Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-025-05841-5 OPEN

Check for updates

Exploring film therapy in digital health: text mining study of Google search data

Guoqing Wang[1], Yang Yu[1], Yipei Wang[2] & Kamal Sabran[1]✉

Against the backdrop of the rapid development of digital health technology and the increasing prominence of mental health problems, film therapy as an innovative intervention method is gaining more and more attention. This study systematically collected information related to film therapy on the Google search platform and used multidimensional text mining technology to conduct an in-depth analysis of data sources, aiming to reveal the presentation characteristics and development trends of film therapy in cyberspace. Eight core topic categories were identified using the latent Dirichlet allocation (LDA) model, covering a complete spectrum from innovative treatment methods to diversified application scenarios, among which character emotional resonance and professional guidance became the two most prominent themes. Sentiment calculation analysis found that positive attitudes dominated the online discussions, with up to three-quarters of the content expressing positive evaluations and no negative emotional content appearing, reflecting the public's widespread recognition of this treatment method. The results of the semantic association network construction showed that key concepts such as "therapy," "movie," "use," and "people" formed a close semantic cluster, reflecting the inherent logic of the film therapy theory system and the systematic nature of the practice framework. Vocabulary frequency statistics further confirmed the high integration characteristics of treatment terms and film and television terms, indicating that this interdisciplinary field has formed a relatively mature discourse system. The research findings provide important empirical support and theoretical guidance for the standardized development of film therapy, the promotion of clinical applications, and the construction of a digital mental health service system.

[1]Universiti Sains Malaysia, Gelugor, Malaysia. [2]Universiti Putra Malaysia, Serdang, Malaysia. ✉email: kamalsabran@usm.my

Introduction

Mental health issues have become a major global public health challenge, especially among young people (Cooney, 2023). With the increasing number of mental health problems among young people around the world, mental health has become a top priority in the field of global health (Panayiotou et al., 2023). Although various mental health workers continue to work hard to promote the mental health of individuals and families (Yan et al., 2024), traditional psychotherapy still faces many challenges. Some patients feel ashamed or depressed during treatment, and it is difficult to understand the treatment content or communicate with the therapist (Reupert, 2024). With 76–85% of people in low-income countries unable to access mental health treatment due to barriers of cost and stigma, several alternative approaches offer viable and cost-effective solutions to address this care gap (Kolenik and Gams, 2021b). Especially in the context of the rapid development of Internet technology in the twenty-first century, the use of new technologies to carry out mental health interventions has become a research area that has attracted much attention (Marciano et al., 2024). Technical intervention has advantages such as cost reduction, round-the-clock accessibility, reduction of stigmatization, and promotion of prevention (Kolenik et al., 2024).

The advancement of digital technology has made digital health an important means to improve the level of medical services. Digital health covers various applications that use digital technology to provide objective data to clinicians, caregivers, and service recipients to improve medical outcomes (Geifman et al., 2023). In recent years, digital health has shown a significant growth trend worldwide and is considered by multiple stakeholders to be a necessary innovation to improve the quality and economic benefits of medical services (Kim and Choi, 2023). In the field of digital health, digital interventions for mental health problems have received increasing attention. Studies have shown that virtual communities have played an important role in maintaining people's mental health during the COVID-19 pandemic (Golz et al., 2022). Through digital mental health interventions, children and adolescents can improve sleep problems and reduce mental health symptoms, which are especially important for the physical and mental health of contemporary young groups (Lawrence-Sidebottom et al., 2024). Digital mental health is gradually becoming an important supplement to the existing mental health care system, effectively providing help to those in need (Kolenik, 2022). Various technologies enhance mental health interventions through: (1) machine learning-based assessment models using Random Forest and neural networks for symptom classification, (2) personalized intervention systems incorporating cognitive behavioral therapy techniques and motivational interviewing, and (3) conversational AI platforms that provide real-time, accessible mental health support through text-based interactions (Kolenik and Gams, 2021a). Among various digital mental health interventions, film therapy as an emerging treatment method is receiving increasing attention.

As an art form with high popularity and relatively low cost, films have shown unique advantages in mental health interventions (Goodwin et al., 2021). Studies have shown that mental health education through film media can effectively reduce stigma and improve the public's understanding and acceptance of mental health issues (Kao et al., 2023). In particular, films that present mental health conditions in a humane way can help the audience build empathy and reduce prejudice (Pirkis et al., 2006). Studies have shown that film therapy can effectively improve a variety of mental health problems, such as anxiety, depression, and eating disorders (Koushiou et al., 2018), and has achieved remarkable results in the field of mental health intervention in recent years. In addition, film therapy has the advantages of high cost-effectiveness and easy implementation and is particularly suitable for college students with limited economic conditions (Du-Nan et al., 2020).

The development of the Internet and social media has made user-generated content an important data source for studying public opinion. However, how to effectively extract valuable information and insights from massive unstructured text data has become an important challenge in current research (Farkhod et al., 2021). As an unsupervised machine learning method, topic modeling can automatically discover the potential topic structure in a text collection. Among them, the LDA model has become one of the most widely used topic modeling methods due to its simplicity and effectiveness (Shi et al., 2009). As an important branch of natural language processing, sentiment analysis is dedicated to identifying emotional tendencies in texts. Semantic network analysis based on tools such as ROST Content Mining System Version 6.0 (ROST CM6.0) can reveal the co-occurrence relationship and semantic distance between keywords, helping researchers better understand the deep semantics of the text (Sun et al., 2022). In recent years, the method of combining topic modeling with sentiment analysis has received widespread attention (He et al., 2024).

Although film therapy shows good application prospects, there is still a lack of systematic research on its use as a mental health intervention. Especially in the Internet age, people are increasingly obtaining mental health-related information through online platforms (Chen et al., 2020). As the world's largest search engine, Google plays an important role in disseminating and influencing people's access to mental health information (Birnbaum et al., 2018). Therefore, it is of great significance to systematically evaluate the quality and reliability of information on film-based mental health interventions in *Google search results*.

To systematically evaluate the information on film-based mental health interventions in *Google search results*, this study comprehensively used methods such as LDA topic modeling, sentiment analysis, and semantic network analysis to conduct an in-depth analysis of *Google search results* related to film therapy. The multidimensional analysis method can not only discover the potential topic structure in the text but also reveal the sentiment tendency and semantic association under different topics. This study will not only help understand the current status of film therapy-related information in the current online environment but will also provide an important reference for improving the quality of online mental health resources. At the same time, the research results can provide valuable guidance for mental health professionals, policymakers, and the public.

Methods

The study was divided into six steps: (1) Collecting film therapy-related topics through *Google search results*. (2) Extract and filter through the Octoparse tool and remove duplicate websites (Octoparse, 2025); (3) Segment the collected text into words; (4) Construct public opinion social network through LDA topic analysis of content; (5) Public opinion is evaluated based on the dual dimensions of word frequency statistics and sentiment analysis; (6) The ROST CM6.0 semantic network analysis tool is used to visualize the association relationship between the main feature words. The specific steps are as follows:

Data collection. On July 26, 2024, a search was conducted using the terms "film therapy," "cinema therapy," "video therapy," and "movie therapy" to obtain *Google search results* via www.google.com. These search terms were selected based on their common usage in academic

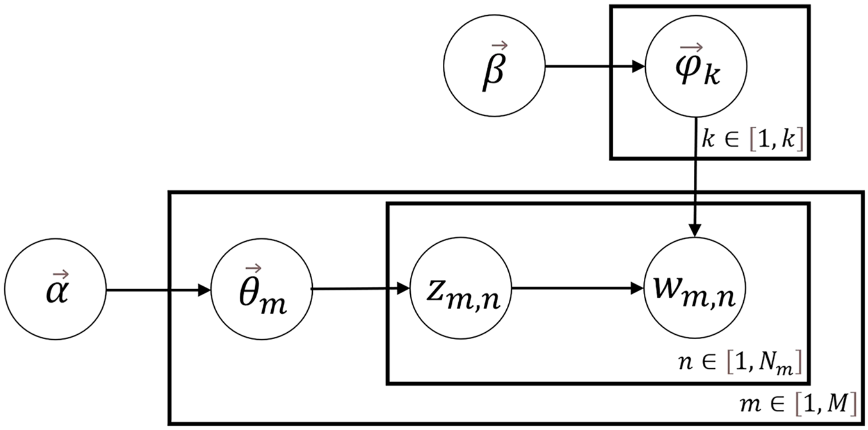

Fig. 1 LDA model structure.

literature and clinical practice to capture the broadest range of relevant content.

