag). Next to that, the word is used in the context of checking in on someone daily. Another keyword that indicates time is "per week", which is mentioned in the context of the number of times visiting the therapist after hospitalization. It is also mentioned with the context of number of times experiencing certain complains after hospitalization. "Per week" seems to be only written about after the hospitalization period, whilst "every day" is used for the time in and out of the hospital. The keyword of therapist is mentioned by two persons, and dietician is mentioned by one person. One person writes about the therapist and dietician that helped her recover after hospitalization. The other person writes about the absence of the therapist consults due to their admission to the hospital. However, with both persons it is unclear what kind of therapist they meant. Another keyword to the topic is the admission to the hospital. Which is mentioned by four people. This word seems to be related to three other keyword combinations, namely "general practitioner later pulmonologist"; "general practitioner and general practice center"; "general practice center eventually admitted". In which the general practitioner, pulmonologist and general practice center referred the patient for admission to the hospital. An example of this is "had often consultation with general practitioner and general practice

center. Eventually admitted to the hospital...” (had vaak contact met de huisarts en de huisartsenpraktijk. Uiteindelijk ben ik opgenomen...). Lastly, “quite” (wel erg) is one of the topic keywords. “Quite” is mostly followed by another word, such as quite long, relating to the duration of the day when quarantined. Another example is quite lonely. The topic seems to be about different professionals that admitted the patient to the hospital and helped the patient with the recovery after hospitalization. However, that theme does not fit with all the topic keywords. For example, that theme does not cover the time-related keywords. Combined, the topic seems to be about indispensable and important professions during and after COVID-19 infection.

Topic 3 is mentioned by two respondents. The first topic keyword is literally translated as nothing hand (niks hand), however is used in the sentence as “nothing wrong”, indicating that shortly before the period of sickness nothing seemed to be physically wrong. Another keyword indicated the permission by the doctors to go home after the hospitalization period. The keyword of “made a thorax X-ray” (longfoto genomen) is written by two respondents. One in the context of making a thorax X-ray in the hospital and afterwards being diagnosed with COVID-19 lungs, and one would like to undergo a X-ray to see if everything is ok with the lungs after hospitalization. The keyword of “terrible skin rash” (vreselijke huiduitslag) is the result of certain medication. The keywords of “Admission little” (opname weinig) and “admission little of” (opname weinig af) are written by the same person in the sentence of: “I know little of the admission” (Ik weet van die hele opname weinig af). Indicating that the person barely remembers the hospitalization period. The two keywords of “died hallway” (overleden gang) and “eye closed screaming” (oog dichtgedaan geschreeuw) seem to be related. Two persons wrote about hearing the caretakers in the hallway talking about people that died due to COVID-19. One person wrote that at the same time she heard people screaming in the hallway out of fear and did not sleep because of it.

The conversations in the hallway, and the screaming caused sleeplessness for the respondents. The different topic keywords seem to contradict each other. For example, they write about nothing being wrong and going home, however other topic keywords relate to physical illness such as taking an X-ray or being admitted to the hospital. When looking for an overall view of the topic keywords, most topics seem to be about unexpected shocking experiences and outcomes during and before hospitalization. Therefore, the overall label of the keywords is impactful confrontational unexpected experiences and outcomes.

Topic 4 has keywords about the nurses and visitors. In which the respondents write about how nice and great the care in the hospital was. In addition, they write about the absence of visitors and loneliness. The overall theme of this topic could be experiencing loneliness but receiving good care during hospitalization.

In short, the high scoring PTSD group write mostly about topic 1, which is about procedures and implications. Next to that they write about professions during course of disease, the hospitalization period, and personal surroundings during the hospitalization period.

Table 5

Topics with belonging tokens of high scored group.

Topic	Frequency (%)	Topic keywords	Topic label
1	4 (40%)	MRI-scan, tired emotional, four hip operations, two times, times per, still very, very emotional, days in bed, stuffy and shortness of breath	Repetitive procedures and long-term persuasive implications
2	3 (30%)	Every day, therapist dietician, admission hospital, quite, per week, general practitioner and general practice center, general practitioner later pulmonologist, general practice center eventually admitted.	Indispensable and important professions during and after COVID-19 infection

