Can Linguistic Analysis be used to Identify whether Adolescents with a Chronic Illness are Depressed?

Lauren Stephanie Jones[1,2], **Emma Anderson**[1], **Maria Loades**[1,2], **Rebecca Barnes**[1], **Esther Crawley**[1]

[1]Centre for Academic Child Health, Bristol Medical School, University of Bristol, UK

[2]Department of Psychology, University of Bath, UK

Abstract

Comorbid depression is common in adolescents with chronic illness. We aimed to design and test a linguistic coding scheme for identifying depression in adolescents with Chronic Fatigue Syndrome/Myalgic Encephalomyelitis (CFS/ME), by exploring features of e-consultations within online cognitive behavioural therapy treatment. E-consultations of 16 adolescents (aged 11 – 17) receiving FITNET-NHS treatment in a national randomised controlled trial were examined. A theoretically-driven linguistic coding scheme was developed and used to categorise comorbid depression in e-consultations using computerised content analysis. Linguistic coding scheme categorisation was subsequently compared to classification of depression using the Revised Children's Anxiety and Depression Scale (RCADS) published cut-offs (t-scores 65, 70). Extra linguistic elements identified deductively and inductively were compared with self-reported depressive symptoms after unblinding. The linguistic coding scheme categorised three (19%) of our sample consistently with self-report assessment. Of all 12 identified linguistic features, differences in language use by categorisation of self-report assessment were found for 'past-focus' words (mean rank frequencies: 1.50 for no depression, 5.50 for possible depression, and 10.70 for probable depression; $p < .05$) and 'discrepancy' words (mean rank frequencies: 16.00 for no depression, 11.20 for possible depression, and 6.40 for probable depression; $p < .05$). The linguistic coding profile developed as a potential tool to support clinicians in identifying comorbid depression in e-consultations showed poor value in this sample of adolescents with CFS/ME. Some promising linguistic features were identified, warranting further research with larger samples.

Introduction

Depression is a common mental health problem in adolescents (Merikangas, Nakamura, & Kessler, 2009), affecting 2.7% of 11 to 16 year olds in England (Vizard et al., 2018). It is characterised by depressed or irritable mood and/or loss of interest for a minimum two-week period, with symptoms resulting in significant distress or reduced functioning (American Psychological Association [APA], 2013).

Approximately 15% of adolescents experience a chronic illness (Van Der Lee, Mokkink, Grootenhuis, Heymans, & Offringa, 2007), such as asthma, diabetes, cancer, migraine, and chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME). Chronic illness poses risk for developing mood disorders (including depression) in adolescence (Pinquart & Shen, 2010; Bennett, Shafran, Coughtrey, Walker, & Heyman, 2015). Associated comorbid mood disorders should be identified during treatment for chronic illnesses (National Institute for Health and Care Excellence (NICE), 2011; NICE, 2017). Cognitive Behavioural Therapy (CBT) has a strong evidence-base for various paediatric psychiatric disorders, and there is growing interest in using CBT for treating somatic conditions (Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012).

Many adolescents with chronic conditions in the UK do not have access to face-to-face treatments such as CBT due to geographical barriers (Vigerland et al., 2016). Internet-delivered CBT with telephone support, video conferencing, or written email messages (e-consultation) has the potential to increase access to treatment for adolescents with chronic illness (Andersson, 2009; Vigerland et al., 2016). Low mood is generally highly visible in the traditional (face-to-face) clinical encounter, for example by appraising a patient's appearance, body language, and tone of speech (Beck, 1967; Hassan, McCabe, & Priebe, 2007). A significant proportion of communication is nonverbal (Mehrabian, 1972), which is lost in therapeutic e-consultations.

Naturally-occurring features of language use present a unique opportunity to assess an individual's psychological state (Pennebaker, 2011), and could be used for therapeutic interactions delivered by e-consultation. Based on Cartesian philosophical assumptions of the relationship between language and thought (Chomsky, 1966), simple words used in everyday speech can reflect underlying mental states (Pennebaker, Mehl, & Niederhoffer, 2003; Tausczik & Pennebaker, 2010; Weintraub, 1989). Linguistic analysis – in contrast to self-report assessment methods – has the advantage of tapping into implicit signals of low mood in adolescents with chronic illness (Hughes et al., 2016). Building upon Beck's (1967) cognitive model of depression and Leventhal's self-regulatory model (Leventhal, Diefenbach, & Leventhal, 1992), adolescents with chronic illness who are depressed might use more negative and catastrophising language in e-consultations than those who are not

depressed, reflecting distorted or unhelpful thoughts and extremely negative illness perceptions about their condition (see Table 1 and Table 2).

Existing linguistic analyses have investigated linguistic features that align with the theoretical assumption of negative styles of thinking in depression (Grant, 2010; Table 1). The literature is dominated by cross-sectional studies using a computerised content-analysis method (the Linguistic Inquiry and Word Count (LIWC; Pennebaker, Booth, & Francis, 2007)) to compare the specific word use of adults with and without depression in structured written tasks. The negative content of language and self-referent speech have received the most empirical attention, with several studies finding depression to be associated with greater use of negative emotion (e.g. "sad") words (Rude, Gortner, & Pennebaker, 2004; Tausczik & Pennebaker, 2010; Dirkse, Hadjistavropoulos, Hesser, & Barak, 2015), fewer positive emotion (e.g. "nice") words (Sloan, 2005), and greater use of first-person singular pronouns (e.g. "I"; Rude et al., 2004; Zimmerman, Wolf, Bock, Peham, & Benecke, 2013; Zimmerman, Brockmeyer, Hunn, Schauenburg, & Wolf, 2017; Holtzman, 2017). This may be because a person with depression is expected to have persistent negative thoughts that mainly centre around the self (Brockmeyer et el., 2015).

Frequent use of causation (e.g. "because") and insight (e.g. "think") words may further suggest a ruminative thinking style, whilst discrepancy words (e.g. "should") might be indicative of a (self-)critical inflexible thinking style (Grant, 2010). Using the LIWC, greater use of these cognitive mechanism words has been considered a marker of depression (Rodriguez et al., 2010), although this evidence is less conclusive (Warner et al., 2005; van der Zanden et al., 2014). Furthermore, consistent with the notion that depression is linked to "being stuck in the past" and difficulty in seeing a future (Holman & Silver, 1998; Habermas, Ott, Schubert, Schneider, & Pate, 2008), depression has been associated with an elevated use of the past tense (e.g. "did"; Rodriguez et al., 2010). Finally, a recent study examining an all-or-nothing thinking style found that depression was associated with greater use of absolutist words (e.g. "always"; Al-Mosaiwi & Johnstone, 2018).

There have been no previous attempts to create a linguistic coding profile of depressive symptoms. Furthermore, it is unknown whether linguistic analyses are useful as a diagnostic tool in paediatric populations (Cornaggia et al., 2016), and more specifically for identifying whether adolescents with chronic illness are depressed using therapeutic e-consultations. Of the few studies using the LIWC in adolescent chronic illness, the focus has been on linguistic markers of change in cognitive processing (i.e. causation and insight words) during interventions of expressive writing (Warner et al., 2005; Gillis, Lumley, Mosley-Williams, Leisen, & Roehrs, 2006). We are aware of only two studies examining patient language use in e-consultations during internet-delivered CBT that have analysed how specific words are related to affective disorders (Dirkse et al., 2015; Van der Zanden et al., 2014). Both studies were conducted in adults with primary anxiety or depressive symptoms, and were interested in patients' word use in relation to treatment adherence and outcome. In the current study, the focus is on early identification of comorbid depressive symptoms rather than psychotherapeutic change over time.

FITNET-NHS, a UK adaptation of the Dutch FITNET (Nijhof, Bleijenberg, Uiterwaal, Kimpen, & van de Putte, 2012), is currently being tested within a national randomised controlled trial (RCT) in the UK (Baos et al., 2018). The full trial is investigating whether internet-delivered CBT (FITNET-NHS) is an effective and cost-effective treatment for paediatric CFS in the NHS, compared to Activity Management (delivered via Skype). The integral e-consultation component of FITNET-NHS treatment provides a unique opportunity for the current study to explore the identification of comorbid depressive symptoms, of which are common in paediatric CFS (Loades, Rimes, Ali, Lievesley, & Chalder, 2017), from patient language use.