This study used the Octoparse tool to extract the content of 67 pages (130 sources) and imported them into Microsoft Excel for analysis. The Octoparse configuration included the following specifications: (1) extraction of title, URL, description, and full text content from each search result; (2) duplicate removal was implemented using URL matching—any pages with identical URLs were automatically excluded; (3) content filtering excluded non-English pages, advertisement-only content, and pages with fewer than 100 words to ensure substantive content analysis.

Because the number of pages searched for each keyword was less than 20, this study comprehensively screened and analyzed all pages returned for each keyword to maximize data coverage. The page selection criteria included: (1) relevance to film therapy; (2) accessibility (non-paywalled content); (3) substantial textual content (minimum 100 words); and (4) English language content. After eliminating duplicate topics through automated URL matching and manual verification of content similarity, the final dataset provided the data foundation for further processing, analysis, and discussion.

Data processing. In the data processing stage, this study mainly uses Python and ROST CM6.0 for data analysis. In the Python environment, we first segmented the document and used the English stop word list and the custom stop word list of the Natural Language Toolkit (NLTK) library to effectively remove irrelevant words in the document. Then, we used the Gensim library to create dictionaries and corpora and converted the processed documents into a bag-of-words model to lay the foundation for subsequent LDA analysis. At the same time, we used the English word segmentation function of ROST CM6.0 to process *Google search results* and provide data support for subsequent semantic network and sentiment analysis.

Data analysis

Data analysis of LDA. In this study, we used the LDA model for text topic modeling analysis. LDA is a topic modeling method used to discover latent topic structures from a collection of documents (Blei et al., 2003). The most evident feature of LDA is that it can encode and classify documents into multiple topics.

As shown in Fig. 1, the LDA model consists of two main components: topic level and document level. At the topic level, the model assumes that there are K topics, and each topic φ_k represents a probability distribution over the vocabulary, generated by a Dirichlet distribution controlled by a hyperparameter β. At the document level, the corpus contains M documents, each with N_m words. For each document, θ_m represents the distribution of the document over K topics, generated by a Dirichlet distribution controlled by a hyperparameter α.

The generation process of the model is: for each word position in the document, first select a topic $z_{m,n}$ according to the topic distribution θ_m of the document, and then generate the observed word $w_{m,n}$ according to the word distribution φ_k of the topic. Through this probabilistic generation model, LDA can infer the topic composition of each document and the word distribution contained in each topic, thereby realizing the topic modeling and semantic analysis of the document.

It is worth noting that the number of topics here needs to be manually set with the help of perplexity and coherence curves. As shown in (1), the mathematical expression of the perplexity of the topic model was calculated (Lv et al., 2022). Through the key indicator of perplexity, we can evaluate the prediction ability of the topic model for new documents and thus judge the model's performance. In the formula: $log_P(w_d)$ is the log-likelihood probability of all word sequences in document d, indicating the model's ability to predict document d. N_d represents the number of words in document d, and M represents the total number of documents. In general, the lower the perplexity, the better the model performance. The perplexity obtained by calculation was 8 (as shown in Fig. 2). This relatively low perplexity indicates that the model has good topic prediction ability (Han et al., 2022), which provides a reliable basis for subsequent analysis.

$$Perplexity = \exp\left\{-\frac{\sum_{d=1}^{M} log_P(w_d)}{\sum_{d=1}^{m} N_d}\right\} \tag{1}$$

The coherence of the LDA model is an important indicator for evaluating the quality of the topic model. Coherence has several calculation methods: UMASS, C_V, and UCI. In this study, we use the UMASS method. This method is based on document frequency and word frequency information, and evaluates coherence by calculating the similarity of words in the topic. The higher the coherence score, the higher the similarity of words within the topic, and the more focused the topic. The calculation formula is as follows:

$$Coherence_{umass}(wi, wj) = \log\frac{D(wiwj) + 1}{D(wi)} \tag{2}$$

Among them (2), $D(wiwj)$ is the number of documents in which words wi and wj appear together, and $D(wi)$ is the number of documents in which word wi appears. If wi and wj appear together in documents more often than they appear alone, the

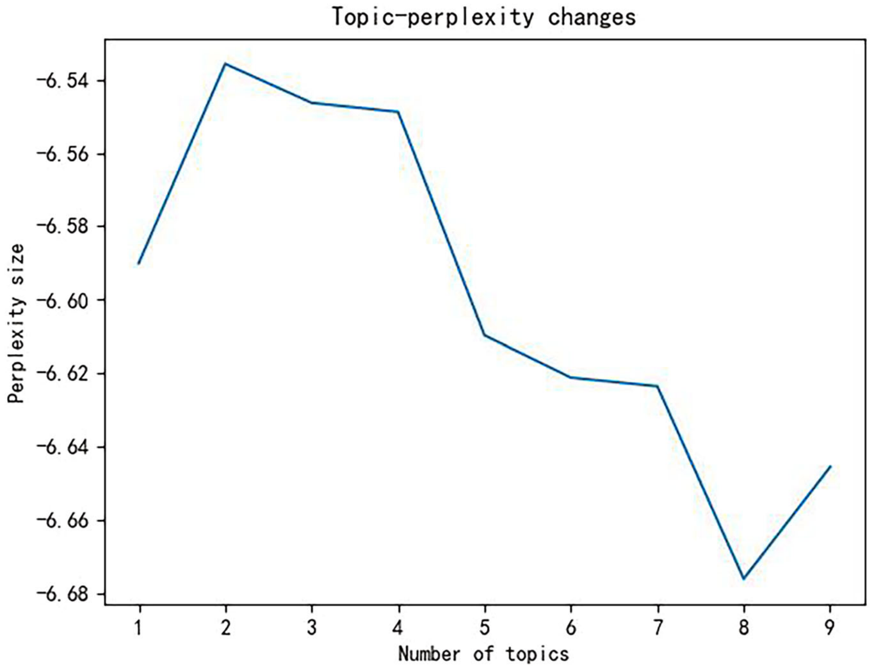

Fig. 2 Topic-perplexity changes.

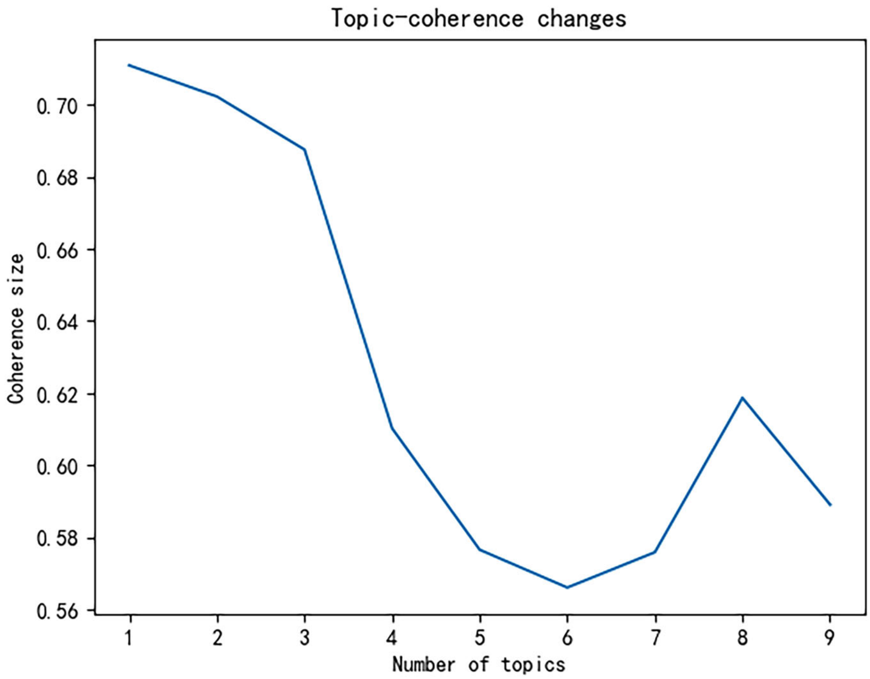

Fig. 3 Topic-coherence changes.

score is higher, which makes sense as a measure of topic coherence.