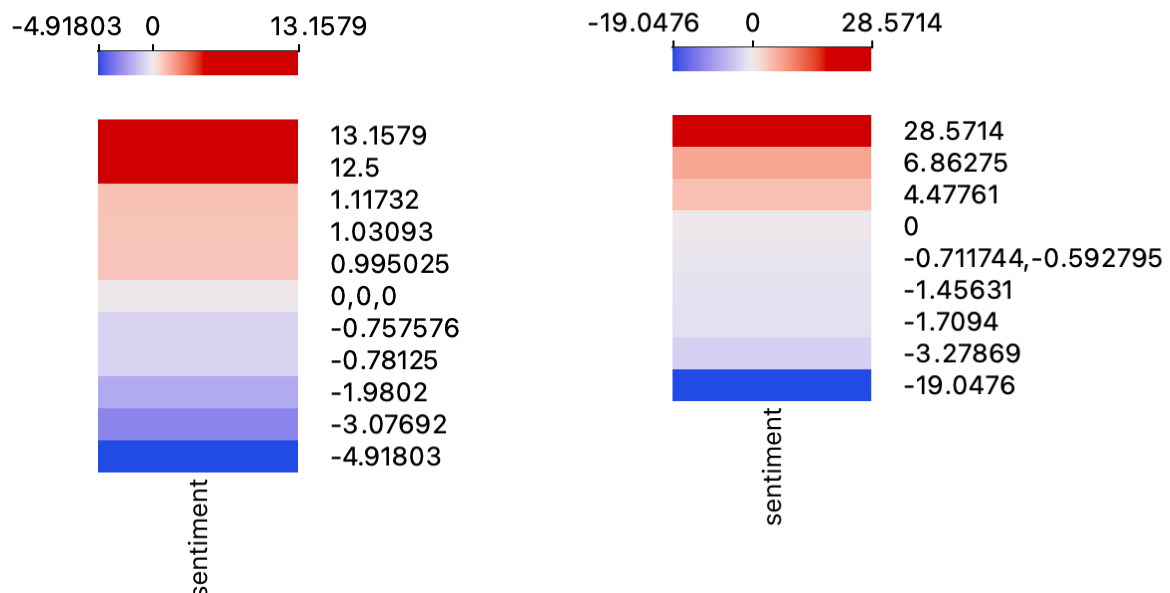
3	2 (20%)	Nothing wrong, allowed home, made a thorax X-ray, terrible skin rash, admission little, admission little of, died hallway, eye closed screaming.	Impactful confrontational unexpected experiences and outcomes
4	1 (10%)	Friendly nurses, time caretaking, lonely visitors, nursing lonely, caretaking fine.	Experiencing loneliness, but receiving good care during hospitalization

Sentiment

Figure 2 displays the sentiment of the low and high scoring PTSD group in a heatmap. The numbers on the side of the heatmap describe the sentiment scores per individual. The colors represent the height of the sentiment scores. Matching colors represent matching sentiment scores.

Figure 2

Heat map on sentiment of low scoring PTSD group (left) and high scoring PTSD group (right)



In the low scoring PTSD group five respondent's texts include positive sentiment, with two outliers with a score of 12.5 or higher. Three person's text had a neutral sentiment.

Lastly, the other 5 persons overall had a more negative sentiment in their written texts. The whole low scoring PTSD group had on average a positive sentiment ($M = 1.31$, $SD = 5.32$). However, as said earlier there are big discrepancies in the data ($Range = 17.73$). In the low scoring PTSD group, the same number of people scored high as low on sentiment. A person that scored high on sentiment and has therefore a positive sentiment had tokens as: ‘‘be alright’’ (weer bovenop) and ‘‘very good’’ (erg goed). Respondents with high sentiment did not seem to write about a specific emotion. However, the word ‘‘good’’ (goed) was mentioned three times. A person that scored low on sentiment (negative) had tokens as: ‘‘felt lonely’’ (eenzaam gevoeld), and ‘‘anxious’’ (angstig) but also tokens as ‘‘very satisfied’’ (erg tevreden). In addition, a certain emotion that was written about was fear, described as ‘‘anxious’’.

The high scoring PTSD group scored between -19.05 and 28.57 on sentiment in the text. The range between the answers of the respondents is big ($range = 47.61$), with two outliers in the data of -19.05 and 28.57 as can be seen in figure 2 by looking at the colors. The average sentiment from the texts for this group is positive ($M = 1.31$, $SD = 11.76$). However, there are more negative sentiment texts than positive. An example of a text with a high score (positive) on sentiment is: ‘‘Not quite well’’ (niet helemaal goed) and ‘‘quite well’’ (wel goed). A person with a negative sentiment used words as: ‘‘lonely’’ (eenzaam), ‘‘very sick’’ (erg ziek), ‘‘crying’’ (huilen), ‘‘anxious’’ (angstig) and ‘‘control’’ (controle).

The data is widespread. Therefore, a non-parametric test is used as well. The outcome of the non-parametric test is that there is no significant difference between the groups (Mann-Whitney U test = 58.50 , $p = .693$). Based on the outcomes it can be stated that there is no significant difference between the high and low scored group on sentiment. However, the group of high PTSS scores have relatively more people with negative sentiments, than the low PTSS scoring group. Next to that, the group of low PTSS scores have relatively more

people with positive sentiments than the other group. Furthermore, both groups have outliers and therefore both groups have wide ranges.

Discussion

For this study, it was argued that the textual characteristics of word count, topic modeling and sentiment could differentiate the high and low scoring PTSD group and could therefore indicate PTSD symptoms. Those characteristics could help indicate the risk of high PTSD symptomatology in written text after hospitalization in the future. The conclusions of this study could be used by hospitals to predict and prevent PTSD early. There were no significant differences found, yet the analysis showed qualitative differences in how patients narrate their experiences. The differences substantiate words and topics used that could indicate PTSD symptomatology.