The current study aimed to develop a theory-based linguistic coding profile of depressive symptoms in adolescents with chronic illness by studying the e-consultations of adolescents with CFS/ME, and provide methodological recommendations to pursue in future work. The research question is: can we create a diagnostically useful linguistic coding tool to identify comorbid depression from e-consultation messages of adolescents with CFS/ME in the early stages of internet-delivered CBT?

Method

Study design and setting

This research is nested within a national UK RCT of internet-delivered CBT for CFS/ME (Baos et al., 2018), which is a UK adaptation of the Dutch trial (Nijhof et al., 2012). The trial will examine treatment effects in the subgroup of adolescents with CFS/ME with comorbid mood disorders. A specialist paediatric CFS/ME service delivers the online CBT treatment within the main trial, and adolescents receive the treatment at home via the Internet. Participants are supported through the CBT program with one-to-one therapeutic e-consultations.

In the present study, we analysed the first four email messages sent by adolescents with CFS/ME to a specialist paediatric CFS/ME psychologist within the trial (see Figure 1).

Participants

Adolescents referred to the specialist paediatric CFS/ME service by their GP were eligible for the full trial if they were aged 11 to 17 years, had CFS/ME as defined by NICE (2007) guidance, had no access to a local specialist service, CFS/ME was the main presenting problem to be treated, and any mood disorder was not a cause of the fatigue. Email e-consultations of participants in the full trial were analysed if they were randomly allocated to online CBT treatment, had given consent for the analysis of their e-consultations, and had started the treatment and engaged in the e-consultations (see Figure 2).

Measures

Depression—At eligibility assessment for the full trial, participants completed the Revised Children's Anxiety and Depression Scale (RCADS; Chorpita, Yim, Moffitt, Umemoto, & Francis, 2000), a validated measure of paediatric depressive symptoms (Chorpita, Moffitt, & Gray, 2005). The RCADS depression sub-scale consists of 10 items such as "I feel sad or

empty", and participants answer using a 4-point rating scale where 0 = Never and 3 = Always. Higher scores indicate more depressive symptoms. A t-score (i.e. age- and gender-adjusted clinical thresholds) of 65 is described as possible depression and a t-score of 70 as probable depression.

Other patient-reported measures—At baseline assessment, participants completed validated measures of fatigue (Chalder Fatigue scale; Chalder et al., 1993; Morriss, Wearden, & Mullis, 1998) and pain (Pain Visual Analogue scale; Hawker, Mian, Kendzerska, & French, 2011), and reported their length of illness, symptoms (based on NICE (2007) criteria), and typical school attendance.

Procedure

Data collection—Data were collected from eligible patients within the online CBT treatment arm of the trial from November 2017 to June 2018. The e-consultation data was extracted from the trial platform on 01/06/2018. Potentially identifying information mentioned in the e-consultations, such as names of people and places, were pseudoanonymised (replaced with artificial identifiers) prior to analysis.

Generation of Linguistic Coding Scheme—We searched PubMed and Google Scholar using the search terms: "linguistic analysis", "content analysis", "LIWC", "word use", "depression", "CFS/ME", "written", and "differential diagnosis". We identified nine linguistic features to potentially include in the linguistic coding scheme based on the published literature of language use and depression: negative emotion, positive emotion, first-person pronouns, causation, insight, discrepancy, past-focus, future-focus, and absolutist words (Rude, et al., 2004; Molendijk et al., 2010; Arntz et al., 2012; Zimmermann et al., 2017; Al-Mosaiwi & Johnstone, 2018). Furthermore, the e-consultations were read to identify linguistic features that may be unique to the clinical setting (e-consultations) and patient population (paediatric chronic illness). We identified one linguistic feature to potentially include in the linguistic coding scheme: achieve words (e.g. "goal").

The initial hypothesised list of linguistic features identified deductively and inductively was refined through consultation with specialist paediatric CFS/ME psychologists and consideration of the literature, which lead to the exclusion of six linguistic features: causation, insight, discrepancy, past-focus, future focus, and achieve words. These features were excluded for (at least) one of the following reasons: potentially affected by the study context, potentially reflective of both CFS/ME and depression, a lack of previous evidence or a less conclusive evidence-base, and available linguistic data for personality disorder only (see Table 3).

For instance, it was indicated through the consultation process that an adolescent might be less reflective in the current study, and consequently use fewer causation words (e.g. "because"), due to two key aspects of the study context: i) the therapeutic interactions are task-focussed, such that an adolescent is more likely to use specific language about their progress with the online CBT chapters, and ii) the adolescent has never seen their therapist face-to-face, which might impact on their willingness to reflect and share in their written

email messages to their therapist. In turn, causation words were excluded from the current coding scheme.

This resulted in a focused linguistic coding scheme (see Appendix A) using four linguistic features with a theoretical basis and empirical evidence with depressed individuals: negative emotion, positive emotion, and first-person singular pronouns (Rude et al., 2004; Zimmerman et al., 2017; Molendjik et al., 2010; Arntz, Hawke, Bamelis, Spinhoven, & Molendijk, 2012), and absolutist (Al-Mosaiwi & Johnstone, 2018) words.

The theoretically-driven coding scheme was developed to characterise the frequencies of each linguistic feature as "0" = non-case, "1" = borderline case, and "2" = case, matching the RCADS categorisation (no/possible/probable depression) with similar but distinctive names for clarity. The cut-off frequencies we used to define these categories (non-case/ borderline case/case) were based on frequency scores from the previous literature as follows:

To form discrete categories (non-case/borderline case/case) for each linguistic feature, the strongest evidence in the literature was prioritised (see Appendix A). We assessed the strength of previous evidence and its relevance to the current study, including closeness of sample characteristics (age, 11-17 years; mental health, depression) and modality of language (e-consultation or written language). Based on this assessment, we selected Rude et al.'s (2004) study to inform three linguistic features (negative emotion, positive emotion, and first-person singular pronouns) and Al-Mosaiwi and Johnstone's (2018) study to inform the fourth linguistic feature (absolutist) on the linguistic coding scheme. Specifically, we constructed non-case and case category boundaries based on the mean frequencies reported for non-depressed and depressed groups by these studies, with intermediate frequencies categorised as borderline.

Summary scores for all four linguistic features had a possible range of 0 to 8. By dividing this range into three equal categories – with no previous evidence to justify alternative category boundaries – we categorised these scores as: non-case 0-2, borderline case 3-5, and case 6-8. Again, the borderline category was included to mirror the RCADS categorisation.

Analysis—The LIWC2015 (Pennebaker, Boyd, Jordan, & Blackburn, 2015), a computerised approach of content analysis (Krippendorff, 1980), was used to analyse the e-consultations, a method chosen for its ability to examine the linguistic style of written patient communication. The e-consultations were analysed whilst blinded to the participant's depressive symptoms as measured by the RCADS. The LIWC program calculates the prevalence of words for a given dictionary word category as a percentage of the total number of words analysed. The LIWC2015 master dictionary was run in the current analysis, which is composed of 90 word-categories. A user-created dictionary was added to the program, namely the absolutist dictionary developed by Al-Mosaiwi and Johnstone (2018).

Primary outcome: To examine the match between the hypothetical comorbid depression status of participants generated by the linguistic coding scheme and that identified via self-report depression assessment, we used four word categories from the LIWC2015 and absolutist dictionaries (see Generation of Linguistic Coding Scheme): negative emotion,

positive emotion, first-person singular pronouns, and absolutist. The linguistic coding scheme was applied to the LIWC output for the first four messages of each participant, giving each participant a total score and generating a hypothetical comorbid depression status (non-case/borderline case/case). For a sensitivity analysis, the coding scheme was subsequently applied to the LIWC output for all messages sent by each participant. On completion of coding the LIWC output for each participant using the linguistic coding scheme, the research team were unblinded to participants' self-reported depression scores (RCADS).

