

Reading between the lines - Artificial Intelligence-based Co-Facilitator (AICF) detects and monitors group cohesion outcome in online cancer support groups: A single armed trial study

Yvonne W Leung, Elise Wouterloot, Achini Adikari, Jinny Hong, Veenaajaa Asokan, Lauren Duan, Claire Lam, Carlina Kim, Kai P Chan, Daswin De Silva, Lianne Trachtenberg, Heather Rennie, Jiahui Wong, Mary Jane Esplen

Submitted to: JMIR Cancer
on: October 03, 2022

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 21

0..... 21

Preprint
JMIR Publications

Reading between the lines - Artificial Intelligence-based Co-Facilitator (AICF) detects and monitors group cohesion outcome in online cancer support groups: A single armed trial study

Yvonne W Leung^{1,2,3} PhD; Elise Wouterloot¹ PhD; Achini Adikari⁴ PhD; Jinny Hong¹ PhD; Veenaajaa Asokan¹ BSCN; Lauren Duan¹; Claire Lam¹ BA; Carlina Kim¹ MDS; Kai P Chan¹ MEng, DMin; Daswin De Silva⁴ PhD; Lianne Trachtenberg^{1,5} PhD; Heather Rennie^{1,6} MA, MSc, RCC, RMFT; Jiahui Wong^{1,1} MD, PhD; Mary Jane Esplen¹ PhD

¹de Souza Institute University Health Network Toronto CA

²Department of Psychiatry Temerty Faculty of Medicine University of Toronto Toronto CA

³College of Professional Studies Northeastern University Toronto CA

⁴Centre for Data Analytics and Cognition La Trobe University Melbourne AU

⁵Centre for Psychology and Emotional Health Toronto CA

⁶BC Cancer Agency Vancouver CA

Corresponding Author:

Yvonne W Leung PhD
de Souza Institute
University Health Network
222 St Patrick St Office 503
Toronto
CA

Abstract

Background: Commonly offered as supportive care, therapist-led online support groups (OSGs) are a cost-effective way to provide support to individuals affected by cancer. One important indicator of a successful OSG session is group cohesion; however, monitoring group cohesion can be challenging due to the lack of non-verbal cues and in-person interactions in text-based OSGs. The Artificial Intelligence-based Co-Facilitator (AICF) was designed to contextually identify therapeutic outcomes from conversations and produce real-time analytics.

Objective: The aim of this study was to develop a method to train and evaluate AICF's capacity to monitor group cohesion.

Methods: Human scorers used a confusion matrix to evaluate the performance of AICF. AICF performance was also compared against the natural language processing software Linguistic Inquiry Word Count (LIWC).

Results: AICF was trained on 80,000 messages obtained from Cancer Chat Canada. We tested AICF on 34,048 messages. Human experts scored 6,797 (20%) of the messages to evaluate the ability of AICF to classify group cohesion. Results showed that machine-learning algorithms combined with human input can detect group cohesion, a clinically meaningful indicator of effective OSGs. After re-training with human input, AICF reached a F1-score of 0.82. AICF performed slightly better at identifying group cohesion compared to LIWC.

Conclusions: AICF has the potential to assist therapists by detecting discord in the group amenable to real-time intervention. Overall, AICF presents a unique opportunity to strengthen patient-centered care in virtual settings by attending to individual needs.

(JMIR Preprints 03/10/2022:43070)

DOI: <https://doi.org/10.2196/preprints.43070>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>



Original Manuscript

Reading between the lines - Artificial Intelligence-based Co-Facilitator (AICF) detects and monitors group cohesion outcome in online cancer support groups: A single armed trial study

Abstract

Background: Commonly offered as supportive care, therapist-led online support groups (OSGs) is a cost-effective way to provide support to individuals affected by cancer. One important indicator of a successful OSG session is group cohesion; however, monitoring group cohesion can be challenging due to the lack of non-verbal cues and in-person interactions in text-based OSGs. The Artificial Intelligence-based Co-Facilitator (AICF) was designed to contextually identify therapeutic outcomes from conversations and produce real-time analytics.

Objective: The aim of this study was to develop a method to train and evaluate AICF's capacity to monitor group cohesion.

Methods: AICF used a text classification approach to extract the mentions of group cohesion within conversations. A sample of data was annotated by human scorers, which was used as the training data to build the classification model. The annotations were further supported by finding contextually similar group cohesion expressions using word embedding models as well. AICF performance was also compared against the natural language processing software Linguistic Inquiry Word Count (LIWC).

Results: AICF was trained on 80,000 messages obtained from Cancer Chat Canada. We tested AICF on 34,048 messages. Human experts scored 6,797 (20%) of the messages to evaluate the ability of AICF to classify group cohesion. Results showed that machine-learning algorithms combined with human input could detect group cohesion, a clinically meaningful indicator of effective OSGs. After re-training with human input, AICF reached the F1-score of 0.82. AICF performed slightly better at identifying group cohesion compared to LIWC.

Conclusions: AICF has the potential to assist therapists by detecting discord in the group amenable to real-time intervention. Overall, AICF presents a unique opportunity to strengthen patient-centered care in virtual settings by attending to individual needs.

Keywords: group cohesion; LIWC; online support groups; natural language processing; emotion analysis; machine learning; sentiment analysis; emotion detection; integrating human knowledge; emotion lining

Introduction

Virtual care has become increasingly important in healthcare delivery as a means to accessibly reduce emotional distress. Online support groups (OSG) offer a convenient solution to those who cannot attend in-person support groups [1–3]. Professionally-led OSGs occur in real-time with participants engaging with a therapist and other participants in the group. Therapists facilitate the sharing of personal experiences to foster a mutually supportive environment. OSG participants report an increased sense of empowerment and control, as well as improved knowledge about their conditions [4].