Figure 2 presents the perplexity change curves across different topic numbers, reflecting the relationship between the number of topics and perplexity values. The perplexity reaches the lowest value when there are 8 topics (perplexity = 8). Figure 3 presents the topic coherence change curve, demonstrating how the UMass coherence score varies with the number of topics. The coherence score is high at the beginning (0.72), and then decreases as the number of topics increases, reaching the lowest point at 5–6 topics, and rising back to 0.62 at 8 topics. Combining the coherence and perplexity indicators, the eight topics have the lowest perplexity while maintaining good coherence, indicating that they have achieved the best balance between topic distinctiveness and model fitting effect. After we set the number of topics, we can use the computing power of the LDA algorithm to obtain the topic distribution, document distribution, and other results.

Word frequency and sentiment analysis. During the research process, we performed English word segmentation and invalid word filtering on the film therapy webpage content collected from *Google search results* and used ROST CM6.0 software to implement word frequency statistics and sentiment analysis. ROST

CM6.0 can quickly analyze massive amounts of network text (Liu et al., 2017). Word frequency analysis is a traditional content analysis method in bibliometrics. Its basic principle is to determine hot spots and their changing trends by counting the frequency of word occurrences and analyzing the number of occurrences of important words. It is an important means of text mining (Wang et al., 2024).

In the field of humanities and social sciences research in China, ROST CM6.0 occupies a unique position due to its characteristics as the only large-scale free social computing platform (Chen et al., 2020). ROST CM6.0 allows users to customize dictionaries so that they can accurately identify and extract relevant high-frequency words and sentiment words for specific research. This feature is critical for ensuring that sentiment analysis accurately reflects the emotional tone of the text, whether positive, negative, or neutral (Chen et al., 2023). The emotional intensity threshold setting adopts a seven-level classification standard, covering the three emotional dimensions of positive, neutral, and negative. Positive emotions are divided into three levels: general positive ($0 < \text{score} \le 15$), moderate positive ($15 < \text{score} \le 25$) and highly positive ($25 < \text{score} < +\infty$); neutral emotions are defined as a balanced state with a score equal to 0; negative emotions are divided into three levels: general negative ($-15 \le \text{score} < 0$), moderate negative ($-25 \le \text{score} < -15$) and highly negative ($-\infty < \text{score} < -25$). This hierarchical threshold system can accurately quantify emotional expressions of different types and intensities, providing a standardized evaluation framework for emotional analysis.

To ensure the reliability of the ROST CM6.0 sentiment analysis results, this study used a dual validation mechanism for systematic validation. Cross-validation machine learning method: TF-IDF feature extraction and logistic regression classifier were used to build a prediction model (Li, 2012). TF-IDF Vectorizer was used for feature extraction (maximum number of features = 1000, English stop words were removed), and the coherence between the ROST analysis results and the machine learning algorithm prediction results was evaluated by 5-fold cross-validation. Sentiment threshold analysis method: Valence Aware Dictionary and Sentiment Reasoner (VADER) sentiment analyzer was used for independent sentiment polarity analysis. VADER is specifically used for social media text sentiment analysis. The optimal classification threshold was determined by systematically testing different thresholds (0.0–0.2 range) and then compared with the ROST CM6.0 results for validation. The validation dataset was constructed based on 130 source texts collected. The two validation methods were independent of each other, and both used automated processing procedures to ensure the objectivity and repeatability of the results.

Through the analysis of *Google search results*, we completed three tasks: keyword frequency statistics, sentiment analysis of webpage topics, and quantitative calculation of sentiment scores.

Semantic network analysis. This study used ROST CM6.0 to perform semantic network analysis of the texts. The system can perform accurate semantic network analysis on a variety of texts, such as microblogs, papers, and chat records, and can not only generate keyword matrices but also draw semantic networks (Gao et al., 2021). A semantic network is a graphical representation of knowledge based on meaningful relationships between written texts, structured as a network of words that are cognitively related to each other (Alhajj and Rokne, 2014), and in this study, it refers to film therapy-related information. In a semantic network, nodes are words that represent concepts in a text, such as "therapy," "movie," "people," and other core concepts. The connections between nodes are called edges, which represent the relationship between connected concepts. For example, the association between "movie" and "therapy" reflects the conceptual connection of movies as a therapeutic tool. Semantic networks extract meaningful ideas by identifying emerging concept clusters rather than analyzing the frequency of isolated words (Doerfel, 1998). In this way, analyzing the film therapy webpage content from Google search can enhance the understanding of complex mental health intervention behaviors, especially the understanding of the application mechanism and effect cognition of film therapy.

We first performed word segmentation and stop word processing on the data containing all the film therapy content from *Google search results*. Then, through the "Social Network and Semantic Network Analysis" module in ROST CM6.0 "Functional Analysis," the row feature word list, co-occurrence semantic network, and co-occurrence matrix word list were generated in turn to reveal the deep semantic associations and network structure characteristics between concepts in the field of film therapy.

Results

The LDA verification results showed good model reliability and theoretical validity. Computational verification results: As shown in Figs. 2 and 3, the relatively low value of perplexity of 8 indicates that the model has good topic prediction ability, providing a reliable basis for subsequent analysis. UMASS coherence analysis determined that 8 topics were the optimal number of topics, with the lowest perplexity while maintaining good coherence, indicating that it achieved the best balance between topic discrimination and model fitting effect. Theoretical comparison verification results: The 8 identified topics show a high degree of conceptual correspondence with the existing film therapy theoretical framework, among which "innovative film psychotherapy" and "audience-participatory therapy" extend the Berg-Cross classification system, and "narrative therapy effect" and "therapist guidance" are highly consistent with the Powell mechanism framework, verifying the theoretical contribution value of topic discovery.

Topic modeling via LDA. By analyzing the texts through topic modeling and comparing the perplexity and coherence index, we found that grouping the topics into eight topic categories was the best option. Thus, based on these findings, we identified the following eight topics (see Fig. 4 and Table 1). Figure 4 presents the thematic distance distribution based on multidimensional scaling analysis, showing the semantic relationships among the eight film therapy topics in a two-dimensional space. The size of the circle indicates the importance of the topic in the corpus, and the spatial position reflects the conceptual similarity between the topics. The clear separation between the topics verifies the effectiveness of topic identification, and the clustering pattern reveals the conceptual association.

Topic 1: Innovative psychotherapy based on films. This topic focuses on the innovative use of movies in mental health treatment. Keywords include therapy, cinema, mental health, patients, and addiction, reflecting the potential application of this new treatment method for mental health and addiction problems. Frequently appearing word combinations show the innovativeness and practicality of this method.

Topic 2: Emotion and relationship interaction. This topic explores the development of emotional experiences and interpersonal relationships. Core vocabulary includes help, feel, experience, relationship, and emotions. The content of the topic emphasizes the importance of promoting emotional understanding and relationship building through experiential learning.