Secondary outcomes: To explore differences in language use by participants' comorbid status according to self-report depression assessment, we used 11 linguistic features identified deductively as relevant to depression and/or chronic illness and one linguistic feature identified inductively:

- Negative emotion

- Positive emotion

- First-person singular pronouns

- Absolutist

- Causation

- Insight

- Discrepancy

- Past-focus

- Future-focus

- Health

- Social Processes

- Achieve

Statistical Analysis

The hypothetical depression status of participants generated by the linguistic coding scheme in the primary and sensitivity analyses was compared to that identified via RCADS baseline assessment (no/possible/probable depression), using a percentage match.

Baseline demographic and clinical data of the sample were compared with the 'main trial' sample (which excluded participants of the current sample), using Mann-Whitney U tests for continuous variables and Pearson's Chi-Square tests (or Fisher's exact test when expected values were below five) for categorical variables.

The Kruskal-Wallis test was used to examine differences in use of linguistic features by participants' comorbid status according to RCADS assessment (no/possible/probable depression). Pairwise comparisons were used for follow-up analyses, for which effect sizes (*r*) were calculated. An alpha value of .05 was used for all statistical analyses.

Ethical Approval

The study received full ethical approval from a University Research Ethics Committee (REC) in March 2018 (*REC ref*). The trial received ethical approval from the NHS Health Research Authority in November 2016 (REC ref) and for the amendment to analyse the email consultations in October 2017 (ref, Substantial Amendment 1).

Results

Participant Characteristics

Table 4 describes our sample of 16 participants. Twelve were female, and the age range for the overall sample was 12 to 15 years ($M = 13.69$, $SD = 1.01$). On average, our sample of 16 participants experienced eight of nine CFS symptoms ($M = 7.88$, $SD = 1.36$). Participants were quite different from the main trial sample, in that they had a more recent onset of CFS symptoms (mean rank 46.25 vs 66.56, $p = .039$), higher fatigue scores (mean rank 83.72 vs 59.96, $p = .014$), more moderate-severe pain symptoms (94% vs 57%, $p = .004$), and a higher prevalence of depression defined using the RCADS with probable depression represented in 63% of current study sample versus 44% in main trial sample, and possible depression represented in 31% of current versus 9% of main trial sample (see Table 5).

Fifteen of our sample of 16 participants had four e-consultation messages. One participant had only two e-consultation messages because the participant's parent intervened in writing the emails for the adolescent.

Primary Outcome: Match between Linguistic Coding Scheme and RCADS

Table 6 shows the linguistic dimension for each participant. According to the linguistic coding scheme, one (6%) participant was categorised as a case for comorbid depression, and five (31%) were categorised as a borderline case. Ten (63%) participants were categorised as a non-case.

Table 7 shows the percentage match between the linguistic coding scheme and the RCADS in the primary analysis. Our linguistic coding scheme developed as a potential tool for identifying comorbid depression categorised three (19%) of our sample of 16 (adolescents with CFS/ME) consistently with the RCADS. Of the ten adolescents with probable comorbid clinical depression as categorised by the RCADS, our linguistic coding scheme categorised one (10%) as a case. Our coding scheme categorised over half (N = 9; 60%) of the 15 adolescents who had possible/probable comorbid depressive symptoms according to the RCADS as a non-case.

Sensitivity Analysis

In the sensitivity analysis, the coding scheme was applied to the LIWC output for all messages sent by each participant (i.e. the analysis was not limited to the first four messages). Table 7 shows the percentage match between the linguistic coding scheme and the RCADS. Including all messages, our linguistic coding scheme categorised two (13%) adolescents in accordance with RCADS categories, thus reducing the match between the coding scheme and RCADS.

Differences in Linguistic Dimensions by RCADS Category

Table 8 shows the mean frequencies for the 12 identified LIWC categories and word count for the RCADS groups (no/possible/probable depression).

A Kruskal-Wallis test showed that there was a difference in the use of past-focus words (e.g. "did") between RCADS groups, $H(2) = 6.292$, $p = .043$, with a mean rank frequency of 1.50 for no depression, 5.50 for possible depression, and 10.70 for probable depression. Pairwise comparisons, with a Bonferonni adjustment ($p = .0167$), showed no significant differences in the use of past-focus words between no depression and possible depression RCADS groups ($p = 1.00$, $r = -.31$) or probable depression ($p = .196$, $r = -.56$). There was also no significant difference between the possible and probable depression RCADS groups ($p = .138$, $r = -.52$).

There was a difference in the use of discrepancy words (e.g. "should") between RCADS groups, $H(2) = 6.044$, $p = .049$, with a mean rank frequency of 16.00 for no depression, 11.20 for possible depression, and 6.40 for probable depression. Pairwise comparisons, with a Bonferonni adjustment ($p = .0167$), showed no significant differences in the use of discrepancy words between no depression and possible depression RCADS groups ($p = 1.00$, $r = .38$) or probable depression ($p = .163$, $r = .58$). There was also no significant difference between the possible and probable depression RCADS groups ($p = .196$, $r = .48$).

Discussion

This was a first attempt to use a novel coding scheme we developed to identify possible/probable depression in young people with CFS/ME by analysing the linguistic content of their e-consultation emails. The depression category results generated by the coding scheme did not match the possible/probably depression categories identified by RCADS self-assessment in this group of patients

We created the coding scheme drawing on the best theory and empirical evidence available, which led us to the four linguistic features included (negative emotion, positive emotion, first-person singular pronouns, absolutist) as related to depression (e.g. Rude et al., 2004; Al-Mosaiwi & Johnstone, 2018). It may be that the evidence is currently insufficient and more empirical evidence is needed to inform and refine a future coding scheme before it can have predictive validity.

One explanation for the under-identification of depressive symptoms is that the linguistic coding scheme – informed by previous findings with adult populations (Rude et al., 2004; Al-Mosaiwi & Johnstone, 2018; Zimmerman et al., 2017; Molendjik et al., 2010; Arntz et al., 2012) – was not appropriate for the adolescent sample (Cornaggia et al., 2016). Although the core symptoms of depression are similar for adolescents and adults (APA, 1980), previously-identified linguistic features might not be useful or sufficient for identifying whether adolescents with chronic illness are depressed. In comparison to Al-Mosaiwi and Johnstone's (2018) study, the current sample used fewer absolutist words. Greater flexibility and less rigidity in thinking styles might be expected in adolescence compared to adulthood due to brain plasticity peaking at this developmental stage (Dahl, 2004; Tamnes et al., 2017), and this cognitive process might explain differences in communication patterns (Bell &

Condren, 2016; Chomsky, 1966). Further linguistic analyses in adolescent and chronic illness populations may be needed to inform a future iteration of a coding scheme.

Another explanation for the under-identification of depressive symptoms by the linguistic coding scheme is the study context. Although the coding scheme prioritised previous studies that examined written language (i.e. written essays and online forums), the unique context of a clinical e-consultation might have affected the appropriateness of the category boundaries. In comparison to previous studies (Rude et al., 2004; Zimmerman et al., 2017; Molendjik et al., 2010), more positive emotion words were used by adolescents in their e-consultations. In the trial context, there is the possibility of social desirability bias, whereby participants may respond more positively in their answers to their therapists. Future research could use a sensitivity analysis to explore whether the coding scheme is improved by the removal of positive emotion words.

Limitations

It is possible that the sampling affected the main findings. The proportion of adolescents with CFS/ME in the current sample categorised as having possible/probable comorbid depression by self-report assessment was much higher than we anticipated based on previous literature (Garralda & Rangel, 2005; Loades et al., 2017; Bould et al., 2013). Our sample also had greater levels of fatigue and pain compared to the wider cohort of the trial. It is difficult to draw conclusions from this unrepresentative sample.