Cancer Chat Canada (CCC) offers online professionally-led, synchronous, text-based support groups

to cancer patients/caregivers across six Canadian provinces with a text-based nature allowing for anonymity while reaching people in rural areas. All groups provided via CCC are manual-based, consisting of 8-12 sessions. Each session focuses on a specific theme, homework readings, and interactive activities. Participants exchanged their experiences and ideas through a chatbox on the CCC platform. During sessions, therapists facilitate discussions based on the weekly readings, address issues/concerns, attend to emergent emotional needs of the members, and employ therapeutic techniques that promote a continuous sense of mutual support among the 6-10 members [5].

For group interventions to be effective, therapists encourage authentic emotional expression while effectively monitoring and addressing signals of distress [6]. However, the absence of visual cues, along with the simultaneous entries by multiple participants, can impose challenges for therapists to attend to all participants' needs during the session[4]. Therapists' failure to recognize and respond to participants' expressions of distress can reduce the participants' perceived level of support, safety, and trust in the group, leading to disengagement and attrition [7].

One way to reduce attrition and improve OSG services is through tracking and monitoring group cohesion[8, 9]. A cohesive group experiences a sense of warmth, comfort, acceptance, affiliation, and support from other members they value [5]. Group cohesion is associated with positive participant outcomes, including reductions in distress and improvements in interpersonal functioning [5].

Traditionally, group cohesion is measured by participant self-report instruments, such as the Harvard Community Health Plan Group Cohesiveness Scale[10] and the Group Cohesion Scale Revised [11]. Alternatively, it can be measured by content analysis, where analysts assign ratings to the participants' statements [12]. While useful, these approaches have limitations of participant recall bias, measurement fatigue in self-reports, and time and cost of labor in post-hoc qualitative analyses.

Previous studies demonstrate that a higher frequency of first-person singular pronouns use (i.e., I, my), also referred to as "iTalk" or self-referential language, is a linguistic marker of general distress and is associated with negative psychological outcomes such as depression and suicidal behaviors [13–15]. In contrast, collective identity language use (i.e. our group, us) was instrumental to group attachment [16]; with greater uses of references to the group as a whole and to other members predicting reduced symptoms of grief [12]. Aside from content analysis, such as Psychodynamic Work, Object Rating System [17], many studies adopted computerized textual analysis systems such as dtSearch [18], Linguistic Inquiry, and Word Count (LIWC) to track levels of cohesion through text [19–22]. In particular, Lieberman et al. detected group cohesion by combining LIWC to count the proportion of group referential language use and dtSearch to count words indicative of positive connotations (i.e., hope, altruistic, accept, affection) within 10 words of such group referential language in an OSG for Parkinson's patients [21]. However, Alpers et al. questioned the software's ability to process complex communications, suggesting that future studies should develop systems that analyze the context of discourse for real-time analysis [20].

Given the evidence, group cohesion can be systematically measured by a well-designed computer analytical system. We designed the Artificial Intelligence-based Co-Facilitator (AICF) to contextually identify therapeutic outcomes, including group cohesion from conversations, and produce real-time analytics [23–26]. AICF can track basic emotions, including joy, sadness, anger, trust, fear, anticipation, disgust, surprise, and psychological outcomes such as distress, group cohesion, and hopelessness for each participant in the OSGs [23, 24]. AICF extracted emotions from the text by parsing through over 120,000 lines of chat messages from a training dataset to multiple levels of granularity: word, phrase, sentence, post, and posts by each user [27]. AICF employed several NLP techniques, such as word2vec [28] and text classification models. Classification models

were trained to classify posts containing group cohesion mentions to determine the level of group cohesion in this online conversation setting. Each level of extraction served as an input for calibration for the subsequent extraction to increase accuracy [27, 29, 30]. AICF could therefore track and inform facilitators of each participant's level of cohesive statement use in their posts.

We hypothesized that AICF could detect first-person plural pronoun use (e.g. we, our) in OSGs and group-references language use (a.k.a “we-talk”) as group cohesion, machine learning-based Natural Language Processing (NLP) could also identify a broader definition of group cohesion, including expressing gratitude, mutual support, and sense of belonging.

Objective:

The current study is focused on the development of a method to train and evaluate AICF's ability to detect group cohesion among cancer OSG members.

AICF Training and Development

The steps involved in the training and development of AICF's cohesion detection is outlined below:

Collecting Design Specifications

Experienced CCC therapists participated in phase I and phase III focus groups to obtain design specifications for which clinically meaningful outcomes AICF should capture and provide real-time analytics for, as well as the pros and cons after experiencing AICF clinically. All therapists who responded to our request to participate were involved in the study and are experienced in their field. In addition to the individual emotion tracking feature, the therapists expressed interest in tracking group processes with a particular emphasis on group cohesion. Therapists described group cohesion as a high frequency of posting by members with a sense of interconnectedness through replying to others. A successful group session results in members feeling supported and acknowledged by other group members. The results herein this manuscript excluded the results of these focus groups as they were published elsewhere.

Ethical Considerations: The study protocol including the human subject recruitment method was approved by the University Health Network Research Ethics Board. All identifiable information was removed from the quotes in this report. Participants were compensated with a \$100 gift card upon the completion of the focus group.

Scoring Guide Development

A literature-based guide was developed to ensure that group cohesion statements were consistently identified and annotated by the human team.

Group Cohesion is the sense of warmth, acceptance, support, and belongingness to the members [5], a sense of closeness, and participation [31]. It is measured by statements that reflect a sense of belonging and support in the group.