Findings on patient textual differences

Examining word count was done to explore textual characteristics that could differentiate the two groups and would therefore indicate PTSD symptomatology. According to the outcomes of this study there is no significant difference between the high PTSD and low PTSD scoring group on word count. Within the groups there are very big differences in word count, with some outliers. However, although not significant, the high scoring PTSD group seem to use more words when expressing their experiences. In this study it was expected that there would be a difference between the groups. Since that has also been found in the study of He and colleagues (2012), in which respondents with PTSD used more words in the self-narratives in comparison to respondents without PTSD. In addition, a study of Gray and Lombardo (2001) did not find a significant difference in word count but did find that the lengthiest narratives belonged to the PTSD group. In the study of He and colleagues (2012) the PTSD group had the shortest texts as well as the longest text, which is also seen in the results of this study. In the results of the same study, it was stated that the average length of self-narratives from people with PTSD was 150 words, which was also expected in this study. However, the average number of words in this study is larger than 150 words for the

high PTSD group. The results of this study seem to contradict some literature found in previous studies but is in line with the findings of the study of Gray and Lombardo. In which the high scoring PTSD group seem to use more words when writing texts. There is no evidence that word count can differentiate the two groups and therefore does not seem to indicate PTSD symptoms. However, there are some interesting outcomes found regarding that the lengthiest seem to belong to the high scoring group. In future research, the length of the texts could be further compared to investigate what underlying factors result in the difference in length of text, for example, it could be investigated if high scoring PTSD patients feel the need to write longer text to explain their inner world and their experiences and why.

Another aim of this study was to investigate the topics of the groups in order to find deviant topics that could indicate PTSD symptoms. The low scoring PTSD group writes about physically doing better which sometimes results in going back to work. The group also writes about a fast recovery from the disease. In comparison, the high scoring PTSD group write about that they are physically not feeling well and implications. In the literature it was found that there is an association between poor health and PTSD, in which poor health results in a negative affect, such as anxiety and anger and causes PTSD (Schnurr & Green, 2004). In addition, it was found in multiple longitudinal studies that implications could lead to a negative illness representation for PTSD (e.g., Sheldrick et al., 2006). Which includes a negative emotional representation or negative perception of consequences of diagnosis. Therefore, the physically poor health of the respondents scoring high on PTSD could impact their affect negatively and cause higher PTSD symptom severity. In addition, experiencing multiple implications could enhance the negative affect increasing the chances on PTSD symptomatology. Additionally, with this new knowledge a new hypothesis could be that fast recovery helped the patients not to obtain PTSD symptoms. Another topic written by the high

scoring PTSD group was the absence of visitors and loneliness, which was not mentioned by the low scoring PTSD group. This outcome was not expected, however, according to Fox and colleagues (2022) there is an association between loneliness and PTSD symptoms, in which the feeling of loneliness seems to cause heights in PTSD symptomatology. Another topic from the high scoring group is about procedures they endured. Procedures could be considered as stressors; however, do not necessarily cause traumatization. Especially, more intrusive medical procedures are risk factors for PTSD (Vilchinsky et al., 2017). Procedures is often written by the high scoring group. Therefore, it seems that procedures depending on the intrusiveness could indicate PTSD symptoms.

In the literature it was found that individuals with PTSD use more frequently neutral and negative words (Sawalha et al., 2022). In this study the low scoring PTSD group seem to use more positive words, such as “good”, “fast” and “again”. In addition, they use superlatives in a positive way. Meaning that the superlatives are mostly connected to positive words and in the context of a positive experience regarding the physical condition and recovery. In addition, the high scoring group uses superlatives, which are mostly used in a negative context, such as that the experience was very hard or regarding their troubling physical condition. It can be concluded that there is a qualitative difference in topics between the two groups.

Differentiating the two groups on a qualitative level the low scoring PTSD group write more about getting better and being able to go back to work. They write about the days in the context of spending days in the hospital or use the word to indicate a process over time. Furthermore, they write more about good caretaking in the hospital and that their physical state improved quickly. They also write about going home more often. The high scoring PTSD group write more frequent about loneliness and being alone. They also write more about sickness or being sick. In the literature it was found that sickness or illness could cause

PTSD (Khakpour, 2018, as cited in Zhou, 2020), which might cause the relatively high PTSD symptomatology within the high scoring group. PTSD seems highly connected to death relating words and anxiety related words (Alvarez-Conrad et al., 2001; D'Andrea et al., 2012), the usage of sick or sickness of the high PTSD group could be indirectly related to the anxiety or fear of dying. Interestingly, both groups use superlatives, however, the contexts in which the superlative is used differentiates between the groups. It is complicated to find specific literature on PTSD and superlatives. However, in one study it was found that the responses of individuals with PTSD had negative adjectives, in self-referential processing (Ford, & Courtois, 2013). This seems to be also the case for the results in this study. The low scoring group use superlatives to indicate a positive experience, whilst the high scoring group uses the superlatives to indicate a negative experience. This study confirms the more frequently use of superlatives in a negative context as an indicator of PTSD.