In addition to the sampling limitation, several further methodological limitations of the current study merit consideration. First, this study examined a small dataset. Given that more data (particularly from those with no depressive symptoms) may be needed for refining the coding scheme, and then more testing to be useful, it would be desirable to examine the identified linguistic features with larger samples. This exploratory study was as inclusive as possible. However, including participants with fewer words than what is recommended (> 50 words) in their early e-consultation messages may have compromised the accuracy of analysis. The focus on the first four e-consultation messages was important for applying findings to early identification of depression in real-life therapeutic e-consultations. The sensitivity analysis increases our confidence that our findings are not explained by chance because the pattern of results found when all the messages (and more words) were analysed showed consistency with the results of the primary analysis. We further addressed study reliability by using a blinded coding procedure.

Second, the high proportion of adolescents categorised as having possible/probable depression by self-report assessment (RCADS) indicates potential for a reliability issue in the RCADS depression scale for paediatric CFS/ME. We know from several different studies using different assessment methods (Garralda & Rangel, 2005; Loades et al., 2017; Bould et al., 2013), that the estimated prevalence of depression is around 30%, yet in our small sample it was 63% (probable depression). Importantly, we do not yet know if the thresholds on the RCADS to identify possible/probable depression in healthy samples are also the optimum thresholds to apply to adolescents with chronic illness like CFS/ME to maximise sensitivity and specificity (Thase, 1991; Larkin & Martin, 2017). It is possible that CFS/ME symptoms and the RCADS assessment categories overlap which then would inflate the

incidence of possible/probably depression in a CFS/ME sample using this measure. However, the RCADS has strong psychometric properties and has shown greater correspondence to depression as a diagnostic disorder compared to traditional measures in otherwise healthy populations (Chorpita et al., 2005).

Third, the assumption that words provide insight into the cognitive processes associated with particular affective disorders has not gone unchallenged (Carley, 1990; Pennebaker & King, 1999; Pennebaker et al., 2003). A transdiagnostic approach to mental health highlights that the same core psychological processes (e.g. the repetitive occurrence of negative thoughts) underlie various common mental health problems (Kaplan et al., 2018). Some linguistic features are not specific to depression, but rather are equally related to other mental health groups in comparison to controls, including anxiety, suicidal ideation, borderline personality disorder, and eating disorder (Al-Mosaiwi & Johnstone, 2018). It is therefore possible that the results of applying the linguistic coding scheme are influenced by a range of different problems, in addition to depression, thereby reducing the specificity of the tool.

Fourth, word-count strategies are unable to consider context or nuances of language (Pennebaker & King, 1999).Furthermore, in this study we only analysed word categories from predetermined dictionaries, which constrained our exploration of linguistic features that are theoretically relevant to depression. Both the self-regulatory model (Leventhal, Diefenbach, & Leventhal, 1992) and the cognitive model of depression (Beck, 1967) include catastrophic thinking, with negative and catastrophic illness beliefs and thinking patterns hypothesised to be related to strong emotion in adolescents with CFS/ME (Gray & Rutter, 2007). References to third parties (e.g. a family member or friend) to catastrophise an illness experience (e.g. "my teacher was really worried about me… I'll never get better") have been explored in neurological patient populations (Robson, Drew, Walker, & Reuber, 2012). However, a word-count analysis of third-party references would not consider the context with which third-party references are used – specifically, whether the patient is catastrophising or normalising their experience. Nevertheless, the LIWC was an efficient process and is currently the most widely used computerised approach to linguistic analysis (Tausczik & Pennebaker, 2010).It is possible that e-consultations could be analysed by a coding scheme grounded in conversation analysis, in order to characterise contextual features (e.g. the patient-therapist interaction), in addition to the linguistic style (Stivers, 2015). The way adolescents respond to the generally long narrative and questions of therapists in e-consultations and the extent to which they work collaboratively as an active agent in internet-delivered CBT could provide useful insights into mood. For example, response time to a therapist's questions and spontaneous volunteering of information, as investigated in neurological patient populations (e.g. Jones et al., 2016; Robson et al., 2012; Plug, Sharrack, & Reuber, 2009), could be useful for identifying comorbid depression in paediatric chronic illness.

Future Directions

At present, there is insufficient evidence to create a useful linguistic coding scheme that could support clinicians in identifying whether adolescents with chronic illness are depressed during e-consultations. The current research points to the following four areas to

test in future: i) gain a more robust idea of the language used by adolescents with and without comorbid depression by exploring language use in larger samples, in different chronic illness populations, and different language modalities; ii) ii) further test the four theory-driven word categories within the current coding scheme and the other hypothesised elements that were excluded from the current coding scheme; iii) go beyond the words that an adolescent uses, to also consider whether the way that an adolescent interacts with their therapist during e-consultations can help us to identify low mood; and iv) for simplicity and greater clinical relevance, compare the depression status (depressed/not depressed) generated by a linguistic coding scheme with classification of depression using a diagnostic interview, such as the gold standard Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS; Kaufman et al., 1997).

Implications

Clinicians should be aware that, at the moment, it is very difficult to identify co-morbid depression in the early stages of e-consultations. This is a concern because adolescents with chronic illness might be cautious about disclosing low mood, particularly for any adolescents who have perceived scepticism and experienced lack of validation for a condition of unknown aetiology such as CFS/ME (Jelbert et al., 2010; Hareide, Finset, & Wyller, 2011). It is possible that this could result in missed or delayed identification of low mood during internet-delivered CBT. Clinicians should utilise multiple sources of information, including self-report inventories shown to be valid in the adolescent age group, and also information from informants such as parents and teachers. Clinicians should have a low threshold for face-to-face assessment if they are concerned about an adolescents' mood.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Al-Mosaiwi M, Johnstone T. In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. Clinical Psychological Science. 2018

Andersson G. Using the Internet to provide cognitive behaviour therapy. Behaviour research and therapy. 2009; 47:175–180. [PubMed: 19230862]

Arntz A, Hawke LD, Bamelis L, Spinhoven P, Molendijk ML. Changes in natural language use as an indicator of psychotherapeutic change in personality disorders. Behaviour Research and Therapy. 2012; 50:191–202. [PubMed: 22317755]

Baos S, Brigden A, Anderson E, Hollingworth W, Price S, Mills N, et al. Crawley E. Investigating the effectiveness and cost-effectiveness of FITNET-NHS (Fatigue In Teenagers on the interNET in the NHS) compared to Activity Management to treat paediatric chronic fatigue syndrome (CFS/ME)/

myalgic encephalomyelitis (ME): protocol for a randomised controlled trial. Trials. 2018; 19:136. [PubMed: 29471861]

Beck, AT. Depression: Clinical, experimental, and theoretical aspects. University of Pennsylvania Press; 1967.

Bell J, Condren M. Communication Strategies for Empowering and Protecting Children. The Journal of Pediatric Pharmacology and Therapeutics: JPPT. 2016; 21:176–184. [PubMed: 27199626]

Bennett S, Shafran R, Coughtrey A, Walker S, Heyman I. Psychological interventions for mental health disorders in children with chronic physical illness: a systematic review. Archives of disease in childhood. 2015; 100:308–316.

Bould H, Collin SM, Lewis G, Rimes K, Crawley E. Depression in paediatric chronic fatigue syndrome. Archives of Disease in Childhood. 2013; 98:425–428. [PubMed: 23619200]

Brockmeyer T, Zimmermann J, Kulessa D, Hautzinger M, Bents H, Friederich HC, et al. Backenstrass M. Me, myself, and I: self-referent word use as an indicator of self-focused attention in relation to depression and anxiety. Frontiers in Psychology. 2015; 6:1564. [PubMed: 26500601]

Carley, K. Content analysis. The encyclopedia of language and linguistics. Pergamon Press; Edinburgh: 1990. 725–730.

Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, Wright D, Wallace EP. Development of a fatigue scale. Journal of Psychosomatic Research. 1993; 37:147–153. [PubMed: 8463991]

Chomsky, N. Cartesian Linguistics: Chapter in the Historz of Rationalist Thought. Harper and Row; 1966.