This belonging and support could be expressed with the statement themes below [32]. The following examples were from the CCC chat training data.

- Reassurance/encouragement between peers

- Expressing support/feeling supported
- Deepening emotional disclosure and trust
- A sense of belonging
- Gratitude for the group
- Finding shared experiences and commonalities
- Looking forward to future sessions or connecting outside of the group
- Reflecting on the positive aspects of the group

Creating Training Data

To train AICF to identify group cohesion, 1000 examples of cohesive statements from 10 OSG sessions were annotated by two human group therapists (E.W. and J.H.). These annotated examples were used for training the algorithm.

Algorithm Development

Feature Selection of Group Cohesion Expressions

First, a corpus of CCC chat sessions (~80,000 messages) was used to train a word embedding model using Word2Vec using the Gensim library in Python (`gensim.models.Word2Vec (documents, size=100, window=10, min_count=2, workers=10))`). This enabled the creation of a vector representation for each word in the corpus. This positioned semantically similar expressions in closer proximity to generate contexts of cancer OSG discussion. Second, to expand the group cohesion mentions, the annotated samples were fed into the trained Word2Vec model as inputs to query for neighboring words. This resulted in a set of semantically similar, contextually relevant group cohesion expressions. This enabled AICF to identify statements representing group cohesion, including keywords such as “us,” “we,” and “our group,” as well as themes such as expressing gratitude, eagerness to attend upcoming group sessions, chatting outside of group time, mutual support, and a sense of belonging.

Training the Classifiers of Group Cohesion

To produce the probability of each post containing group cohesion, three models, multinomial naive Bayes, logistic regression, and Multi-layer Perceptron classifier with the group cohesion features selected were trained using the training dataset.

Before training the classifier, a series of feature engineering steps were followed. Feature

engineering is the process of creating features by extracting information from the data. For this purpose, the TF-IDF (term frequency-inverse document frequency) approach was used. TF-IDF is a statistical measure that evaluates how relevant a word is to a document in a collection of documents. It is performed by multiplying the term frequency and inverse document frequency of the word across a set of documents. In this classification task, the TF-IDF vectorizer was used with a limit of 5000 words capturing both unigrams (single words) and bi-grams (two words that occur together). Next, the vectorizer was applied to the pre-processed training dataset.

Once the dataset was transformed, it was used to train multiple classifiers; Naïve Bayes, Random Forest, Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), and Logistic regression models. The objective of training multiple classifiers was to increase the performance of the final classification by incorporating multiple high-performing classifiers. This technique is called ‘soft voting’, which is an ensemble machine learning technique that combines predictions from multiple models. The following table shows the F1 score of the trained classifiers.

Classifier	F1 Score
SVM	0.63
Naïve Bayes	0.79
MLP	0.77
Random Forest	0.72
Logistic Regression	0.82

Group Cohesion Score Calculation

Based on this, the top three classifiers were selected; Naïve Bayes, MLP, and Logistic Regression. The outcomes of all three classifiers were used to make the final prediction. If a post is classified into the same label by 2/3 classifiers, then the output label is used as the final outcome. A confidence value was generated for each classification based on the weighted f1 scores of each classifier. The average F1 score using three classifiers was 0.8.

In order to improve the performance of the model, an active learning approach [ref] was used where human input is used as feedback to fine-tune the models. Therapists examined 20% of the outputs using a confusion matrix (See Human Scoring). The scoring results were then used as a feedback loop to improve the list of keywords of queried expressions. Lastly, to fine-tune group cohesion extraction, linguistic rules were hand-coded to handle exceptions such as past tense and empathetic questions by participants (Figure 1).

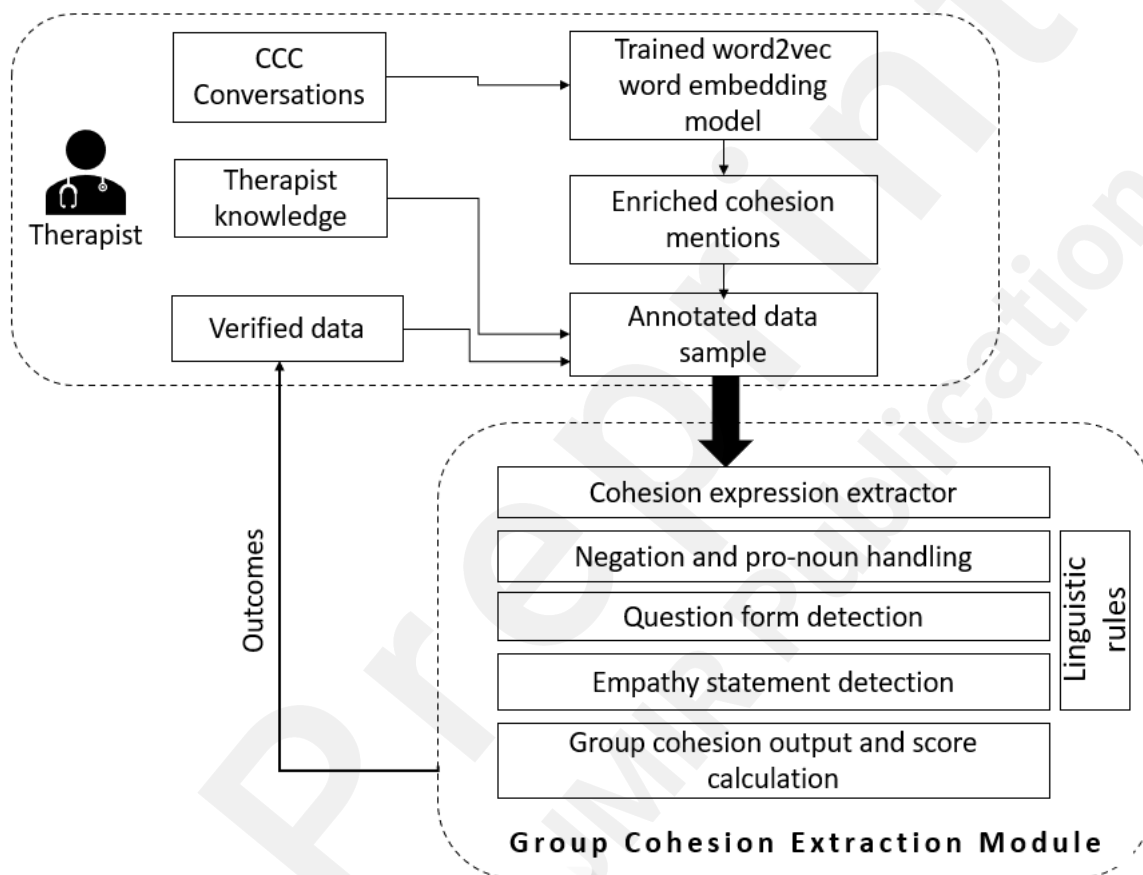
Linguistic Rules for Group Cohesion Score Adjustment