The topic keywords show underlying topic labels. The low scoring PTSD group have an underlying structure of three topics and the high scoring group have an underlying structure of four topics. The low scoring group writes most about challenges after hospitalization and evaluating the hospitalization process and illness trajectory. In addition, they write about medication and the recovery resulting from the medication. The high scoring group write most about repetitive procedures and long-term persuasive implications. Next to that, they write about indispensable and important professions during and after COVID-19 infection; Impactful, confrontational, unexpected experiences and outcomes; and experiencing loneliness but receiving good care during hospitalization. Most topic keywords seem to fit in the overall label, however not all topic keywords seem to fit the overall labels. It seems that writing about repetitive procedures and long-term persuasive implications; indispensable and important professions during and after COVID-19 infection; Impactful, confrontational, unexpected experiences and outcomes; and experiencing loneliness but

receiving good care during hospitalization indicate PTSD. This is also in line with the results of our finding that procedures and implications, and loneliness seem to be important indicators for PTSD.

There seem to be no significant differences between the sentiments of the group with high and low PTSD symptomatology. However, there are large discrepancies in sentiment within the groups, so in-group variation seems to exceed between group variation. According to Sawalha and colleagues (2022) it was expected that individuals with PTSD use more negative and neutral words. This was also confirmed by another study, in which PTSD expressed more negative emotions (Orsillo et al., 2004). Although the high scoring PTSD group have relatively more people with negative sentiments than the low scoring PTSD scoring group, and the low scoring PTSD group seem to have relatively more people with positive sentiments, this difference is not significant.

Limitations and future research

Considering the current study, several considerations could be mentioned that limit this study. As concluded above, there are no differences found in sentiment and word count. An explanation for this is that within the groups there are big discrepancies in scores, with some outliers. If the within-group variance is bigger than the between-group variance, then there is no evidence for any differences between the groups (Likao, 2019). Therefore, overall it might be possible that the differences within the groups are bigger than between the groups, which results in the absence of a significant difference between the groups on word count and sentiment analysis. In addition, when looking at the PTSD scores within the groups, there are big discrepancies of total scores on the PCL-5 within the high scoring group, which makes them less generalizable. Next to that, there is the question if there are two groups present or that there is another underlying structure between the individuals. When looking at the big differences within the group the question arose if the individuals can be categorized by PTSD

or that other ways of categorizing the discharged COVID-19 patients would give other interesting differences or patterns. In addition, in this study the average age of the respondents is quite high, which makes the study less representative for groups that have lower ages. In future research the similarities and differences within the groups could be investigated further. In addition, the data of the discharged COVID-19 respondents could be explored further on other similarities and differences in text between individuals, with the aim to find meaningful underlying categories or patterns. By investigating these other patterns for discharged COVID-19 patients, practical implications could be formulated and applied to increase wellbeing after hospitalization.

Sickness and implications are frequently written topics by the high scoring PTSD group. In the literature it was found that illness or sickness could cause PTSD on itself (eg., Schnurr & Green, 2004; Khakpour, 2018, as cited in Zhou, 2020). Therefore, another limitation of this study is the cause-and-effect correspondence of PTSD, sickness, and hospitalization. It might be that the correlation between hospitalization and PTSD is not because of the traumatic hospital experience, but by the experience of being sick. However, indicating PTSD is mostly done after hospitalization, which could give a false causal connection between hospitalization and PTSD, whilst the causation of PTSD is the disease or illness. Future research could investigate if PTSD symptoms occurred because of the hospitalization period or because the individual was sick.

Another limitation is that in this study the open-ended questions were investigated together, so valuable information may have been lost. One question was about the hospital admission and the other question was about life after hospitalization. These are two different questions and investigating them separately results in specific practical implications on both subjects. It is relevant for the hospital to investigate the hospital stay, eventually for

improving their services. The recovery after hospitalization is relevant for elaborating on the follow-up care that could be provided by the hospital or other healthcare organizations.

According to the literature, some people feel over surveyed by the number of questionnaires they receive (Porter et al., 2004). In addition, respondents could feel overwhelmed by the number of questions which could cause demotivation. Therefore, in this study another way of analyzing patients' data was used. The respondents wrote texts about their experiences instead of questionnaires, which could counteract the feeling of over surveying and prevent demotivation. A strength of this study is that giving the option to write about experiences openly, flexibility and creatively could prevent demotivation and give respondents the option to write about their inner world. In addition, for the psychologist analyzing textual data could be less time consuming than item development, which involves multiple time-consuming steps (Data collection, data cleaning, field trial, and examination of reliability and validity; He et al., 2019). Therefore, using textual data could have positive impacts on respondents and researchers.