Chorpita BF, Moffitt CE, Gray J. Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. Behaviour Research and Therapy. 2005; 43:309–322. [PubMed: 15680928]

Chorpita BF, Yim L, Moffitt C, Umemoto LA, Francis SE. Assessment of symptoms of DSM-IV anxiety and depression in children: A revised child anxiety and depression scale. Behaviour Research and Therapy. 2000; 38:835–855. [PubMed: 10937431]

Cornaggia CM, Di Rosa G, Polita M, Magaudda A, Perin C, Beghi M. Conversation analysis in the differentiation of psychogenic nonepileptic and epileptic seizures in pediatric and adolescent settings. Epilepsy & Behavior. 2016; 62:231–238. [PubMed: 27494360]

Dahl RE. Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. Annals of the New York Academy of Sciences. 2004; 1021:1–22. [PubMed: 15251869]

Dirkse D, Hadjistavropoulos HD, Hesser H, Barak A. Linguistic analysis of communication in therapist-assisted internet-delivered cognitive behavior therapy for generalized anxiety disorder. Cognitive Behaviour Therapy. 2015; 44:21–32. [PubMed: 25244051]

Garralda ME, Rangel L. Chronic fatigue syndrome of childhood. Comparative study with emotional disorders. European Child & Adolescent Psychiatry. 2005; 14:424–430. [PubMed: 16341498]

Gillis ME, Lumley MA, Mosley-Williams A, Leisen JC, Roehrs T. The health effects of at-home written emotional disclosure in fibromyalgia: A randomized trial. Annals of Behavioral Medicine. 2006; 32:135–146. [PubMed: 16972811]

Grant, A. Cognitive behavioural interventions for mental health practitioners. Learning Matters: 2010.

Gray SE, Rutter DR. Illness representations in young people with Chronic Fatigue Syndrome. Psychology and Health. 2007; 22:159–174.

Habermas T, Ott LM, Schubert M, Schneider B, Pate A. Stuck in the past: Negative bias, explanatory style, temporal order, and evaluative perspectives in life narratives of clinically depressed individuals. Depression and Anxiety. 2008; 25

Hareide L, Finset A, Wyller VB. Chronic fatigue syndrome: a qualitative investigation of young patient's beliefs and coping strategies. Disability and rehabilitation. 2011; 33:2255–2263. [PubMed: 21473686]

Hassan I, McCabe R, Priebe S. Professional-patient communication in the treatment of mental illness: a review. Communication & medicine. 2007; 4:141–152. [PubMed: 18052814]

Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual analog scale for pain (vas pain), numeric rating scale for pain (nrs pain), mcgill pain questionnaire (mpq), short-form mcgill pain questionnaire (sf-mpq), chronic pain grade scale (cpgs), short form-36 bodily pain
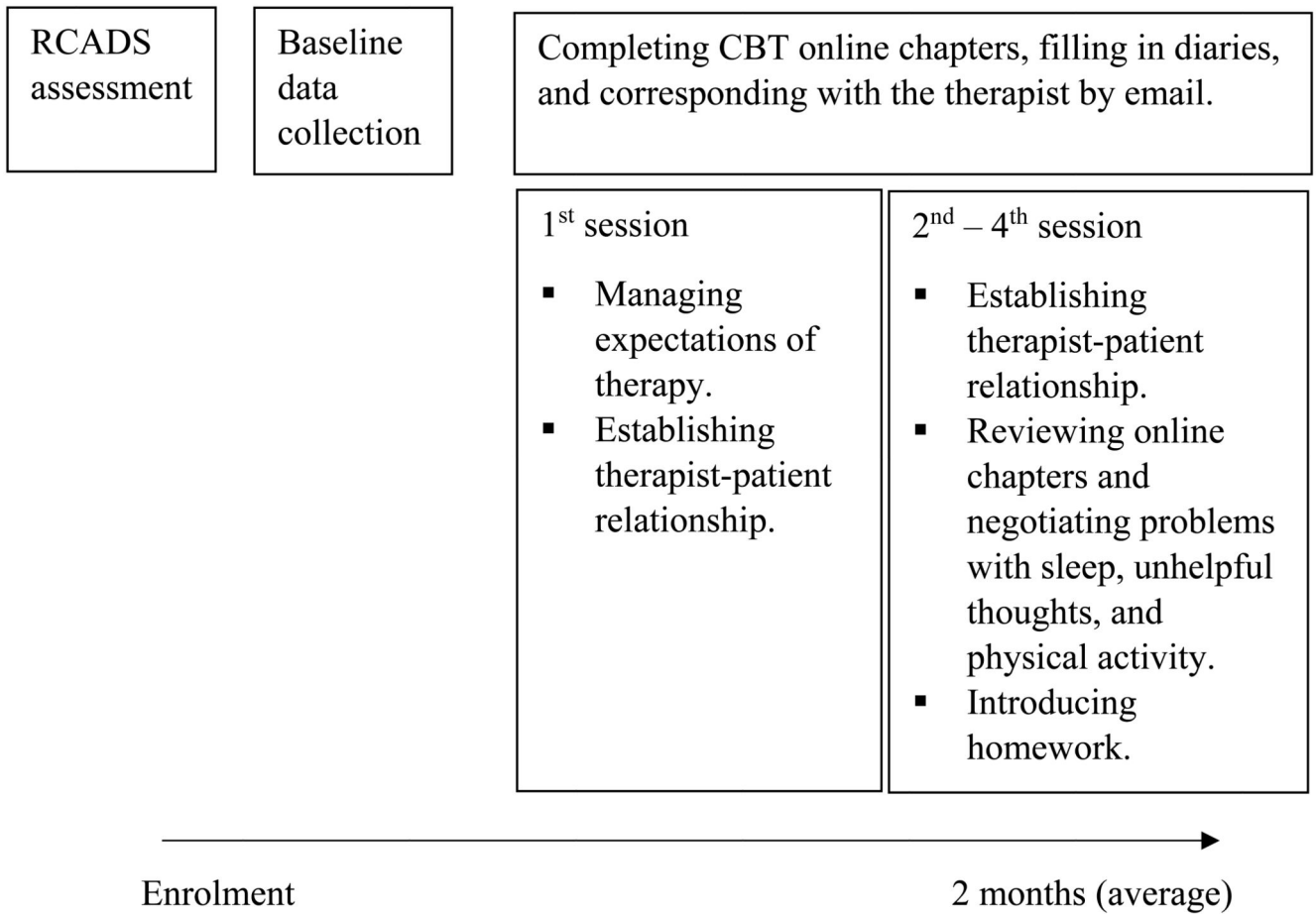
scale (sf-36 bps), and measure of intermittent and constant osteoarthritis pain (icoap). Arthritis Care & Research. 2011; 63:S240–S252. [PubMed: 22588748]

Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: A review of meta-analyses. Cognitive Therapy and Research. 2012; 36:427–440. [PubMed: 23459093]

Holman EA, Silver RC. Getting" stuck" in the past: temporal orientation and coping with trauma. Journal of Personality and Social Psychology. 1998; 74:1146. [PubMed: 9599436]

Holtzman NS. A meta-analysis of correlations between depression and first person singular pronoun use. Journal of Research in Personality. 2017; 68:63–68.

Hughes A, Hirsch C, Chalder T, Moss-Morris R. Attentional and interpretive bias towards illness-related information in chronic fatigue syndrome: A systematic review. British Journal of Health Psychology. 2016; 21:741–763. [PubMed: 27329758]

Jelbert R, Stedmon J, Stephens A. A qualitative exploration of adolescents' experiences of chronic fatigue syndrome. Clinical Child Psychology and Psychiatry. 2010; 15:267–283. [PubMed: 20179018]

Jones D, Drew P, Elsey C, Blackburn D, Wakefield S, Harkness K, Reuber M. Conversational assessment in memory clinic encounters: interactional profiling for differentiating dementia from functional memory disorders. Aging & Mental Health. 2016; 20:500–509. [PubMed: 25803169]

Kaplan DM, Palitsky R, Carey AL, Crane TE, Havens CM, Medrano MR, et al. O'Connor MF. Maladaptive repetitive thought as a transdiagnostic phenomenon and treatment target: An integrative review. Journal of Clinical Psychology. 2018

Kaufman J, Birmaher B, Brent D, Rao UMA, Flynn C, Moreci P, et al. Ryan N. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. Journal of the American Academy of Child & Adolescent Psychiatry. 1997; 36:980–988. [PubMed: 9204677]

Krippendorf, K. Content analysis: An introduction to its methodology. Beverly Hills, CA: Sage Publications; 1980.