The following rules were added after the first round of scoring based on therapists’ feedback:

- 1) Intensifiers - We have used the intensifiers from a pre-trained library, VADER [33], which considers intensity boosters such as ‘very’ and ‘so much’ to enhance the valence.

- 2) Past-tense (in the part-of-speech tagging via the NLTK Python library)- The score would be multiplied by 0.5 if past-tense was present as the event had happened in the past, we assume that the effect of the event on the person would subside.
- 3) Negation - the calculated cohesion score would be set to zero in case of a negation expression.
- 4) First-person tagging - This was set to be 'False' if 2nd/3rd person pronouns were found.
- 5) Empathy - If an empathy statement were found, then the calculated group cohesion score would be doubled to denote the intensity.

Figure 1: Process of group cohesion extraction



Finally, an aggregated score of group cohesion (β) was calculated for specific time intervals using the following formula:

$$\beta = \frac{|C_{(t,t+T)}|}{|A_{(t,t+T)}|} \quad \forall t \in T$$

Where:

β = Group cohesion score

T= Specified time interval (30 minutes)

$A(t,t+T)$ = Set of all posts occurred during the time t to t+T

$C(t,t+T)$ = Set of cohesion mentioned posts occurred during the time t to t+T

A group cohesion score was displayed and updated at 30-minute intervals on the 90-minute timeline

in a real-time dashboard for therapists.

Methods

Active learning via Human Scoring

Outputs were scored by undergraduate students (responsible for basic emotions), graduate students, and clinical experts (responsible for clinical and process outcomes). The team scored 20% of the output to inform AICF development, which was improved in light of the human scoring results. The updated AICF was run on the data of a new OSG (test data). Each AICF version was saved before training with new data. The team scored the output using definitions/examples from well-established psychometric measures such as the Group Cohesiveness Scale and Group Openness and Cohesion Questionnaire. A confusion matrix was used to score AICF outputs. The scoring process was based on Recall, Precision, and F1 measures. Scorers' feedback using their domain expertise was used to improve AICF's performance until it achieved F1=80% before deploying in real-time OSG for beta-testing [34].

LIWC Evaluation

The Linguistic Inquiry Word Count (LIWC) software [35], considered the Gold-Standard of Psychology-Based NLP, was used as a validation tool. LIWC reads a given text and calculates the percentage of total words in the text that match each of the LIWC dictionary categories. We tried to capture the concept of group cohesion using multiple LIWC dictionary categories: 'we', 'positive emotion', 'family', 'friend', and 'affiliation' as the measurement criteria. We classified the text as an instance of group cohesion when at least 3 out of 5 criteria were met.

Results

The results herein only focus on the human evaluation of the AICF system and its ability to detect group cohesion. We compared AICF to LIWC using human judgment using the confusion matrix and F1 score to measure accuracy and precision. AICF was run on 34,048 messages of CCC chat history to generate outputs for human scoring. Every fifth message was scored, totaling 6,797 messages (20%). The precision, recall, and F1 scores are reported in Table 1. In this first round, AICF missed a high number of group cohesive statements (Table 2). All scored statements were incorporated into AICF for improvement.

In the second round, the team checked another 296 of 1,208 messages (20%) from a separate set of CCC group conversations. AICF was able to improve the false negative rate (Recall) from 0.52 to 0.70.

We also ran LIWC on new OSG data (12,034 messages) from the CCC platform. Precision, recall, and F1 scores are listed in Table 2.

Table 1. Themes, keywords, and examples of AICF outputs

True Positive Themes	Examples
Reassurance among peers	"Good for you, that must have been really difficult to do"