Some differences in the characteristic of word and topics were found, however, not on the other characteristics. In this study an explorative approach is used to find textual differences, however this has some disadvantages. The first disadvantage is that some findings with topic modeling were vague, which made interpretation of the topics difficult. This resulted in not all topic keywords being account for by the model. Another disadvantage of the explorative approach is finding relevant literature for elaborating on the findings in the study. The last disadvantage is that the analysis is less structured due to not using predefined dictionaries, with predefined categories. An example of this is the study of Kleim of colleagues (2018) where it was found that the category of cognitive processing words was correlated with elevated heights of PTSD. Therefore, another method could be used to explore characteristics in texts more structured, less vague, and more accounted for by

literature. Using another method could classify the words in categories. This could give a clear overview of what is talked about and leaves the researcher out of the interpretation, which makes the outcome more reliable. Such an interesting other way of analyzing text is by using Linguistic Inquiry and Word Count (LIWC). It is a computerized text analysis program, which was developed by Tausczik and Pennebaker (2010). LIWC exists of a variety of hierarchical structured dictionaries, that classifies words and categories them (Pennebaker et al., 2015). The newest version of the app is the LIWC22 (LIWC, n.d.), which comes with an extensive manual. It comes with 100-dictionaires created to capture the social and psychological states of people. LIWC compares every word with the dictionaries and calculates the percentage of total words that match with the category of the dictionary. For example, there is a category of “cognitive processes” which is about the person actively processing through information, by using 1000 entries in the dictionary. This gives a clear overview of the different categories in the text and could therefore be a very informative and reliable analysis tool to use.

Practical implications

The conclusions of this study could be used by hospitals to predict and prevent PTSD early. According to He and colleagues (2017) it is critical to identify PTSD early in order to initiate an intervention quickly. Next to that, early identification could prevent personal suffering (Bo et al., 2021). The hospital can be attentive of the feeling of loneliness regarding isolation during admission. Predicting could be done by talking about the feelings of the person, especially regarding feeling of loneliness, absence of visitors, physical condition, implications and procedures. Next to that, the hospital could be attentive of the usage of negative formulated words and the usage of negative superlatives. Predicting could also be done by giving patients the opportunity to write about experiences and analyzing those questions on the results described above or by composing a questionnaire with subjects as

resulted above. However, there is a possibility that the causation of PTSD symptoms is a consequence of the sickness itself, which could make it hard to prevent the PTSD. Therefore, preventing sickness could be more valuable than preventing hospitalization. This means that hospitals could contribute more to the screening of PTSD early to treat and prevent illness from becoming chronic, debilitating, and difficult to treat (Todorov et al., 2018, as cited in He et al., 2019). According to Anxiety and Depression Association of America (2017) preventing PTSD symptoms could be done by demolishing avoidance behaviors, such as trying not to think or talk of the event. The hospital could demolish PTSD symptoms by actively talking about the illness, hospital admission period and the feelings of the patient. Nurses could contribute to demolish the feeling of loneliness by checking in on the patient more frequently. In addition, when professional help is needed, hospitals could give a referral for professional care to patients.

Final remark

In this study some qualitative patterns were found in how people hospitalized for COVID-19 narrate their experiences during and after hospitalization. These text patterns could indicate PTSD in the future. Screening patient-generated text data for words and topics, PTSD-related narration could be screened early and help to prevent development of PTSD symptomatology. While textual data has advantages for the respondents and researchers, text-mining analysis approaches should be further developed, to become easier to interpret by the researcher, and therefore become more reliable.

Literature

- Adams, R. E., & Boscarino, J. A. (2006). Predictors of PTSD and delayed PTSD after disaster: The impact of exposure and psychosocial resources. *Journal of Nervous and Mental Disease, 194*, 485–493. <http://dx.doi.org/10.1097/01.nmd.0000228503.95503.e9>
- Alvarez-Conrad, J., Zoellner, L. A., & Foa, E. B. (2001). Linguistic predictors of trauma pathology and physical health. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 15*(7), S159-S170. <http://doi.org/10.1002/acp.839>
- Anxiety and Depression association of America (2017). *How to prevent trauma from becoming PTSD*. <https://adaa.org/learn-from-us/from-the-experts/blog-posts/consumer/how-prevent-trauma-becoming-ptsd>
- Bo, H., Li, W., Yang, Y., Wang, Y., Zhang, Q., Cheung, T., Wu, X., & Xiang, Y. (2021). Posttraumatic stress symptoms and attitude toward crisis mental health services among clinically stable patients with COVID-19 in China. *Psychological Medicine, 51*(6), 1052-1053. <https://doi.org/10.1017/S0033291720000999>
- Boeschoten, M. A., Bakker, A., Jongedijk, R. A., & Olf, M. (2020). *PTSS checklist voor de DSM-5*. ARQ Nationaal Psychotrauma centrum.
- Boldenthusiast (2019, February 7). *Sentiment analysis – the lexicon based approach*. <https://www.alphabold.com/sentiment-analysis-the-lexicon-based-approach/>
- Chamberlain, S., Grant, J., Trender, W., Hellyer, P., & Hampshire, A. (2021). Post-traumatic stress disorder symptoms in COVID-19 survivors: Online population survey. *BJPsych Open, 7*(2), E47. <https://doi.org/10.1192/bjo.2021.3>
- D'Andrea, W., Chiu, P. H., Casas, B. R., & Deldin, P. (2012). Linguistic predictors of post-traumatic stress disorder symptoms following 11 September 2001. *Applied Cognitive Psychology, 26*(2), 316-323. <https://doi.org/10.1002/acp.1830>
- Ford, J. D., & Courtois, C. A. (2013). *Treating complex traumatic stress disorders in children and adolescents: Scientific foundations and therapeutic models*. Guilford Press.