Larkin D, Martin CR. The interface between chronic fatigue syndrome and depression: A psychobiological and neurophysiological conundrum. Neurophysiologie Clinique/Clinical Neurophysiology. 2017; 47:123–129. [PubMed: 28314518]

Leventhal H, Diefenbach M, Leventhal EA. Illness cognition: using common sense to understand treatment adherence and affect cognition interactions. Cognitive Therapy and Research. 1992; 16:143–163.

Leventhal H, Leventhal EA, Contrada RJ. Self-regulation, health, and behavior: A perceptual-cognitive approach. Psychology and Health. 1998; 13:717–733.

Loades ME, Rimes KA, Ali S, Lievesley K, Chalder T. The presence of co-morbid mental health problems in a cohort of adolescents with chronic fatigue syndrome. Clinical Child Psychology and Psychiatry. 2017

Mehrabian, A. Nonverbal Communication. New Brunswick: Aldine Transaction; 1972.

Molendijk ML, Bamelis L, van Emmerik AA, Arntz A, Haringsma R, Spinhoven P. Word use of outpatients with a personality disorder and concurrent or previous major depressive disorder. Behaviour Research and Therapy. 2010; 48:44–51. [PubMed: 19819423]

Morriss R, Wearden A, Mullis R. Exploring the validity of the Chalder Fatigue scale in chronic fatigue syndrome. Journal of Psychosomatic Research. 1998; 45:411–417. [PubMed: 9835234]

National Institute for Health and Care Excellence (NICE). Common mental health problems: identification and pathways to care (NICE guidelines CG123). 2011. Retrieved February 1, 2019, from: https://www.nice.org.uk/guidance/cg123/chapter/1-guidance#step-1-identification-and-assessment

National Institute for Health and Care Excellence (NICE). Depression in children and young people: identification and management (NICE guidelines CG28). 2017. Retrieved August 24, 2018, from: https://www.nice.org.uk/guidance/cg28

Nijhof SL, Bleijenberg G, Uiterwaal CS, Kimpen JL, van de Putte EM. Effectiveness of internet-based cognitive behavioural treatment for adolescents with chronic fatigue syndrome (FITNET): a randomised controlled trial. The Lancet. 2012; 379:1412–1418.
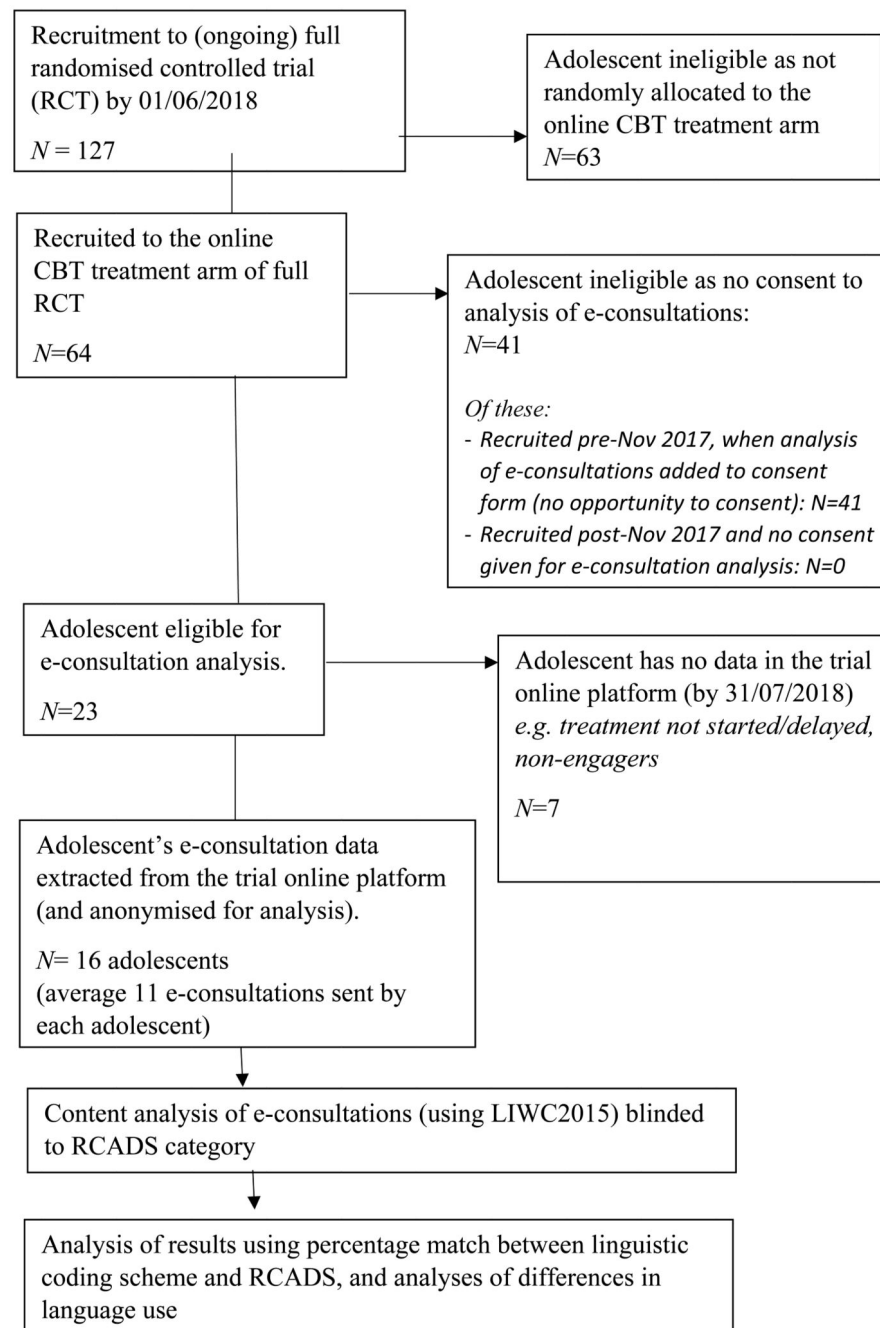
Pennebaker JW. The secret life of pronouns. New Scientist. 2011; 211:42–45.

Pennebaker, JW, Booth, RJ, Francis, ME. Linguistic inquiry and word count: LIWC [Computer software]. Austin: TX liwc. net; 2007.

Pennebaker, JW, Boyd, RL, Jordan, K, Blackburn, K. The development and psychometric properties of LIWC2015. The University of Texas; Austin: TX: 2015.

Pennebaker JW, King LA. Linguistic styles: Language use as an individual difference. Journal of Personality and Social Psychology. 1999; 77:1296. [PubMed: 10626371]

Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language use: Our words, our selves. Annual review of psychology. 2003; 54:547–577.

Pinquart M, Shen Y. Depressive symptoms in children and adolescents with chronic physical illness: an updated meta-analysis. Journal of Pediatric Psychology. 2010; 36:375–384. [PubMed: 21088072]

Plug L, Sharrack B, Reuber M. Seizure, fit or attack? The use of diagnostic labels by patients with epileptic or non-epileptic seizures. Applied Linguistics. 2009; 31:94–114.