Expressing support/ feeling supported	<i>“sending hugs”; “thanks everyone for your support <3 “</i>
Deeping emotional disclosure and trust	<i>“At my last treatment, I felt really scared, I didn’t want to tell my family this. But I can tell this group”</i>
Sense of belonging	<i>“I am glad to be part of this group”</i>
Gratitude for the group and peers	<i>“Thank you everyone, this is such a great group.”</i>
Shared experiences	<i>“It is amazing how much we all have in common”</i>
Reflecting on the positive aspects of the group	<i>“This group has been a great resource venue for meeting and I will continue down the road with fond memories of the time spent here”</i>
Anticipating future groups/chats	<i>“I can’t wait to chat with you all next week”</i>
Keywords: we, us, our, you all / all of you, thanks / thank you, time, chat, group, support, miss	<i>“<u>we</u> are friends”; “thinking of <u>you all</u>!”; “time went by fast!”; “A very good <u>chat</u> session”; “I’m so glad I found this <u>group</u>”; “<u>thank you</u> for your <u>support</u> and input”; “I am going to <u>miss</u> this <u>group</u>”</i>
False Positive Themes	Example
False detection of “we” or “us” when it is not indicative of group cohesion	<i>“When I mentioned that last week, when we all said what our situations were, no one even acknowledged it. I was very hurt by that”</i>
Talking about support from non-group members	<i>“When I put my things in order I involved my children and as challenging and emotional as it was it made it easier for me and for them - we laughed and cried but it really made me feel supported”</i>
False Negative Themes	Example
Idioms/expressions	<i>“been there, am there... got the t-shirt”; “I hope you don't think you've been placed in a hot seat. It's just that I missed you and worried about you.”; “we are joining our circle“</i>
Missed detection of providing empathy, encouragement, appreciation, and support to other members	<i>“I am with you, be strong please”; “Are you going to be alright? I want you to know how much I appreciate your presence.”</i>
Missed detection of agreeing with or relating to others’ stories	<i>“I can relate”; “I feel the same way”;</i>
Missed statements because of lack of “we” language	<i>“I am sure there will be a sense of connection, so much sharing already”; “I am sure I will think of you often“</i>

Within the “true positive” instances identified by AICF in agreement with the human scorers, several thematic categories and keywords emerged. They closely align with established measures of group cohesion [32], including expressions of support or a sense of belonging (Table 2). Moreover, some keywords consistently emerged within the true positive statement classifications (e.g. ‘we’, ‘us’, ‘our’, ‘group’, ‘support’). Among the false positive identifications, it was typically due to a missed subtle negation within the sentence or when a participant wrote about a supportive person or activity from outside the group (Table 2).

Where AICF missed a classification of group cohesion (i.e. a false negative), it was typically also due to nuanced conversational features on which it had not yet been trained, such as local expressions/idioms, supportive responses to others or statements missing identified group cohesion keywords (such as ‘we,’ ‘us,’ and ‘our’) (Table 2). These correct and incorrect classifications were used to refine AICF detection of group cohesion as the algorithm progresses in development.

Table 2. AICF performance evaluation for identification of Group Cohesion

Scoring Round/ Method	Precision	Recall	F1
1st	0.99	0.52	0.68
2nd	0.98	0.70	0.82
LIWC	0.36	0.23	0.28

Discussion

Principal Results

AICF, an ensemble of NLP and machine-learning algorithms combined with annotation and human scoring, offers a novel way of measuring the group cohesion changes for each group member and alerting the therapist of these changes in real-time. This affords therapists the opportunity to allocate their attention and resources for effective facilitation. The objective was to determine whether AICF can detect group cohesion beyond the first-person plural pronunciation use. The findings indicate that it is feasible to measure group cohesion in text-based complex human interactions using AICF. The level of congruency with human scoring suggests that it can be a helpful tool to therapists in improving the group cohesion outcome.

This study has opened an avenue to person-centered and process-outcome research using AI combined with human inputs to improve the quality of care, which otherwise is a labor-intensive research process. Initially, after being trained with 1000 annotated group cohesion statements processed by word embeddings and the domain expertise from therapists, AICF was able to achieve reasonable F1, Precision and Recall scores. Furthermore, training the algorithm using only word embeddings allows AICF to identify the various cohesion themes that emerged, which are consistent with previous research [32, 34]. These themes include expressing support, reassurance, a sense of belonging, trust, deepening emotional disclosure, gratitude, remarking on shared experiences, reflecting on positive aspects of the group, and anticipating future chats. The findings suggest that training AICF to monitor therapeutic responses in virtual care is promising.

When combined with the human scoring examples in the algorithm, as little as 20% of the outputs, AICF obtained a high F1 score. The human rater detected both false positive examples (e.g. “Just have to find what works for you, I listen to a lot of audible books while I do chores, it's a mental distraction and really helps me”) and false negative examples (e.g. “Thanks so much to all of you, for being in this moment. You've helped me get ready for yet another week.”). These examples contributed to the rule-based algorithms as a second layer of analysis. While Precision values remained relatively low in both rounds (0.99 vs. 0.98), the Recall value improved from 0.52 to 0.70 due to a reduction in false negative classifications. These increases strongly suggest that a continuous effort to train AICF using human input can lead to a higher level of accuracy in detecting group cohesion.

After running LIWC on a test dataset, its performance was evaluated by a human scorer. The Precision, Recall, and F1 scores were lower compared to the performance of AICF. Unlike AICF, which is capable of identifying group cohesion expressions and idioms, LIWC is programmed to identify certain keywords. For a false negative example, LIWC was unable to detect the following quote as an instance of group cohesion due to the lack of the keyword ‘we’: “I feel like I’ve suddenly inherited a whole group of sisters.” Another instance of an LIWC false positive was that LIWC dictionary categories ‘family’, ‘positive emotion’, and ‘affiliation’ falsely detected group cohesion from this quote: “My husband has helped me see that it isn't something I did, or who I am.”

Comparison with Prior Work

The current study successfully trained a machine-learning system to detect cohesive statements in contrast to qualitative content analysis, which tends to be onerous and prone to human errors when dealing with large amounts of data [12]. Emerging computer programs such as Discourse Attributes Analysis Program (DAAP)[36] and LIWC [35, 37] offer an iterative psycholinguistic approach to coding transcripts of psychotherapy for therapeutic moments [19, 22]. For example, DAAP is based on a weighted dictionary that assigns weights to different words instead of solely detecting them as belonging to various categories where all matching words contribute equally to the generated scores. This method allows for greater accuracy in measuring different concepts compared to human coding while processing large amounts of data. However, these weighted dictionary approaches can be limited by a fixed number of instances that can be detected, and only one keyword can be considered in each matching rather than taking contexts into account. Also, they do not consider emerging words, phrases, idioms of expressions, word order, negation, and context-dependent factors, as well as their post-hoc nature [20]. In the current study, the word embedding approach was used to create contextual variables from the keywords to successfully detect a reasonably broad definition of cohesiveness. Thus, work will continue toward improving the accuracy of AICF in upcoming OSG sessions.