- FormPlus (2018, February 19). Here's why more people are not completing your online surveys. *Formplus*. <https://blog.formpl.us/heres-why-more-people-are-not-completing-your-online-surveys-6898106f986e>
- Fox, R., Hyland, P., Coogan, A. N., Cloitre, M., & McHugh Power, J. (2022). Posttraumatic stress disorder, complex PTSD and subtypes of loneliness among older adults. *Journal of Clinical Psychology*, 78(2), 321-342. <https://doi.org/10.1002/jclp.23225>
- Franklin, C. L., & Thompson, K. E. (2005). Response Style and Posttraumatic Stress Disorder (PTSD): A Review. *Journal of Trauma & Dissociation*, 6(3), 105-123. https://doi.org/10.1300/J229v06n03_05
- Fu, L., Wang, B., Yuan, T., Chen, X., Ao, Y., Fitzpatrick, T., ... & Zou, H. (2020). Clinical characteristics of coronavirus disease 2019 (COVID-19) in China: a systematic review and meta-analysis. *Journal of Infection*, 80(6), 656-665. <https://doi.org/10.1016/j.jinf.2020.03.041>
- GGZ standaarden (2021). *Psychotrauma- en stressor gerelateerde stoornissen*. <https://www.ggzstandaarden.nl/zorgstandaarden/psychotrauma-en-stressorgerelateerde-stoornissen/introductie>
- Goldmann, E., & Galea, S. (2014). Mental health consequences of disasters. *Annual Review of Public Health*, 35, 169–183. <https://doi.org/10.1146/annurev-publhealth-032013-182435>
- Gray, M. J., & Lombardo, T. W. (2001). Complexity of trauma narratives as an index of fragmented memory in PTSD: a critical analysis. *Applied Cognitive Psychology*, 15(7), 171–186. <https://doi.org/10.1002/acp.840>
- Hawryluck, L., Gold, W. L., Robinson, S., Pogorski, S., Galea, S., & Styra, R. (2004). SARS control and psychological effects of quarantine, Toronto, Canada. *Emerging infectious diseases*, 10(7), 1206–1212. <https://doi.org/10.3201/eid1007.030703>
- He, Q., Veldkamp, B. P., & de Vries, T. (2012). Screening for posttraumatic stress disorder using verbal features in self narratives: A text mining approach. *Psychiatry research*, 198(3), 441-447. <https://doi.org/10.1016/j.psychres.2012.01.032>

- He, Q., Veldkamp, B. P., Glas, C. A. W., van den Berg, S. M. (2019). Combining tekst mining of long constructed responses and item-based measures: A hybrid test design to screen for Posttraumatic Stress Disorder. *Frontiers in Psychology, 10*:2358. <https://doi.org/10.3389/fpsyg.2019.02358>
- He, Q., Veldkamp, B. P., Glas, C. A., & de Vries, T. (2017). Automated assessment of patients' self-narratives for posttraumatic stress disorder screening using natural language processing and text mining. *Assessment, 24*(2), 157-172. <https://doi.org/10.1177/1073191115602551>
- Howes, C., Purver, M., & McCabe, R. (2014, June 27). Linguistic indicators of severity and progress in online text-based therapy for depression. In P. Resnik, R. Resnik, & M. Mitchell (Eds.), *Proceedings of the workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality* [Workshop]. Association for Computational Linguistics, Baltimore, Maryland, USA. <https://aclanthology.org/W14-32.pdf>
- Hussein, D. M. E. D. M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences, 30*(4), 330-338. <https://doi.org/10.1016/j.jksues.2016.04.002>
- Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media, 8*(1), 216-225. <https://ojs.aaai.org/index.php/ICWSM/article/view/14550>
- Imel, Z. E., Atkins, D. C., Steyvers, M. (2015). Computational psychotherapy research: Scaling up the evaluation of patient-provider interactions. *Journal of Psychotherapy, 52*(1), 19-30. <http://dx.doi.org/10.1037/a003684119>
- Itech (2021). *Which of the 3 algorithms models should you choose for sentiment analysis.* <https://itechindia.co/blog/which-of-the-3-algorithms-models-should-you-choose-for-sentiment-analysis-2/>
- Kang, L., Ma, S., Chen, M., Yang, J., Wang, Y., Li, R., Yao, L., Bai, H., Cai, Z., Xiang Yang, B., Hu, S., Zhang, K., Wang, G., Ma C., & Liu, Z. (2020). Impact on mental health and perceptions of psychological care among medical and nursing staff in Wuhan during the 2019 novel coronavirus disease outbreak: A cross-sectional study. *Brain, behavior, and immunity, 87*, 11-17. <https://doi.org/10.1016/j.bbi.2020.03.028>