Robson C, Drew P, Walker T, Reuber M. Catastrophising and normalising in patient's accounts of their seizure experiences. Seizure-European Journal of Epilepsy. 2012; 21:795–801. [PubMed: 23040370]

Rodriguez AJ, Holleran SE, Mehl MR. Reading Between the Lines: The Lay Assessment of Subclinical Depression From Written Self-Descriptions. Journal of Personality. 2010; 78:575–598. [PubMed: 20433631]

Rude S, Gortner EM, Pennebaker J. Language use of depressed and depression vulnerable college students. Cognition & Emotion. 2004; 18:1121–1133.

Sloan DM. It's all about me: Self-focused attention and depressed mood. Cognitive Therapy and Research. 2005; 29:279–288.

Stivers T. Coding social interaction: A heretical approach in conversation analysis? Research on Language and Social Interaction. 2015; 48:1–19.

Tamnes CK, Herting MM, Goddings AL, Meuwese R, Blakemore SJ, Dahl RE, et al. Crone EA. Development of the cerebral cortex across adolescence: A multisample study of interrelated longitudinal changes in cortical volume, surface area and thickness. J Neurosci. 2017; doi: 10.1523/JNEUROSCI.3302-16.2017

Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology. 2010; 29:24–54.

Thase ME. Assessment of depression in patients with chronic fatigue syndrome. Reviews of Infectious Diseases. 1991; 13

Van Der Lee JH, Mokkink LB, Grootenhuis MA, Heymans HS, Offringa M. Definitions and measurement of chronic health conditions in childhood: a systematic review. JAMA. 2007; 297:2741–2751. [PubMed: 17595275]

Van der Zanden R, Curie K, Van Londen M, Kramer J, Steen G, Cuijpers P. Web-based depression treatment: Associations of clients′ word use with adherence and outcome. Journal of Affective Disorders. 2014; 160:10–13. [PubMed: 24709016]

Vigerland S, Lenhard F, Bonnert M, Lalouni M, Hedman E, Ahlen J, et al. Ljótsson B. Internet-delivered cognitive behavior therapy for children and adolescents: a systematic review and meta-analysis. Clinical Psychology Review. 2016; 50:1–10. [PubMed: 27668988]

Warner LJ, Lumley MA, Casey RJ, Pierantoni W, Salazar R, Zoratti EM, et al. Simon MR. Health effects of written emotional disclosure in adolescents with asthma: A randomized, controlled trial. Journal of Pediatric Psychology. 2005; 31:557–568. [PubMed: 16014820]

Weintraub, W. Verbal behavior in everyday life. New York: Springer; 1989.

Zimmermann J, Wolf M, Bock A, Peham D, Benecke C. The way we refer to ourselves reflects how we relate to others: Associations between first-person pronoun use and interpersonal problems. Journal of Research in Personality. 2013; 47:218–225.

Zimmermann J, Brockmeyer T, Hunn M, Schauenburg H, Wolf M. First-person Pronoun Use in Spoken Language as a Predictor of Future Depressive Symptoms: Preliminary Evidence from a Clinical Sample of Depressed Patients. Clinical Psychology & Psychotherapy. 2017; 24:384–391. [PubMed: 26818665]

Key Practitioner Message

- It is important to identify comorbid depression to aid and inform the clinical care of young people with chronic illness.

- Use of online interventions is increasing in healthcare, and evidence is growing for the effectiveness of online treatment of paediatric CFS/ME.

- Naturally-occurring features of language use within emails present an opportunity to assess psychological state in an online clinical setting, in which other indicators (such as tone of voice and body language) are not available.

- This study presents a first attempt to develop and test a theory and evidence-based linguistic profile of depression in a novel clinical setting (e-consultations) and patient population (paediatric CFS/ME).

- Recommendations to pursue in future work with larger samples are provided.

RCADS assessment	Baseline data collection	Completing CBT online chapters, filling in diaries, and corresponding with the therapist by email.	
		1st session ▪ Managing expectations of therapy. ▪ Establishing therapist-patient relationship.	**2nd – 4th session** ▪ Establishing therapist-patient relationship. ▪ Reviewing online chapters and negotiating problems with sleep, unhelpful thoughts, and physical activity. ▪ Introducing homework.

Enrolment → 2 months (average)

Figure 1. The trial intervention and patient journey.

| Recruitment to (ongoing) full randomised controlled trial (RCT) by 01/06/2018 *N* = 127 | → | Adolescent ineligible as not randomly allocated to the online CBT treatment arm *N*=63 |

Recruited to the online CBT treatment arm of full RCT

N=64

→

Adolescent ineligible as no consent to analysis of e-consultations: *N*=41

Of these:
- *Recruited pre-Nov 2017, when analysis of e-consultations added to consent form (no opportunity to consent): N=41*
- *Recruited post-Nov 2017 and no consent given for e-consultation analysis: N=0*

Adolescent eligible for e-consultation analysis.

N=23

→

Adolescent has no data in the trial online platform (by 31/07/2018) *e.g. treatment not started/delayed, non-engagers*

N=7

Adolescent's e-consultation data extracted from the trial online platform (and anonymised for analysis).

N= 16 adolescents (average 11 e-consultations sent by each adolescent)

Content analysis of e-consultations (using LIWC2015) blinded to RCADS category

Analysis of results using percentage match between linguistic coding scheme and RCADS, and analyses of differences in language use

Figure 2. Study flowchart.

Table 1

The Self-Regulatory Model: Linguistic Features associated with Chronic Illness

Illness representations (dimension)	Description	Linguistic feature
Identity	How symptoms are labelled by the patient.	Health
Cause	Beliefs about what causes the illness.	Causation; Insight; Social processes; Past-focus
Consequences	The perceived impact of the illness.	Health; Social processes
Timeline	Expectations about the chronicity and course of the illness.	Future-focus; Past-focus; Social processes
Controllability	Beliefs about the controllability of symptoms, through lifestyle management or medical treatment.	Causation; Insight; Past-focus; Social processes

Note. The self-regulatory model (Leventhal et al. 1992).

Linguistic features were taken from the LIWC2015 (Pennebaker et al., 2015).

Table 2

The Cognitive Model of Depression: Linguistic Features associated with Low Mood

Unhelpful thinking patterns	Description	Linguistic feature
Mental filter	The tendency to dwell on the negatives and discount any positive information.	Negative emotion Positive emotion
All-or-nothing thinking	A rigid 'black or white' perception of the world.	Absolutist
Rumination	The endless repetitive process of going over thoughts.	First-person pronouns Causation Insight
Hopelessness	The tendency to expect the worst.	Negative emotion Positive emotion Past-focus Future-focus
Should statements	A very (self-)critical inflexible thinking style.	Discrepancy First-person pronouns

Note. The cognitive model of depression (Beck, 1967).

Linguistic features were taken from the LIWC2015 (Pennebaker et al., 2015) and absolutist (Al-Mosaiwi & Johnstone, 2018) dictionaries.

Table 3

Reasons for Inclusion and Exclusion of Hypothesised Linguistic Features

Linguistic Feature	Reason for Inclusion/Exclusion
Included	
Negative emotion[a]	Theoretical basis and empirical evidence
Positive emotion[a]	Theoretical basis and empirical evidence
First-person singular pronoun[a]	Theoretical basis and empirical evidence
Absolutist[a]	Theoretical basis and empirical evidence
Excluded	
Causation[a]	• Potentially affected by the study context • Potentially reflective of CFS/ME or depression • A less conclusive evidence-base • Available linguistic data for personality disorder only
Insight[a]	• Potentially reflective of CFS or depression • A less conclusive evidence-base • Available linguistic data for personality disorder only
Discrepancy[a]	• A less conclusive evidence-base • Available linguistic data for personality disorder only
Past-focus[a]	• Potentially reflective of CFS or depression • Available linguistic data for personality disorder only
Future-focus[a]	• Potentially reflective of CFS or depression • Available linguistic data for personality disorder only
Achieve[b]	• Potentially reflective of CFS or depression • A lack of previous evidence

Note. Linguistic features were taken from the LIWC2015 (Pennebaker et al., 2015) and absolutist (Al-Mosaiwi & Johnstone, 2018) dictionaries.

[a] Identified deductively.