Limitations

AICF is based on a previously trained ensemble called Patient-Reported Information Multidimensional Exploration (PRIME) that was primarily trained on Australian online forum data [30]. Thus, Canadians may have used expressions/idioms that were unfamiliar to the original PRIME system and, therefore, not detected (e.g. “My head is swirling” to describe feeling overwhelmed or “the clock is ticking” to describe an impending end of life). The local idioms and expressions were

handled by the rule-based approach; ideally, AICF would be (re)trained with a large amount of local data in order to capture such idioms and expressions.

Currently, the interactional nature of the statement is not incorporated into AICF, including responses to other members' or therapists' statements. Furthermore, AICF cannot consistently distinguish whether participants are speaking about the group or about people outside of the group. When data accumulates, this distinction will become more obvious and refine AICF's detection ability within the context of an OSG.

The performance of AICF's group cohesion classification was evaluated in comparison to scores by two human experts, whose scoring was guided by the same criteria. However, given the nuanced nature of a group process like cohesion, there was still an element of personal judgment and openness to interpretation in the statements. Finally, emojis were not considered in the algorithm; future studies need to incorporate them as expressions of group cohesion.

Future Directions

AICF has been running in the background on three CCC groups and will soon be deployed for beta-testing on 10-12 groups. Participants will be filling out a survey package that includes the psychometrically validated questionnaire that tabs group cohesion for further validation. For algorithm development, sequencing the emotions expressed by each participant will be explored to capture more accurate emotional profiles.

The use of Large Language models (LLMs), such as ChatGPT, has revolutionized natural language understanding in the field of affective computing. Research suggests that an LLM called ROBERTa [38] has been equipped with emotion knowledge that contains 14 human conceptual attributes of emotions, including two affective, six appraisals, and six basic emotions. Future work will incorporate LLMs into our system to enhance AICF's ability to detect group cohesion and other significant clinical outcomes. For example, the LLM has already understood the syntactic difference between first-person and third-person pronoun uses and their contexts. Combining both of these emotional attributes and syntax, we are able to better formulate an equation to calculate the tendency of a writer to be self-focused or other-focused. This will truly improve the accuracy and precision of group cohesion detection.

Lastly, In this study, five LIWC dictionary categories were used to capture the concept of group cohesion. Future studies may test whether there is a way that will improve the performance of group cohesion prediction using LIWC by 1) adding more categories, 2) reducing some categories, and 3) adding weighting to each criterion.

AICF will explore ways to measure multiple processes comprising group climate, including the level of participation, expression of emotion, signs of cohesion, avoidance, and therapeutic factors such as conflict, altruism, universality, interpersonal learning input and output, catharsis, identification, self-understanding, and instillation of hope [5, 9, 32]. If successful, AICF will be applied alongside the mobile health Chatbot technology to provide a scalable, automated monitoring and referral system that screens users for specific symptoms, recommends individualized online and community resources, tracks each user's psychological outcomes through, and refers them to local therapists when necessary.

Conclusions

Optimal OSG delivery requires rapid alerts for therapists to effectively monitor markers of positive and negative responses within the group. This study has demonstrated that advanced machine-learning algorithms combined with human inputs can reasonably detect the clinically meaningful indicator of group cohesion in OSGs. Future research in utilizing LLMs in AICF could enhance the capabilities in understanding the context, given the capability of creating a highly customized model in a short time. Therefore, AICF has the potential to assist therapists by highlighting issues that are amenable to intervention in real-time, which allows therapists to provide greater levels of individualized support.

Acknowledgments

This research is funded by the Ontario Institute for Cancer Research Cancer Care Ontario Health Services Research Network.

The data sets generated during and/or analyzed during this study are not publicly available due to the presence of private health information but are available from the corresponding author on reasonable request.

Author Contributions

Y Leung: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing

E Wouterloot: Formal Analysis + Methodology

A Adikari: Visualization, Writing – original draft, Writing – review & editing, Formal Analysis, Software

J Hong: Formal Analysis + Methodology

V Asokan: Writing – original draft, Writing – review & editing

L Duan: Formal Analysis

C Lam: Formal Analysis

C Kim: Validation + Software

KP Chan: Validation + Software

DD Silva: Supervision + Funding acquisition

L Trachtenberg: Validation and Conceptualization

H Rennie: Validation and Conceptualization

J Wong: Writing – review & editing, Supervision, Data curation, Funding acquisition

MJ Esplen: Writing – review & editing, Supervision, Data curation, Funding acquisition

Conflicts of Interest

None declared.