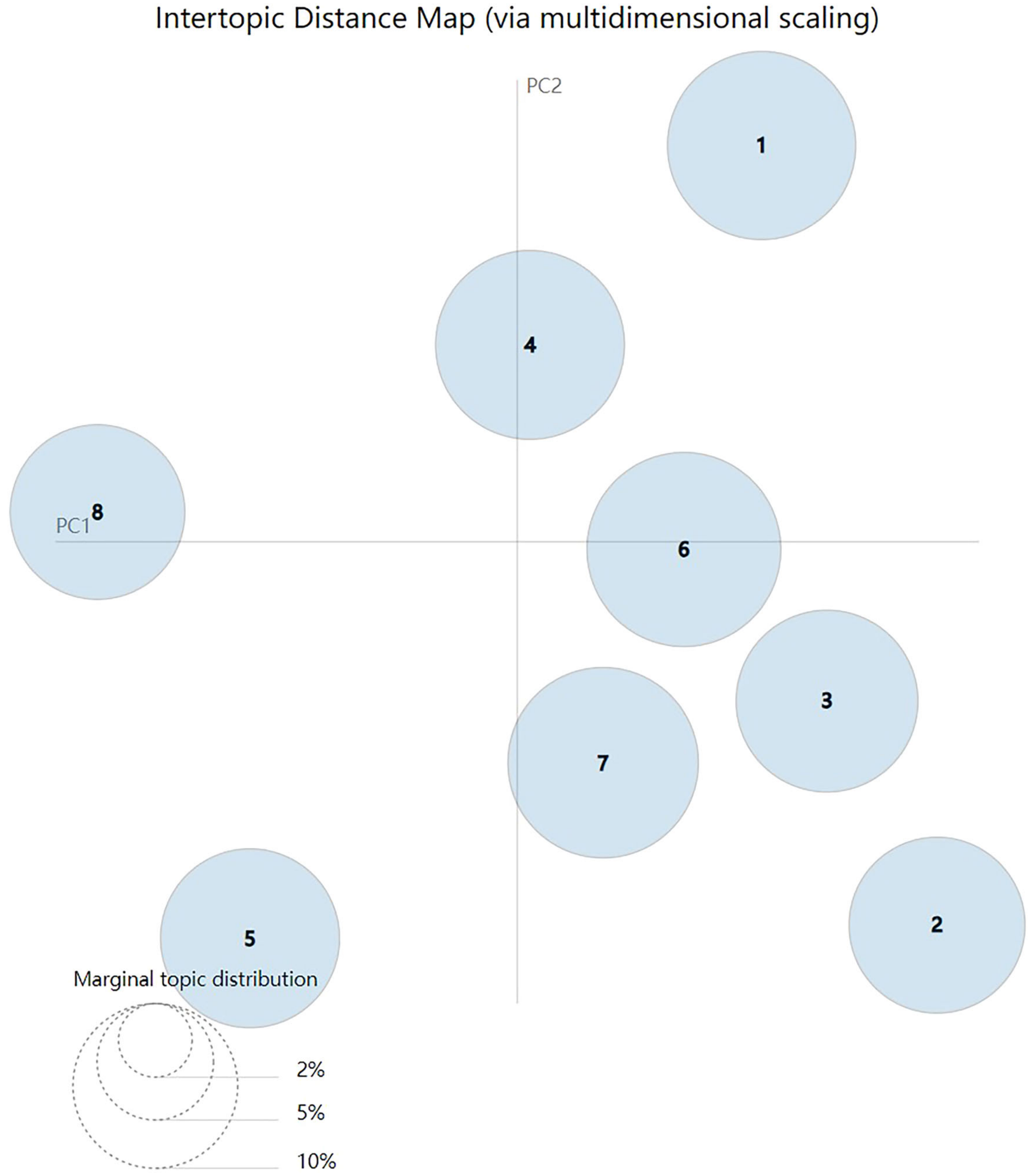

Fig. 4 Topic distance distribution.

Topic 3: Audience-participated therapy. This topic focuses on the interaction between the audience and film therapy. Key terms include cinema, viewers, patients, therapy, and use, reflecting the importance and practical methods of audience participation in film therapy.

Topic 4: Character emotional resonance. This topic explores the establishment of emotional connections through film characters. The keywords include emotions, characters, therapeutic, and group, reflecting the method of using character emotional resonance to promote therapeutic effects.

Topic 5: Therapeutic movie selection. This topic focuses on the importance of movie selection in therapy. Keywords include movie, choose, therapy, and helping, emphasizing the impact of choosing the right movie on the therapeutic effect.

Topic 6: Narrative therapy effect. This topic explores the role of film narrative in therapy. The core words include films, therapeutic, story, and clients, reflecting the method of achieving therapeutic goals through storytelling.

Topic 7: Therapist's guidance. This theme highlights the role of therapists in film analysis. Keywords include analysis, therapist, feel, and know, emphasizing the importance of professional guidance in film therapy.

Topic 8: Multidimensional therapeutic applications. This topic focuses on the application of film therapy in multiple scenarios. Keywords include mental health, group, family, and individual, showing the applicability of film therapy in different therapeutic environments.

Table 1 *LDA* topic modeling results.

Topic #1		Topic #2		Topic #3		Topic #4	
Keywords	**Weight**	**Keywords**	**Weight**	**Keywords**	**Weight**	**Keywords**	**Weight**
therapy	0.0199	help	0.0106	make	0.0091	help	0.0153
cinema	0.0172	see	0.0098	cinema	0.0088	therapy	0.0142
people	0.0084	people	0.0088	find	0.0077	emotions	0.0116
new	0.0060	feel	0.0086	help	0.0073	characters	0.0102
idea	0.0059	experience	0.0072	new	0.0062	films	0.0095
movie	0.0057	relationship	0.0059	viewers	0.0059	mental	0.0095
mental	0.0051	learn	0.0059	patients	0.0053	emotional	0.0090
patients	0.0051	think	0.0059	people	0.0053	cinema	0.0076
work	0.0051	different	0.0055	use	0.0048	therapeutic	0.0069
addiction	0.0050	emotional	0.0047	therapy	0.0048	group	0.0067
Topic #5		**Topic #6**		**Topic #7**		**Topic #8**	
Keywords	**Weight**	**Keywords**	**Weight**	**Keywords**	**Weight**	**Keywords**	**Weight**
movie	0.0218	films	0.0162	therapy	0.0209	mental	0.0167
films	0.0100	therapy	0.0136	movie	0.0175	health	0.0142
watch	0.0100	use	0.0128	analysis	0.0127	group	0.0097
therapy	0.0063	movie	0.0118	different	0.0079	therapy	0.0092
good	0.0058	people	0.0107	help	0.0076	cinema	0.0085
choose	0.0055	life	0.0066	people	0.0071	good	0.0074
help	0.0044	therapeutic	0.0062	know	0.0065	movie	0.0069
helping	0.0044	story	0.0062	feel	0.0065	therapeutic	0.0069
need	0.0042	new	0.0060	therapist	0.0064	family	0.0062
asked	0.0041	clients	0.0060	cinema	0.0059	individual	0.0055

Table 2 Frequency distribution of characteristic words.

No.	Characteristic word	Frequency	No.	Characteristic word	Frequency
1	therapy	951	1	feel	176
2	movie	715	2	will	176
3	people	385	3	use	176
4	cinema	379	4	work	168
5	help	373	5	clients	166
6	life	368	6	watch	165
7	films	336	7	emotional	161
8	mental	257	8	group	160
9	characters	234	9	make	147
10	therapist	216	10	process	145
11	health	199	11	client	145
12	time	193	12	find	143
13	emotions	188	13	see	143
14	therapeutic	185	14	feelings	142
15	experience	177	15	self	141

Word frequency statistics (top 30). The word frequency statistics of the dataset in this study revealed a corpus feature centered on psychotherapy and movies (see Table 2). The data showed that the most frequent words were therapy (951 times), followed by movie (715 times), and people (385 times). Other high-frequency words included cinema (379 times), help (373 times), and life (368 times). The distribution of these keywords clearly reflects the core content of the research topic.

From the perspective of semantic classification, these high-frequency words are mainly distributed in four subject areas: mental health-related (such as therapy, help, mental), film and media related (such as movie, cinema, films), interpersonal and life experience (such as people, life, emotions), and time and process related (such as time, process, use). This distribution feature shows that the research content mainly revolves around film therapy and its application in the field of mental health.

In addition, some verbs such as "feel," "will," "use," and "make" also appear at higher frequencies, which may indicate that the dataset contains discussions about personal experiences, future plans, or specific application methods. This word frequency distribution strongly suggests that the dataset may be closely related to film therapy and the application of movies in the field of mental health. At the same time, it also reflects people's widespread interest in using movies as a means of treatment or auxiliary treatment, and how this method affects people's lives, emotions, and mental health.

Sentiment analysis. The "Emotional Analysis" function in the "Functional Analysis" of the ROST CM6.0 software was used to perform emotional analysis on the text data. The obtained emotional analysis results are shown in Table 3.

Sentiment analysis conducts an in-depth study of texts, and the results show that these texts exhibit obvious emotional tendencies. Through the analysis of the data in the table, it can be concluded that the *online content creators* have a significantly positive attitude towards the topic discussed, with positive emotions accounting for 75.38%, which indicates that the content contributors are generally quite satisfied with the topic manner.

In terms of sentiment distribution, positive sentiments (98, 75.38%) were dominant, followed by neutral sentiments (32, 24.62%). It is worth noting that no negative sentiment was found in the analysis (0, 0.00%). This sentiment distribution pattern clearly shows the generally positive attitude and also reflects the high recognition of the topic among the *online content creators*.

A more detailed classification of positive emotions shows that there are 40 positive *contents* of the general level (0, 15], accounting for 30.77%; 23 *contents* of the medium level (15, 25], accounting for 17.69%; and 35 *contents* of the high level (25, $+\infty$), accounting for 26.92%. This breakdown shows that a significant number of online content creators express positive attitudes toward the topic, with both general and highly positive content accounting for a notable share of the total.