- Kessler R.C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic Stress Disorder in the National Comorbidity Survey. *Arch Gen Psychiatry*, 52(12), 1048–1060.
<https://doi.org/10.1001/archpsyc.1995.03950240066012>
- Kleim, B., Horn, A. B., Kraehenmann, R., Mehl M. R., & Ehlers A. (2018) Early Linguistic Markers of Trauma-Specific Processing Predict Post-trauma Adjustment. *Frontiers in Psychiatry* 9:645.
<https://doi.org/10.3389/fpsy.2018.00645>
- Likao. (2019, January 3). *From a purely statistical standpoint, there is no reason at all why the variance within a group should be larger or smaller than between groups* [Comment on the online forum post *Are variances within a group always larger than between groups? When is this not true?*]. StackExchange. <https://stats.stackexchange.com/questions/385360/are-variances-within-a-group-always-larger-than-between-groups-when-is-this-not>
- Liu, D., Epskamp, S., Isvoranu, A. M., Chen, C., Liu, W., & Hong, X. (2021). Network analysis of physical and psychiatric symptoms of hospital discharged patients infected with COVID-19. *Journal of Affective Disorders*, 294, 707-713. <https://doi.org/10.1016/j.jad.2021.07.043>
- LIWC (n.d.). *How it works*. <https://www.liwc.app/help/howitworks>
- McNally, R. J., Robinaugh, D. J., Wu, G. W. Y., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clinical Psychological Science*, 3, 836-849. <https://doi.org/10.1177/2167702614553230>
- Monkeylearn (n.d.-a). *What is text mining? A beginner's guide*. <https://monkeylearn.com/text-mining/>
- Monkeylearn (n.d.-b). *Sentiment analysis: A Definitive Guide*. <https://monkeylearn.com/sentiment-analysis/>
- Monson, E., Caron, J., McCloskey, K., & Brunet, A. (2017). Longitudinal analysis of quality of life across the trauma spectrum. *Psychological Trauma: Theory, Research, Practice, and Policy*, 9(5), 605–612. <https://doi-org.ezproxy2.utwente.nl/10.1037/tra0000254>
- National center for PTSD (n.d.). *How is PTSD measured?*
https://www.ptsd.va.gov/understand/isitptsd/measured_how.asp
- Orange (n.d.). *Orange*. <https://orangedatamining.com/>

- Orsillo, S. M., Batten, S. V., Plumb, J. C., Luterek, J. A., Roessner, B. M., (2004). An experimental study of emotional responding in women with posttraumatic stress disorder related to interpersonal. *Journal of Traumatic Stress, 17*(3), 241–248.
<https://doi.org/10.1023/B:JOTS.0000029267.61240.94>
- Otis, J. D., McGlinchey, R., Vasterling, J. J., & Kerns, R. D. (2011). Complicating factors associated with mild traumatic brain injury: Impact on pain and posttraumatic stress disorder treatment. *Journal of Clinical Psychology in Medical Settings, 18*(2), 145-154. <https://doi-org.ezproxy2.utwente.nl/10.1007/s10880-011-9239-2>
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1-135. <http://dx.doi.org/10.1561/1500000011>
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. University of Texas.
- Porter, S. R., Whitcomb, M. E., & Weitzer, W. H. (2004). Multiple surveys of Students and Survey Fatigue. In S. R. Porter (Ed.), *Overcoming survey research problems* (pp. 63-73). Jossey-Bass.
- Qiu, J., Shen, B., Zhao, M., Wang, Z., Xie, B., & Xu, Y. (2020). A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations. *General psychiatry, 33*(2), 100-213. <https://doi.org/10.1136/gpsych-2020-100213>
- Rijksoverheid. (2022, March 6). *Ziekenhuisopnames*.
<https://coronadashboard.rijksoverheid.nl/landelijk/ziekenhuis-opnames>
- Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., & Roser, M. (2020). Coronavirus (COVID-19) Hospitalizations. *Our World in Data*. <https://ourworldindata.org/covid-hospitalizations#citation>
- Sawalha, J., Yousefnezhad, M., Shah, Z., Brown, M. R. G., Greenshaw, A. J., & Greiner, R. (2022). Detecting Presence of PTSD Using Sentiment Analysis From Text Data. *Frontiers Psychiatry, 12*, 811392. <https://doi.org/10.3389/fpsy.2021.811392>