Abbreviations

AICF: artificial intelligence co-facilitator

CCC: Cancer Chat Canada

DAAP: Discourse Attributes Analysis Program

JMIR: Journal of Medical Internet Research

LIWC: Linguistic Inquiry and Word Count

PRIME: Patient-Reported Information Multidimensional Exploration

NLP: natural language processing

OSG: online support group

TF-IDF: term frequency-inverse document frequency

References

- [1] D. Gratzner and D. Goldbloom, "Making Evidence-Based Psychotherapy More Accessible in Canada," *Can. J. Psychiatry*, vol. 61, no. 10, pp. 618–623, Oct. 2016, doi: 10.1177/0706743716642416.
- [2] D. A. Male, K. D. Fergus, and J. E. Stephen, "Professional Online Support Group Facilitators: Guarantors of Maximal Group Utility," *Int. J. Group Psychother.*, vol. 67, no. 3, pp. 314–336, Jul. 2017, doi: 10.1080/00207284.2016.1240587.
- [3] E. McCaughan, K. Parahoo, I. Hueter, L. Northouse, and I. Bradbury, "Online support groups for women with breast cancer," *Cochrane Database Syst. Rev.*, Mar. 2017, doi: 10.1002/14651858.CD011652.pub2.
- [4] C. F. van Uden-Kraan, C. H. Drossaert, E. Taal, B. R. Shaw, E. R. Seydel, and M. A. van de Laar, "Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia," *Qual Health Res*, vol. 18, no. 3, pp. 405–17, Mar. 2008, doi: 10.1177/1049732307313429.
- [5] I. D. Yalom, *The theory and practice of group psychotherapy*, 5th ed. in The theory and practice of group psychotherapy, 5th ed. New York, NY, US: Basic Books/Hachette Book Group, 2005.
- [6] W. G. Anderson *et al.*, "'What concerns me is...' Expression of emotion by advanced cancer patients during outpatient visits," *Support. Care Cancer Off. J. Multinat. Assoc. Support. Care Cancer*, vol. 16, no. 7, pp. 803–811, Jul. 2008, doi: 10.1007/s00520-007-0350-8.
- [7] J. Roos and A. Werbart, "Therapist and relationship factors influencing dropout from individual psychotherapy: a literature review," *Psychother. Res. J. Soc. Psychother. Res.*, vol. 23, no. 4, pp. 394–418, 2013, doi: 10.1080/10503307.2013.775528.
- [8] G. Clarke *et al.*, "Overcoming Depression on the Internet (ODIN) (2): A Randomized Trial of a Self-Help Depression Skills Program With Reminders," *J. Med. Internet Res.*, vol. 7, no. 2, p. e16, Jun. 2005, doi: 10.2196/jmir.7.2.e16.
- [9] J. E. Johnson, G. M. Burlingame, J. A. Olsen, D. R. Davies, and R. L. Gleave, "Group Climate, Cohesion, Alliance, and Empathy in Group Psychotherapy: Multilevel Structural Equation Models," *J. Couns. Psychol.*, vol. 52, no. 3, pp. 316–321, 2005, doi: <http://dx.doi.org.myaccess.library.utoronto.ca/10.1637/0022-0167.52.3.310>.
- [10] S. H. Budman, S. Soldz, and *et al.*, "Cohesion, Alliance and Outcome in Group Psychotherapy," *Psychiatry Wash.*, vol. 52, no. 3, pp. 339–50, Aug. 1989.
- [11] T. Treadwell, N. Lavertue, V. K. Kumar, and V. Veeraraghavan, "The Group Cohesion Scale-Revised: Reliability and validity," *Int. J. Action Methods Psychodrama Ski. Train. Role Play.*, vol. 54, no. 1, pp. 3–12, 2001.
- [12] C. A. S. Sierra Hernandez, W. E. Piper, J. S. Ogrodniczuk, A. S. Joyce, and R. Weideman, "Use of referential language in short-term group psychotherapy for complicated grief," *Group Dyn. Theory Res. Pract. Wash.*, vol. 20, no. 1, p. 1, Mar. 2016.
- [13] S. K. Fineberg *et al.*, "Self-reference in psychosis and depression: a language marker of illness," *Psychol. Med.*, vol. 46, no. 12, pp. 2605–2615, 2016, doi: 10.1017/S0033291716001215.
- [14] J. W. Pennebaker and L. A. King, "Linguistic styles: language use as an individual difference," *J. Pers. Soc. Psychol.*, vol. 77, no. 6, pp. 1296–1312, Dec. 1999, doi: 10.1037//0022-3514.77.6.1296.
- [15] A. M. Tackman *et al.*, "Depression, negative emotionality, and self-referential language: A