[b] Identified inductively.

Table 4

Baseline Demographic and Clinical Factors

	Current sample (N = 16)	Main trial sample (N = 111)[a]	Difference (P)[b]
Female, n (%)	12 (75)	66 (60)	.282
Age (years), mean (SD)	13.69 (1.01)	14.14 (1.72)	.207
Number of symptoms, mean (SD)	7.88 (1.36)	7.41 (1.64)	.352
Duration of illness (months), mean (SD)	14.25 (8.74)	23.86 (20.84)	**.039**
Fatigue[c], mean (SD)	26.88 (3.36)	23.67 (5.15)	**.014**
Pain[d], n (%) moderate-severe	15 (94)	63 (57)	**.004**
School attendance, n (%)	40% 9 (56)	61 (55)	1.00

Note. P values significant at the .05 are in bold.

[a]Characteristics are shown for participants recruited to the online CBT treatment arm of the trial by 01/06/2018, excluding participants of the current sample. *N* = 109 for fatigue, pain, and school attendance.

[b]Pearson's Chi-Square for sex and pain, Fisher's exact test for school attendance, Mann-Whitney U tests for age, number of symptoms, duration of illness, and fatigue.

[c]Chalder Fatigue Scale, range 0-33.

[d]Visual Analogue Scale, ratings of >45 = 'moderate-severe' pain.

Table 5

RCADS Baseline Categorisation

	Current sample (N = 16)	Main trial sample (N = 111)
	n (%)	n (%)
None[a]	1 (6)	52 (47)
Possible[a]	5 (31)	10 (9)
Probable[a]	10 (63)	49 (44)

Note. RCADS = Revised Child Anxiety and Depression Scale.

[a]None = RCADS t-score 64; Possible = RCADS t-score 65-69; Probable = RCADS t-score 70)

Table 6

Scoring and Categorisation of Linguistic Dimensions by the Coding Scheme

	P1[a]	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
First-person – %[b] (score)[b]	13.15 (2)	8.84 (0)	13.04 (2)	11.19 (1)	12.01 (2)	12.50 (2)	6.21 (0)	9.09 (0)	4.55 (0)	6.85 (0)	10.99 (0)	5.56 (0)	13.36 (2)	9.34 (0)	6.75 (0)	8.45 (0)
Positive – %[b] (score)[b]	3.48 (0)	10.88 (0)	5.80 (0)	5.22 (0)	2.71 (2)	3.12 (0)	2.48 (2)	2.18 (2)	10.23 (0)	4.11 (0)	2.16 (2)	8.33 (0)	1.18 (2)	2.71 (2)	4.91 (0)	3.52 (0)
Negative – %[b] (score)[b]	2.71 (1)	0.68 (0)	0.00 (0)	0.75 (0)	1.35 (0)	0.00 (0)	1.86 (1)	1.82 (1)	2.27 (1)	1.37 (0)	0.54 (0)	5.56 (2)	2.95 (2)	1.20 (0)	2.45 (1)	1.17 (0)
Absolutist – %[b] (score)[b]	0.58 (0)	0.68 (0)	2.90 (2)	0.00 (0)	1.35 (1)	0.00 (0)	0.62 (0)	1.10 (1)	0.00 (0)	0.00 (0)	0.18 (0)	0.00 (0)	0.39 (0)	0.90 (0)	0.61 (0)	0.7 (0)
Total score[c]	3	0	4	1	5	2	3	4	1	0	2	2	6	2	1	0
Category[d]	B	N	**B**	N	B	N	B	B	N	**N**	N	N	**C**	N	N	N
RCADS-MDD[e]	Pr[f]	Pr	**Po**	Po	Pr[f]	Pr[f]	Pr	Pr	Po	**N**	Po	Pr[f]	**Pr[f]**	Pr	Pr	Po

Note. Category matches between linguistic coding scheme and RCADS are in bold.

[a]P = participant.

[b]% = presented as a percentage of total words by each participant; score = non-case 0, borderline case 1, case 2.

[c]Non-case 0-2, borderline case 3-5, case 6-8.

[d]N = non-case, B = borderline case, C = case.

[e]RCADS-MDD = Revised Child Anxiety and Depression Scale (RCADS) Major Depressive Disorder (MDD) subscale; N= none, Po = possible, Pr = probable depression.

[f]Participants with possible or probable anxiety according to the RCADS Total Anxiety subscale.

Table 7

Percentage Match between Linguistic Coding Scheme and RCADS Categorisation of Depression

		RCADS			
		None [a] n (%)	Possible [a] n (%)	Probable [a] n (%)	
Linguistic coding scheme	Primary analysis [b] Non-case	1 (6)	4 (25)	5 (31)	
	Borderline case	0 (0)	1 (6)	4 (25)	
	Case	0 (0)	0 (0)	1 (6)	
		None n (%)	Possible n (%)	Probable n (%)	
	Sensitivity analysis [b] Non-case	1 (6)	5 (31)	8 (50)	
	Borderline case	0 (0)	0 (0)	1 (6)	
	Case	0 (0)	0 (0)	1 (6)	

Note. RCADS = Revised Child Anxiety and Depression Scale.

[a]None = RCADS t-score ≤ 64; Possible = RCADS t-score 65-69; Probable = RCADS t-score ≥ 70).

[b]Primary analysis = first four messages of each participant; Sensitivity analysis = all messages sent by each participant.

Table 8

Differences in Linguistic Dimensions by RCADS Category

	Total sample (N = 16)	None[a] (n = 1)	Possible[a] (n = 5)	Probable[a] (n = 10)	Difference[b]
	Mean (SD)	Mean	Mean (SD)	Mean (SD)	(P)
Word count	256.75 (201.41)	73.00	254.40 (221.57)	276.30 (203.99)	.585
Linguistic dimensions (deductive)					
S-R model only[c]					
Health	1.22 (1.43)	.00	2.30 (2.05)	.80 (.72)	.103
Social processes	6.72 (2.44)	8.22	5.99 (2.34)	6.93 (2.63)	.525
C model only[c]					
First-person	9.49 (2.92)	6.85	9.64 (3.28)	9.68 (2.94)	.736
Negative emotion	1.67 (1.37)	1.37	.95 (.85)	2.06 (1.53)	.231
	Total sample (N = 16)	None[a] (n = 1)	Possible[a] (n = 5)	Probable[a] (n = 10)	Difference[b]
	Mean (SD)	Mean	Mean (SD)	Mean (SD)	(P)
Positive emotion	4.56 (2.91)	4.11	5.39 (3.06)	4.20 (3.06)	.554
Absolutist	0.63 (0.74)	.00	.76 (1.23)	.62 (.43)	.423
Discrepancy	1.76 (1.60)	5.48	2.53 (1.65)	1.01 (.76)	**.049**
S-R and C model[c]					
Past-focus	3.76 (2.45)	.00	2.27 (1.84)	4.89 (2.08)	**.043**
Future-focus	2.89 (1.81)	1.37	3.32 (1.78)	2.83 (1.91)	.468
Causation	1.22 (0.77)	.00	1.73 (.71)	1.08 (.66)	.156
Insight	3.01 (2.36)	5.48	3.43 (3.07)	2.56 (2.02)	.433
	Total sample (N = 16)	None[a] (n = 1)	Possible[a] (n = 5)	Probable[a] (n = 10)	Difference[b]
	Mean (SD)	Mean	Mean (SD)	Mean (SD)	(P)

Linguistic dimensions (inductive)

| Achieve | 2.16 (1.64) | 1.37 | 2.43 (1.71) | 2.11 (1.74) | .823 |

Note. Linguistic dimensions are presented as a mean percentage of total words used by an adolescent. RCADS = Revised Child Anxiety and Depression Scale. *P* values significant at the .05 are in bold.

[a]None = RCADS t-score 64; Possible = RCADS t-score 65-69; Probable = RCADS t-score 70).

[b]Kruskal-Wallis test (alpha level of .05).

[c]S-R model = Self-regulatory model; C model = Cognitive model of depression.