Table 3 Emotional classification and intensity distribution.

Emotion type	Proportion	Emotional segmentation type	Proportion
Positive emotion	75.38%	Ordinary degree (0, 15]	30.77%
		Moderate degree (15, 25]	17.69%
		Advanced degree (25, +∞)	26.92%
Neutral emotion	24.62%	-	-
Negative emotion	0.00%	Ordinary degree (0, −15]	-
		Moderate degree (−15, −25]	
		Advanced degree (−25, −∞)	
Total	100%	-	-

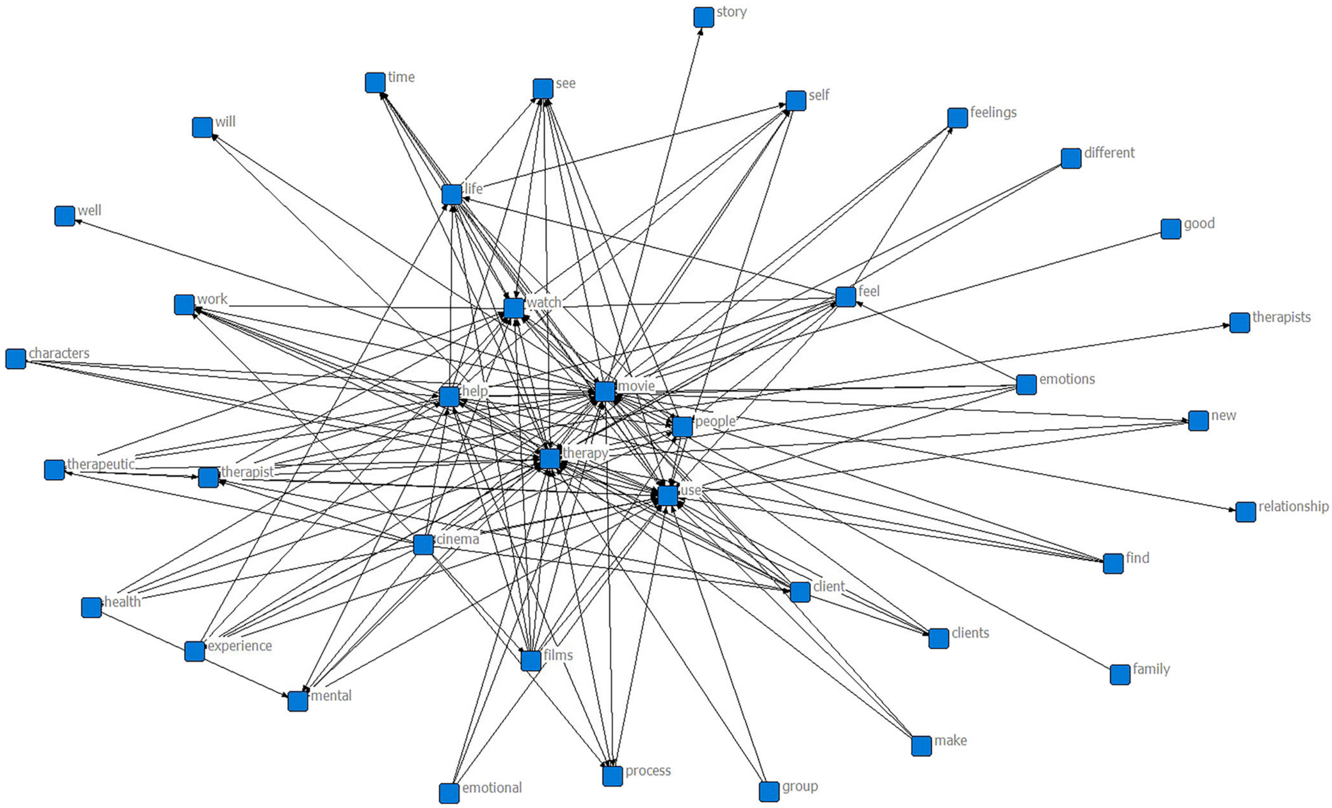

Fig. 5 Semantic network visualization.

It is particularly worth mentioning that in terms of negative emotions, there are no *contents* in each level segment (general, moderate, and high). This phenomenon not only further supports the overall positive tendency of the *contents* but also implies that the topic of discussion may have been recognized by the *online content creators* in all aspects without triggering any significant negative reactions.

The double validation mechanism systematically tested the sentiment analysis results of ROST CM6.0, and the validation results demonstrated a good level of coherence and reliability. Cross-validation machine learning results: The 5-fold cross-validation showed that the prediction accuracy of the logistic regression model based on TF-IDF features for ROST sentiment analysis results was 0.746 (±0.019). The validation scores of each fold were 0.769, 0.769, 0.731, 0.731, and 0.731, respectively, with a standard deviation of 0.019, showing good model stability. This result shows that the machine learning algorithm can reproduce the sentiment classification pattern of ROST to a certain extent, verifying the basic reliability of the sentiment analysis results. Emotion threshold analysis results: The validation accuracy of the VADER sentiment analyzer in the threshold range of 0.0–0.2 was consistent, both of which were 0.615. Comprehensive validation evaluation: The combined average score of the two independent validation methods was 0.681, showing a moderate degree of validation coherence. The validation results support the basic conclusions of the sentiment analysis in this study, but also suggest that methodological limitations need to be considered when interpreting sentiment analysis. This finding emphasizes the importance of multivariate validation methods in sentiment analysis research and provides a more prudent methodological basis for understanding the emotional characteristics of film therapy network discussions.

Semantic network analysis. We constructed a semantic network diagram based on high-frequency words (Fig. 5) to visualize the relationship between concepts related to film therapy. The results showed that the network structure showed decentralized characteristics, with “therapy,” “movie,” “use,” and “people” as core nodes, and the connections between them were the most dense, reflecting the core position of these four concepts in film therapy.

Table 4 Co-word matrix.

...	therapy	use	help	health	mental	therapist	feel	feelings	people	watch	films	work	process	life	...
therapy	-	693	519	-	-	411	300	-	360	342	324	309	279	-	...
use	693	-	429	-	-	-	-	-	303	339	297	-	-	273	...
help	519	429	-	-	-	-	267	-	273	-	-	-	-	264	...
health	-	-	-	-	435	-	-	-	-	-	-	-	-	-	...
mental	-	-	-	435	-	-	-	-	-	-	-	-	-	-	...
therapist	411	-	-	-	-	-	-	-	-	-	-	-	-	-	...
feel	300	-	267	-	-	-	-	366	-	-	-	-	-	-	...
feelings	-	-	-	-	-	-	366	-	-	-	-	-	-	-	...
people	360	303	273	-	-	-	-	-	-	-	-	-	-	-	...
watch	342	339	-	-	-	-	-	-	-	-	-	-	-	-	...
films	324	297	-	-	-	-	-	-	-	-	-	-	-	-	...
work	309	-	-	-	-	-	-	-	-	-	-	-	-	-	...
process	279	-	-	-	-	-	-	-	-	-	-	-	-	-	...
life	-	273	264	-	-	-	-	-	-	-	-	-	-	-	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	

An in-depth analysis shows that the core nodes are closely connected with different semantic groups. The connection between the "Therapy" node and words such as "experience," "process," and "help" reflects the diversity of the treatment process; the association between the "Movie" node and "watch," "characters," and "story" emphasizes the importance of the film content; and the "people" node connects "clients," "group," and "self" to highlight the diversity of the treatment objects. At the same time, the important position of psychology-related words such as "emotions," "feelings," and "mental" indicates that film therapy mainly focuses on the improvement of emotional and mental health.

The semantic network also revealed the wide application value of film therapy. The presence of "Family" and "relationship" reflects that film therapy may involve the improvement of interpersonal relationships, especially family relationships. The existence of the "Health" node further confirms the close relationship between film therapy and overall health. In addition, the presence of descriptive words such as "different" and "many" indicates that film therapy may involve diverse methods and a wide range of application scenarios. The presence of "Good" and "new" may reflect people's positive attitudes towards this treatment method.