- multi-lab, multi-measure, and multi-language-task research synthesis.,” *J. Pers. Soc. Psychol.*, vol. 116, no. 5, pp. 817–834, May 2019, doi: 10.1037/pspp0000187.
- [16] L. M. V. S. Kane Aimée A., “Language and Group Processes: An Integrative, Interdisciplinary Review - Lyn M. Van Swol, Aimée A. Kane, 2019,” *Small Group Res.*, Jul. 2018, Accessed: Sep. 19, 2020. [Online]. Available: <https://journals-sagepub-com.myaccess.library.utoronto.ca/doi/10.1177/1046496418785019>
- [17] W. E. Piper and W. E. McCallum, “The Psychodynamic Work and Object Rating System,” in *The process of group psychotherapy: Systems for analyzing change*, Washington, DC, US: American Psychological Association, 2000, pp. 263–281. doi: 10.1037/10378-010.
- [18] “dtSearch,” *dtSearch*, 1991. <https://dtsearch.com/> (accessed Sep. 19, 2020).
- [19] T. Brockmeyer *et al.*, “Me, myself, and I: self-referent word use as an indicator of self-focused attention in relation to depression and anxiety,” *Front. Psychol.*, vol. 6, Oct. 2015, doi: 10.3389/fpsyg.2015.01564.
- [20] G. W. Alpers *et al.*, “Evaluation of computerized text analysis in an Internet breast cancer support group,” *Comput. Hum. Behav.*, vol. 21, no. 2, pp. 361–376, 2005, doi: 10.1016/j.chb.2004.02.008.
- [21] M. A. Lieberman, A. Wizlenberg, M. Golant, and M. Di Minno, “The impact of group composition on Internet support groups: Homogeneous versus heterogeneous Parkinson’s groups,” *Group Dyn. Theory Res. Pract.*, vol. 9, no. 4, pp. 239–250, 2005, doi: 10.1037/1089-2699.9.4.239.
- [22] R. Mariani, B. Maskit, W. Bucci, and A. De Coro, “Linguistic measures of the referential process in psychodynamic treatment: the English and Italian versions,” *Psychother. Res. J. Soc. Psychother. Res.*, vol. 23, no. 4, pp. 430–447, 2013, doi: 10.1080/10503307.2013.794399.
- [23] Y. Leung *et al.*, “An Extra Pair of Eyes: A Protocol for Developing an Artificial Intelligence-based Virtual Co-Facilitator for Online Cancer Support Groups,” *J. Med. Internet Res.*, Provisionally Accepted.
- [24] R. Plutchik, “Chapter 1 - A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION,” in *Theories of Emotion*, R. Plutchik and H. Kellerman, Eds., Academic Press, 1980, pp. 3–33. doi: 10.1016/B978-0-12-558701-3.50007-7.
- [25] Y. W. Leung *et al.*, “Natural Language Processing–Based Virtual Cofacilitator for Online Cancer Support Groups: Protocol for an Algorithm Development and Validation Study,” *JMIR Res. Protoc.*, vol. 10, no. 1, p. e21453, Jan. 2021, doi: 10.2196/21453.
- [26] Y. Leung *et al.*, “Providing care beyond the therapy session - a natural language processing-based recommender system that identifies cancer patients who experience psychosocial challenges and provides self-care support.” 2021. doi: 10.2196/preprints.35893.
- [27] T. Bandaragoda *et al.*, “The Patient-Reported Information Multidimensional Exploration (PRIME) Framework for Investigating Emotions and Other Factors of Prostate Cancer Patients with Low Intermediate Risk Based on Online Cancer Support Group Discussions,” *Ann. Surg. Oncol.*, vol. 25, no. 6, pp. 1737–1745, Jun. 2018, doi: 10.1245/s10434-018-6372-2.
- [28] J. Hu, F. Jin, G. Zhang, J. Wang, and Y. Yang, “A User Profile Modeling Method Based on Word2Vec,” *2017 IEEE Int. Conf. Softw. Qual. Reliab. Secur. Companion QRS-C*, 2017, doi: 10.1109/QRS-C.2017.74.
- [29] D. De Silva *et al.*, “Machine learning to support social media empowered patients in cancer care and cancer treatment decisions,” *PLOS ONE*, vol. 13, no. 10, p. e0205855, Oct. 2018, doi: 10.1371/journal.pone.0205855.
- [30] W. Ranasinghe, T. Bandaragoda, D. De Silva, and D. Alahakoon, “A novel framework for automated, intelligent extraction and analysis of online support group discussions for cancer related outcomes,” *BJU Int.*, vol. 120, pp. 59–61, Nov. 2017, doi: 10.1111/bju.14036.
- [31] R. R. Dies and K. R. MacKenzie, Eds., *Advances in group psychotherapy: integrating research and practice*. in Monograph series / American Group Psychotherapy Association, no.

- monograph 1. New York, N.Y: International Universities Press, 1983.
- [32] A. J. Borek *et al.*, “Developing and applying a framework to understand mechanisms of action in group-based, behaviour change interventions: the MAGI mixed-methods study,” *Effic. Mech. Eval.*, vol. 6, no. 3, pp. 1–162, Jun. 2019, doi: 10.3310/eme06030.
- [33] C. Hutto and E. Gilbert, “VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text,” *Proc. Int. AAAI Conf. Web Soc. Media*, vol. 8, no. 1, pp. 216–225, May 2014, doi: 10.1609/icwsm.v8i1.14550.
- [34] A. A. Cota, K. L. Dion, and C. R. Evans, “A reexamination of the structure of the Gross Cohesiveness Scale,” *Educ. Psychol. Meas.*, vol. 53, no. 2, pp. 499–506, 1993, doi: 10.1177/0013164493053002019.
- [35] Y. R. Tausczik and J. W. Pennebaker, “The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods,” *J. Lang. Soc. Psychol.*, vol. 29, no. 1, pp. 24–54, Mar. 2010, doi: 10.1177/0261927X09351676.
- [36] B. Maskit and S. Murphy, “The Discourse Attributes Analysis Program (DAAP),” 2016. [http://www.thereferentialprocess.org/the-discourse-attributes-analysis-program-daap#:~:text=The%20Discourse%20Attributes%20Analysis%20Program%20\(DAAP\)%20is%20a%20computer%2D,on%20an%20exponential%20smoothing%20operator](http://www.thereferentialprocess.org/the-discourse-attributes-analysis-program-daap#:~:text=The%20Discourse%20Attributes%20Analysis%20Program%20(DAAP)%20is%20a%20computer%2D,on%20an%20exponential%20smoothing%20operator).
- [37] J. W. Pennebaker, T. J. Mayne, and M. E. Francis, “Linguistic predictors of adaptive bereavement,” *J. Pers. Soc. Psychol.*, vol. 72, no. 4, pp. 863–871, Apr. 1997, doi: 10.1037//0022-3514.72.4.863.
- [38] M. Li *et al.*, “Human Emotion Knowledge Representation Emerges in Large Language Model and Supports Discrete Emotion Inference,” 2023, doi: 10.48550/ARXIV.2302.09582.

Supplementary Files

Untitled.

URL: <http://asset.jmir.pub/assets/ab9737ad976cb56301339a88aeff3591.docx>