- Schaap, C. P. D. R., Sanderman, R., & Ranchor, A. V. (2000). Handboek Psychopathologie, deel 1: Basisbegrippen. In P. M. G. Emmelkamp, W. Vandereycken, & C. A. L. Hoogduin (Eds.), *Handboek Psychopathologie* (pp. 242 - 243). Bohn, Stafleu, Van Loghum.
- Schnurr, P. P., & Green, B. L. (2004). Understanding relationships among trauma, posttraumatic stress disorder, and health outcomes. In P. P. Schnurr, & B. L. Green (Eds.). *Trauma and health: Physical health consequences of exposure to extreme stress* (pp. 217-243). American Psychological Association.
- Sheldrick, R., Tarrier, N., Berry, E., Kincey, J. (2006). Post-Traumatic Stress Disorder and illness perceptions over time following Myocardial Infarction and Sub Arachnoid Hemorrhage. *British Journal of Health Psychology, 11*, 387-400. <https://doi.org/10.1348/135910705X71434>
- Stanley, I. H., Hom, M. A., Spencer-Thomas, S., & Joiner, T.E. (2017). Examining anxiety sensitivity as a mediator of the association between PTSD symptoms and suicide risk among women firefighters. *Journal of Anxiety Disorder, 50*, 94-102. <https://doi.org/10.1037/pas0000254>
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics, 37*(2), 267-307. https://doi.org/10.1162/COLI_a_00049
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology, 29*(1), 24-54. <https://doi.org/10.1177/0261927X09351676>
- Tedstone, J. E., & Tarrier, N. (2003). Posttraumatic stress disorder following medical illness and treatment. *Clinical psychology review, 23*(3), 409-448. [https://doi.org/10.1016/S0272-7358\(03\)00031-X](https://doi.org/10.1016/S0272-7358(03)00031-X)
- Tingey, J. L., Bentley, J. A., & Hosey, M. M. (2020). COVID-19: Understanding and mitigating trauma in ICU survivors. *Psychological Trauma: Theory, Research, Practice, and Policy, 12*(S1), S100–S104. <https://doi.org/10.1037/tra0000884>
- Todorov, G. I., Mayilvahanan, K., Cain, C. K., & Cunha, C. (2018). *Screening word usage in people affected by PTSD: an unbiased, cost effective, and novel screening method?* New York University. <https://doi.org/10.31234/osf.io/y68fx>

- Trusko, B., Rosenbloom, S. T., Montella, D., Jackson, J. C., Fitzhenry, F., Brown, S. H., Elkin, P. L., Fielstein, E., Kotter, K., Tuttle, M., Lanneli, R. J., & Speroff, T. (2010). Are posttraumatic stress disorder mental health terms found in SNOMED-CT medical terminology? *Journal of Traumatic Stress, 23*(6), 794-801. <https://doi.org/10.1002/jts.20591>
- Vilchinsky, N., Ginzburg, K., Fait, K., & Foa, E. B. (2017). Cardiac-disease-induced PTSD (CDI-PTSD): a systematic review. *Clinical Psychology Review, 55*, 92-106. <https://doi.org/10.1016/j.cpr.2017.04.009>
- Vindegaard, N., & Benros, M. E. (2020). COVID-19 pandemic and mental health consequences: Systematic review of the current evidence. *Brain, behavior, and immunity, 89*, 531-542. <https://doi.org/10.1016/j.bbi.2020.05.048>
- WHO: World Health Organization (March, 12, 2019). Coronavirus disease 2019 (COVID-19) Situation Report-51. *WHO*. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10
- Xiang, Y. T., Yang, Y., Li, W., Zhang, L., Zhang, Q., Cheung, T., & Ng, C. H. (2020). Timely mental health care for the 2019 novel coronavirus outbreak is urgently needed. *The lancet psychiatry, 7*(3), 228-229. [https://doi.org/10.1016/S2215-0366\(20\)30046-8](https://doi.org/10.1016/S2215-0366(20)30046-8)
- Xiao, S., Luo, D., & Xiao, Y. (2020). Survivors of COVID-19 are at high risk of posttraumatic stress disorder. *Global health research and policy, 5*, 29. <https://doi.org/10.1186/s41256-020-00155-2>
- Yehuda, R. (2002). Post-traumatic Stress Disorder. *The New England Journal of Medicine, 346*, 108-114. <https://doi.org/10.1056/NEJMra012941>
- Zatzick, D. F., Marmar, C. R., Weiss, D. S., Browner, W. S., Metzler, T. J., Golding, J. M., Stewart, A., Schlenger, W. E., & Wells K. B. (1997). Posttraumatic stress disorder and functioning and quality of life outcomes in a nationally representative sample of male Vietnam veterans. *American Journal of Psychiatry, 154*(12), 1690-5. <https://doi.org/10.1176/ajp.154.12.1690>
- Zhang, J., Lu, H., Zeng, H., Zhang, S., Du, Q., Jiang, T., & Du, B. (2020). The differential psychological distress of populations affected by the COVID-19 pandemic. *Brain, behavior, and immunity, 87*, 49-50. <https://doi.org/10.1016/j.bbi.2020.04.031>

- Zhou, S., Wang, L., Yang, R., Yang, X., Zhang, L., Guo, Z., Chen, J., Wang, J., & Chen, J. (2020). Sleep problems among Chinese adolescents and young adults during the coronavirus-2019 pandemic. *Sleep medicine*, 74, 39-47. <https://doi.org/10.1016/j.sleep.2020.06.001>
- Zhou, T. T. (2020). Illness Experiences and Racial Identity in Sickness: a Memoi. *Cross-Cultural Communication*, 16(1), 93-95. <http://www.cscanada.net/index.php/cc/article/view/11585> DOI: <http://dx.doi.org/10.3968/11585>