The semantic network structure reveals four core features of the conceptual organization of film therapy based on co-occurrence matrix analysis, as shown in Table 4. First, the four main hub nodes demonstrate strong interconnectivity: "therapy" shows the highest co-occurrence with "use" (693), "help" (519), and "people" (360), while "films" maintains significant connections with "use" (297), "films" and "watch" (258), forming a conceptual bridge that connects different semantic groups and reflecting the integrative nature of film therapy as an interdisciplinary intervention method. Second, the treatment process cluster emerges through high co-occurrence values around the "therapy" node, particularly with "therapist" (411), "process" (279), and "work" (309), emphasizing the professional and procedural characteristics of film therapy implementation. Third, the media narrative cluster demonstrates strong internal connectivity, with "films" and "watch" showing substantial co-occurrence patterns (258 with core nodes), and "characters" (81 times) forming narrative-focused connections, indicating that role identification and narrative participation are the core mechanisms of film therapy. Fourth, the emotional and psychological health cluster exhibits robust co-occurrence patterns, with "mental" and "health" showing a strong connection (435), while "feel" and "feelings" demonstrate high co-occurrence (366), and "emotions" (150 times) connecting broadly across clusters, indicating that emotional processing is the fundamental therapeutic mechanism of film therapy. Cross-cluster connectivity analysis reveals that bridging words demonstrate particularly high co-occurrence values: "help" shows strong connections across multiple clusters (519 with therapy, 429 with use, 273 with people), "life" maintains significant co-occurrence with "use" (273) and "help" (264), and "health" forms a substantial connection with "mental" (435), indicating that film therapy operates simultaneously through multiple interconnected mechanisms rather than isolated pathways.

Discussion

Principal results. This study reveals the systematic characteristics and application value of film therapy using multidimensional analysis methods. The study found that film therapy contained eight main thematic dimensions and received significantly positive *contents* in practice (75.38%). Semantic network analysis further confirmed the key role of "therapy," "movie," "use," and "people" as core elements in the treatment process, providing an important theoretical and practical basis for this field.

Multidimensional therapeutic mechanisms and established cinematherapy frameworks. The eight topics identified by LDA topic modeling in this study revealed the multidimensional and systematic characteristics of film therapy, which align closely with established cinematherapy frameworks. According to Berg-Cross et al. (1990), cinematherapy can be categorized into three primary types: evocative (facilitating emotional expression), prescriptive (providing behavioral models), and cathartic (enabling emotional release). Our findings provide empirical support for this theoretical framework while revealing additional dimensions.

Evocative cinematherapy alignment: The identification of "character emotional resonance" (Topic 4) and "emotion and relationship interaction" (Topic 2) directly supports the evocative function of cinematherapy. The dense association of keywords such as emotions, characters, relationship, and emotional suggests that films promote therapeutic effects through role identification and emotional projections. This finding strongly supports Li's (2012) view that films often evoke emotional resonance in viewers, providing a safe space for exploring complex emotions and experiences. The high frequency of emotion-related terms (emotions: 188 times, emotional: 161 times, feelings: 142 times) in our corpus further validates the central role of emotional evocation in film therapy discourse.

Prescriptive cinematherapy alignment: Our topics "innovative psychotherapy based on films" (Topic 1) and "audience-participated therapy" (Topic 3) reflect the prescriptive nature of cinematherapy, where specific films are selected to model desired behaviors or coping strategies. The high frequency of keywords such as therapy, cinema, mental, and patient (therapy: 951 times, cinema: 379 times) reflects the characteristics of film therapy as a structured intervention method. This aligns with Wang and Qin's (2016) findings that film therapy is particularly useful in overcoming psychological resistance during treatment, helping patients relax and release repressed emotions while making them more receptive to therapeutic interventions.

Cathartic cinematherapy alignment: The emergence of "narrative therapy effect" (Topic 6) provides strong evidence for the cathartic function of cinematherapy. The frequent co-occurrence of keywords such as films, therapeutic, story, and clients indicates that narrative structure plays a key role in facilitating emotional release and processing. This finding supports the theoretical foundation that storytelling through film enables therapeutic catharsis.

Extended framework dimensions: Our analysis revealed dimensions that extend beyond the traditional tripartite model. The identification of "therapeutic movie selection" (Topic 5), "therapist's guidance" (Topic 7), and "multidimensional therapeutic applications" (Topic 8) suggests that emphasizes the procedural and contextual aspects of film therapy. This finding echoes Powell et al.'s (2006) emphasis on the critical importance of accurate film selection for treatment effectiveness and Berg-Cross et al.'s (1990) assertion that professional therapist guidance is essential for ensuring therapeutic efficacy.

Contradictions and gaps in existing frameworks: While our findings largely support established cinematherapy frameworks, they also reveal some conceptual gaps. The prominence of "audience-participated therapy" (Topic 3) in online discourse suggests that contemporary film therapy practice may be more collaborative and participatory than traditional models suggest. Additionally, the emergence of "multidimensional therapeutic applications" (Topic 8) with keywords including family, group, and individual indicates that online discourse recognizes film therapy's versatility across different therapeutic contexts more explicitly than classical frameworks typically acknowledge.

Core elements and characteristics of film therapy. Statistical analysis of word frequency revealed the essential characteristics and core components of film therapy. The data show that "therapy," "movie," "use," and "people" as the most frequent words that constitute the basic framework of film therapy. According to Berg-Cross et al. (1990) in the study of "Psychotherapy in Private Practice," the concept of cinematherapy was first proposed, that is, it is a therapeutic technique that allows therapists to select commercial films for clients to watch alone or with designated others. Further analysis found that the high-frequency co-occurrence of "mental," "health," and "emotional" indicates the deep application value of film therapy in the field of mental health. Particularly in narrative therapy, character identification and plot projection in movies significantly contribute to therapeutic effects (Peng, 2021; Zhang, 2019). It is worth noting that the frequent occurrence of "group" and "family" reveals that the application scenarios of film therapy include not only individuals but also groups. However, keywords for film therapy vary significantly in different cultural contexts. As Eric Greene pointed out, interdisciplinary collaboration between psychologists and filmmakers provides an opportunity to explore the therapeutic function of film in society, address cultural issues, and promote recovery (Greene and Gupta, 2022). Word frequency analysis not only reveals the core elements of film therapy but also provides important insights for its future development.

Positive feedback on treatment effectiveness. The sentiment analysis results of this study showed that 75.38% of discussions related to film therapy showed a significant positive emotional tendency, of which 26.92% were highly positive evaluations. This finding is consistent with the existing clinical evidence on the therapeutic effects of films. For example, a study of schizophrenia patients found that depression scores were significantly lower after film therapy than in controls (Lee and Ko, 2013). Another study also showed that film therapy can help reduce anxiety and improve interpersonal relationships in people with schizophrenia (Yu and Bae, 2008). In particular, in the adolescent group, Zhao et al.'s (2020) study found that watching inspirational films was found to significantly increased adolescents' optimism level. Another study showed that film therapy can help manage emotional and behavioral issues, especially in adolescents dealing with family relationships, such as parental divorce (Gagliano et al., 2023). However, it should be noted that the complete lack of negative content from online sources may reflect a selective reporting bias inherent in publicly available web content, and future studies should verify these findings using more diverse data sources and a rigorous methodological design. In general, the high proportion of positive feedback not only reflects the clinical potential of film therapy but also provides an empirical basis for its promotion and application.

Systematic integration model of film therapy. Semantic network analysis revealed the systematic characteristics of film therapy, and its decentralized network structure indicated that this treatment method has formed a relatively complete application system. Studies have shown that the close connection between the four core nodes of "therapy," "movie," "use," and "people" constitutes the basic framework of film therapy. The systematic characteristics of this framework are mainly reflected in three aspects: standardization of the treatment process (Taylor Buck and Hendry, 2016), diversity of intervention techniques (Gentle et al., 2020), and integrity of the evaluation system (Kong, 2024). Especially in clinical practice, Powell and Newgent (2010) found through a meta-analysis that film therapy using a structured application framework significantly improved the treatment effect. However, the current degree of systematization of film therapy varies significantly across different cultural contexts, suggesting the need to establish a more culturally adaptable therapeutic framework. For example, Xu (2024) suggested that cultural differences exist across countries and regions, and for films and television works to gain recognition in foreign cultural contexts, they need culturally adaptive localization to promote cultural integration. Therefore, the systematic characteristics of film therapy provide a basis for its standardized development, but it still needs to be further improved and integrated to enhance its clinical application value.

Clinical and practical implications. The eight thematic dimensions identified in this study provide a structured framework for therapists to optimize the implementation of film therapy. Based on our research findings, we propose the following recommendations for clinical application.

Film selection should follow systematic criteria. The identification of "therapeutic film selection" (Theme 5) and "narrative therapy effectiveness" (Theme 6) suggests that therapists should consider both content appropriateness and narrative structure when selecting films. We suggest a systematic film classification system based on treatment goals: films for emotion regulation (for anxiety and depression), relationship-focused films (for interpersonal problems), and role-driven narrative films (for identity and self-esteem work). The high frequency of keywords such as "selection," "story," and "therapeutic" (Table 2) suggests that structured selection criteria significantly affect treatment effectiveness.

The treatment process requires professional guidance. The theme of "therapist guidance" (theme 7) highlights the key role of professional guidance. Our semantic network analysis showed that "analysis," "therapist," and "feelings" were closely related, suggesting that an effective treatment process should include: (1) a preparatory phase before viewing the film, setting clear treatment goals; (2) guided viewing, providing therapeutic prompts; and (3) a post-viewing processing phase, focusing on emotional reactions and therapeutic insights. The core relationship between the "treatment," "people," and "experience" nodes (Fig. 3) supports a client-centered approach that prioritizes personal meaning construction.

Multidimensional treatment approaches reflect flexibility and adaptability. The theme of "Multidimensional treatment applications" (Theme 8) provides evidence for the versatility of film therapy in different settings. The presence of keywords such as "group," "family," and "individual" suggests that therapists can adapt film therapy to a variety of treatment contexts, and our data support its effectiveness in both individual and group therapy settings.

Evidence supports policy and financial decisions. The overwhelmingly positive sentiment findings (75.38% positive, 26.92% highly positive) provide strong evidence for policy and funding decisions: The evidence base for program development is strong. The lack of negative affect and high rates of positive evaluation support the integration of film therapy into standard mental health services. Policymakers can use these findings to justify pilot film therapy programs and resource allocation, especially in settings where traditional treatment resources are limited.

Standardized protocols need to be developed based on evidence. Our semantic network analysis showed that "treatment," "film," "use," and "population" were core nodes, which provided a basis for the development of standardized protocols. We recommend the establishment of: (1) an evidence-based film resource library classified by treatment goals, (2) a standardized evaluation tool for measuring efficacy, and (3) a quality indicator system for program evaluation.

Limitations and suggestions for future research. This study had several limitations. First, the study was based only on Google search results and may not fully reflect the application of film therapy on other platforms or practice settings. Second, as a study based on text analysis, it is difficult for this study to delve into the actual clinical effects and patient experience of film therapy. In addition, the sentiment analysis results show a complete lack of negative *contents* (0%). This positive tendency may be a selective bias and needs to be verified through more diverse data sources. The ROST CM6.0 model has certain limitations in terms of algorithmic transparency.

Based on the above limitations, future research can be carried out from the following aspects: (1) expand the scope of data sources and include data from multiple platforms such as social media and academic databases in the analysis to obtain a more comprehensive understanding; (2) conduct longitudinal research to track the development trends in the field of film therapy through regular data collection; (3) strengthen cross-cultural comparative research to explore the application characteristics and effect differences of film therapy in different cultural backgrounds; and (4) combine quantitative and qualitative research methods to deeply explore the mechanism and influencing factors of film therapy. Especially in the post-COVID-19 era, future research should focus on the application potential of film therapy in remote mental health services and on how to improve treatment effects through technological innovation. (5) Use more intuitive visualization tools for algorithmic decision-making processes.

Conclusion

With the rapid development of digital health technology and the increasing prominence of mental health problems, film therapy as an innovative intervention has received increasing attention. Based on 67 pages (130 sources) of film therapy-related data collected from Google search results, this study used LDA topic modeling, sentiment analysis, and semantic network analysis methods to explore the application characteristics of film therapy in digital mental health interventions. The results showed that film therapy formed eight thematic dimensions with innovative psychotherapy and emotional relationship interaction as the core, and this conclusion remained reliable after a series of model verifications. The grouping results showed that therapist guidance and film selection strategies determined the professionalism and pertinence of film therapy implementation, respectively. In addition, the positive effect showed significant theoretical fit, which was particularly prominent in the treatment of emotional and relationship problems.

Data availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Received: 15 January 2025; Accepted: 26 August 2025;
Published online: 30 September 2025

References

Alhajj R, Rokne J (2014) Encyclopedia of social network analysis and mining. pp 1–2437. https://doi.org/10.1007/978-1-4614-6170-8/COVER

Berg-Cross L, Jennings P, Baruch R (1990) Cinematherapy: theory and application. Psychother Priv Pract 8(1):135–156. https://doi.org/10.1300/J294V08N01_15

Birnbaum ML, Rizvi AF, Faber K, Addington J, Correll CU, Gerber C, Lahti AC, Loewy RL, Mathalon DH, Nelson LA, Voineskos AN, Walker EF, Ward E, Kane JM (2018) Digital trajectories to care in first-episode psychosis. Psychiatr Serv 69(12):1259–1263. https://doi.org/10.1176/APPI.PS.201800180/SUPPL_FILE/APPI.PS.201800180.DS001.PDF

Blei DM, Ng AY, Edu JB (2003) Latent Dirichlet Allocation Michael I. Jordan. J Mach Learn Res 3:993–1022

Chen T, Peng L, Jing B, Wu C, Yang J, Cong G(2020) The impact of the COVID-19 pandemic on user experience with online education platforms in China. Sustainability 12(18):7329. https://doi.org/10.3390/SU12187329

Chen W, Zheng Q, Liang C, Xie Y, Gu D(2020) Factors influencing college students' mental health promotion: the mediating effect of online mental health information seeking. Int J Environ Res Public Health 17(13):4783. https://doi.org/10.3390/IJERPH17134783

Chen X, Yin Y, Jiang M, Lin H (2023) Deep analysis of the homogenization phenomenon of the ancient water towns in Jiangnan: a dual perspective on landscape patterns and tourism destination images. Sustainability 15(16):12595. https://doi.org/10.3390/SU151612595/S1

Cooney R (2023) Leading the way to reduce mental health stigma. Nat Ment Health 1(7):439–440. https://doi.org/10.1038/S44220-023-00091-3

Doerfel ML (1998) What constitutes semantic network analysis? A comparison of research and methodologies. Connections 21(2):16–26. https://www.researchgate.net/publication/285242354

Du-Nan, Yu-Chuandong, Kim C-Y (2020) Study on the effect of film therapy on the frustration affordability of college students in Northeast China: based on film industry aesthetics in the age of IT. J Multimed Inf Syst 7(2):157–164. https://doi.org/10.33851/JMIS.2020.7.2.157

Farkhod A, Abdusalomov A, Makhmudov F, Cho YI(2021) LDA-based topic modeling sentiment analysis using Topic/Document/Sentence (TDS) Model. Appl Sci 11(23):11091. https://doi.org/10.3390/APP112311091

Gagliano A, Costanza C, Bazzoni M, Falcioni L, Rizzi M, Scaffidi Abbate C, Vetri L, Roccella M, Guglielmi M, Livio F, Ingrassia M, Benedetto L (2023) Effectiveness of an educational filmmaking project in promoting the psychological well-being of adolescents with emotive/behavioural problems. Healthcare 11(12):1695. https://doi.org/10.3390/HEALTHCARE11121695/S1
Gao H, Guo D, Wu J, Zhao Q, Li L (2021) Changes of the public attitudes of China to domestic COVID-19 Vaccination after the vaccines were approved: a semantic network and sentiment analysis based on Sina Weibo texts. Front Public Health 9:723015. https://doi.org/10.3389/FPUBH.2021.723015/BIBTEX
Geifman N, Armes J, Whetton AD (2023) Identifying developments over a decade in the digital health and telemedicine landscape in the UK using quantitative text mining. Front Digit Health 5:1092008. https://doi.org/10.3389/FDGTH.2023.1092008/BIBTEX
Gentle E, Linsley P, Hurley J (2020) "Their story is a hard road to hoe": how art-making tackles stigma and builds well-being in young people living regionally. J Public Ment Health 19(2):109–118. https://doi.org/10.1108/JPMH-10-2019-0087
Golz C, Richter D, Sprecher N, Gurtner C (2022) Mental health-related communication in a virtual community: text mining analysis of a digital exchange platform during the Covid-19 pandemic. BMC Psychiatry 22(1):1–8. https://doi.org/10.1186/S12888-022-04080-1/FIGURES/3
Goodwin J, Saab MM, Dillon CB, Kilty C, McCarthy A, O'Brien M, Philpott LF (2021) The use of film-based interventions in adolescent mental health education: a systematic review. J Psychiatr Res 137:158–172. https://doi.org/10.1016/J.JPSYCHIRES.2021.02.055
Greene E, Gupta N (2022) Introduction to the special issue on the cultural therapeutics of film. J Humanist Psychol 62(6):787–795. https://doi.org/10.1177/00221678211017342
Han W, Han X, Zhou S, Zhu Q (2022) The development history and research tendency of medical informatics: topic evolution analysis. JMIR Med Inform 10(1):e31918. https://doi.org/10.2196/31918
He Q, Li J, Li L (2024) Text mining for ford car review data—based on sentiment analysis, LDA topic model, semantic web analysis. Highlights Sci Eng Technol 118:153–161. https://doi.org/10.54097/FSH1NQ58
Kao SYS, Li ACM, Mak WWS, Cheung RYM (2023) The efficacy of acceptance and commitment therapy on self-stigma reduction among people with mental illness: a quasi-experimental design. Stig Health 9(2):162–172. https://doi.org/10.1037/SAH0000474
Kim H, Choi Y (2023) A practical development protocol for evidence-based digital integrative arts therapy content in public mental health services: digital transformation of mandala art therapy. Front Public Health 11:1175093. https://doi.org/10.3389/FPUBH.2023.1175093/BIBTEX
Kolenik T (2022) Methods in digital mental health: smartphone-based assessment and intervention for stress, anxiety, and depression. Internet Things 105–128. https://doi.org/10.1007/978-3-030-91181-2_7
Kolenik T, Gams M(2021) Intelligent cognitive assistants for attitude and behavior change support in mental health: state-of-the-art technical review. Electronics 10(11):1250. https://doi.org/10.3390/ELECTRONICS10111250
Kolenik T, Gams M (2021b) Persuasive technology for mental health: one step closer to (mental health care) equality? IEEE Technol Soc Mag 40(1):80–86. https://doi.org/10.1109/MTS.2021.3056288
Kolenik T, Schiepek G, Gams M (2024) Computational psychotherapy system for mental health prediction and behavior change with a conversational agent. Neuropsychiatr Dis Treat 20:2465–2498. https://doi.org/10.2147/NDT.S417695
Kong X (2024) The research on individual differentiated mental abnormality of disabled children and the application of art therapy. Int J Clin Exp Med Res 8(2):252–257. https://doi.org/10.26855/IJCEMR.2024.04.011
Koushiou M, Nicolaou K, Karekla M (2018) Inducing negative affect using film clips with general and eating disorder-related content. Eat Weight Disord 23(6):775–784. https://doi.org/10.1007/S40519-018-0485-2/METRICS
Lawrence-Sidebottom D, Huffman LG, Beam A, Parikh A, Guerra R, Roots M, Huberty J (2024) Improvements in sleep problems and their associations with mental health symptoms: A study of children and adolescents participating in a digital mental health intervention. Digit Health 10. https://doi.org/10.1177/20552076241249928
Lee SE, Ko SH (2013) The effects of cinema therapy on depression and self-esteem in people with schizophrenia. J Digit Converg 11(10):585–592. https://doi.org/10.14400/JDPM.2013.11.10.585
Li Y (2012) Emotional outlets: an analysis of the emotional healing function of contemporary films. Master's thesis, Anhui University. https://doi.org/10.7666/d.y2114161
Liu C, Hu M, Lin W, Wang X, Li W, Zeng K (2017) Research on the perceived tourism image of mountain scenic areas: an analysis of online travel notes about Huangshan from 2014 to 2016. Mt Res 35(4):566–571. https://doi.org/10.16089/j.cnki.1008-2786.000254
Lv K, Chen X, Jing J (2022) Research on identification of key technologies in blockchain financial industry based on combined word segmentation method and LDA model. Libr Inf Serv 66(19):110–121. https://doi.org/10.13266/j.issn.0252-3116.2022.19.011
Marciano L, Lin J, Sato T, Saboor S, Viswanath K (2024) Does social media use make us happy? A meta-analysis on social media and positive well-being outcomes. SSM Ment Health 6:100331. https://doi.org/10.1016/J.SSMMH.2024.100331
Octoparse (2025) Octoparse: no-code web scraping tool. https://www.octoparse.com/
Panayiotou M, Black L, Carmichael-Murphy P, Qualter P, Humphrey N (2023) Time spent on social media among the least influential factors in adolescent mental health: preliminary results from a panel network analysis. Nat Ment Health 1(5):316–326. https://doi.org/10.1038/S44220-023-00063-7
Peng L (2021) Exploration of film therapy to improve depression from broken hearts. Master's thesis, Nanning Normal University. https://doi.org/10.27037/d.cnki.ggxsc.2021.000504
Pirkis J, Blood RW, Francis C, McCallum K (2006) On-screen portrayals of mental illness: extent, nature, and impacts. J Health Commun 11(5):523–541. https://doi.org/10.1080/10810730600755889
Powell ML, Newgent RA (2010) Improving the empirical credibility of cinematherapy. Couns Outcome Res Eval 1(2):40–49. https://doi.org/10.1177/2150137810373920
Powell ML, Newgent RA, Lee SM (2006) Group cinematherapy: using metaphor to enhance adolescent self-esteem. Arts Psychother 33(3):247–253. https://doi.org/10.1016/J.AIP.2006.03.004
Reupert A (2024) Mental health promotion, prevention and early intervention in the workplace. Adv Ment Health 22(1):1–3. https://doi.org/10.1080/18387357.2024.2313811
Shi J, Fan M, Li WL (2009) Topic analysis based on LDA model. Acta Autom Sin 35(12):1586–1592. https://doi.org/10.3724/SP.J.1004.2009.01586
Sun X, Hu W, Li Y (2022) Research on the perceived tourism image of ancient villages based on online text analysis: a case study of Mingyuewan Ancient Village in Suzhou. Tour Overv 8:11–14. https://doi.org/10.3969/j.issn.1004-3292.2022.08.004
Taylor Buck E, Hendry A (2016) Developing principles of best practice for art therapists working with children and families. Int J Art Ther Inscape 21(2):56–65. https://doi.org/10.1080/17454832.2016.1170056
Wang X, Qin Q (2016) Research progress on the theory and techniques of film therapy. Chongqing Med 45(17):2413–2416. https://doi.org/10.3969/j.issn.1671-8348.2016.17.040
Wang Z, Zhou Q, Man T, He L, He Y, Qian Y(2024) Delineating landscape features perception in tourism-based traditional villages: a case study of Xijiang Thousand Households Miao Village, Guizhou. Sustainability 16(13):5287. https://doi.org/10.3390/SU16135287
Xu L (2024) Analysis of cultural adaptability in the overseas promotion of film and television works. J Cult 9:32–35. https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2024&filename=WHXU202409007&uniplatform=NZKPT
Yan W, Ohlsen S, Wood E (2024) Factors affecting the retention of mental health occupational therapists in the UK: a qualitative study with framework analysis. Occup Ther Ment Health. https://doi.org/10.1080/0164212X.2024.2423811
Yu I-C, Bae J-A (2008) Effects of cinema therapy on decreasing a sense of uneasiness and improving interpersonal relationship; focusing on schizophrenia patients. J Korea Contents Assoc 8(8):103–110. https://doi.org/10.5392/JKCA.2008.8.8.103
Zhang X (2019) The expression of trauma and the healing function of film: a case study of Room. West Radio Telev 7(13):121–122. https://doi.org/10.3969/j.issn.1006-5628.2019.13.059
Zhao Y, Zhao C, Ao F (2020) The impact of inspirational films on the optimistic psychology of adolescents. West Qual Educ 6(9):78–82. https://doi.org/10.16681/j.cnki.wcqe.202009037

Author contributions

GW designed the project. GW and YW collected data. GW and YY analyzed the data. GW, YY, and YW wrote the manuscript. GW, YY, and YW conducted the visualization. KS supervised the project. KS reviewed the manuscript. All authors have contributed to the manuscript.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Correspondence and requests for materials should be addressed to Kamal Sabran.