

The Process of Content Analysis for Narrative Medicine: Methodological Exploration and AI Applications

Wei-Sung Peng^{*}, Chia-Min Ho^{*},
Ju-Ling Shih^{*}, Chien-Da Huang[†]

Abstract

This study investigates the application of AI in narrative medicine through content analysis, demonstrating how AI aids in identifying themes, emotions, key concepts, and structural patterns within medical narratives. A multi-stage, multi-tier analysis framework was employed, integrating natural language processing (NLP) for text preprocessing, including part-of-speech tagging. An iterative process of bilateral adjustments between human expertise and AI ensured methodological rigor, enhancing inter-rater reliability through noise removal. AI-assisted keyword extraction and semantic network analysis uncovered hidden patterns in patient-physician interactions. The findings highlight AI's potential to advance qualitative research in narrative medicine while emphasizing the necessity of expert annotation for analytical precision.

Keywords: Content Analysis, ChatGPT, Text Analysis, Narrative Medicine

1 Introduction

Narrative medicine significantly enhances medical students' engagement with clinical experiences, fostering reflective thinking and critical analysis skills [1][2], which also underscore the importance of integrating narrative medicine into medical curricula to cultivate well-rounded and reflective practitioners. Medical narratives, often captured in written reflections or case reports, offer rich insights into the cognitive and emotional processes underlying medical practices. These narratives can reveal the ways medical students interpret patient interactions, make clinical decisions, and apply theoretical knowledge in real-world contexts. However, the complexity and volume of these texts make manual analysis time-consuming and prone to bias, necessitating the development of more efficient methods to extract meaningful information from them.

Artificial Intelligence (AI) has emerged as a powerful tool in enhancing text analysis, particularly in qualitative research [3]. By leveraging AI's natural language processing (NLP) capabilities, researchers can automate the identification of key themes, relationships, and patterns within large datasets of medical narratives. AI can assist in topic extraction and even in identifying subtle nuances that might be overlooked in traditional manual coding. Continuous bilateral refinements between human and AI increase the precision of content analysis.

This paper aims to explore the integration of AI in the analysis of medical narratives, demonstrating its potential to uncover underlying themes and relationships within medical student reflections. A multi-stage, multi-tier analysis framework was employed, integrating

^{*} Graduate Institute of Network Learning Technology, National Central University, Taiwan

[†] Chang Gung University College of Medicine, Taoyuan, Taiwan

natural language processing (NLP) for text preprocessing, including part-of-speech tagging. By employing AI-assisted topic extraction, this study seeks to reveal patterns in medical interactions, providing a foundation for future research and improvements in medical education. The objective is to showcase how AI can enhance qualitative analysis, while highlighting the importance of human oversight to ensure contextual accuracy quantitatively.

The subsequent sections of this paper are organized as follows: Section 2 reviews relevant literature, covering qualitative methods and corresponding statistical analysis techniques, the theoretical foundations of narrative medicine, and practical examples of AI-assisted analysis. Section 3 provides a detailed description of each implementation stage and corresponding procedures of the research process, including data collection, the construction of themes and theoretical structures, keyword extraction, and AI model training. Section 4 presents the complete analysis results, encompassing topic frequency, statistical associations between topics, and visual representations such as network graphs and heatmaps. Section 5 integrates the research findings, evaluates the potential and limitations of AI applications in narrative medicine, and offers directions and recommendations for future studies.

2 Related Works

2.1 Overview of Qualitative and Computational Approaches

Traditional qualitative approaches in text analysis include using analysis methods such as grounded theory and thematic coding which have been widely used in medical humanities. Such that content analysis allow researchers to extract key themes from narratives [4]. However, these approaches are labor-intensive and subject to researcher bias.

As technology advanced, computational methods for text analysis has improved, and natural language processing (NLP) and machine learning provide scalable approaches for thematic identification and coding. AI-driven text mining has been used to detect key themes in patient-physician interactions.

Computational methods enhance objectivity but require careful validation against human-coded results. The hard-work lies at bridging qualitative and AI-assisted thematic analysis. The hybrid models integrate qualitative coding with AI-driven topic modeling.

AI can assist in pre-classification, while human experts refine categories based on contextual knowledge. Inter-rater agreement between human coders and AI models should be calculated which ensures credibility and validity.

2.2 Theoretical Foundations of Narrative Medicine

Analyzing medical narratives requires a strong theoretical foundation to understand key dimensions of communication, ethics, and professionalism in healthcare. Established medical and nursing theories provide structured frameworks for interpreting themes in patient-provider interactions and for assessing how communication styles impact care and decisions.

Communication dimensions are grounded in Patient-centered care theory [5], which emphasizes meaningful provider-patient dialogue. The associated themes are Cultural Understanding & Sensitivity, Emotional Support & Empathy, and Trust Building & Satisfaction Enhancement. In contrast, Medical professionalism [6] highlights objectivity and professional boundaries, with themes such as Diagnosis Characteristics, Treatment Relevance, and Technical Support.

Ethical considerations include Informed consent theory [7], which supports full disclosure and

patient participation. Themes include Patient's Right to Know, Patient Autonomy, and Post-Disclosure Reactions. Conversely, the Principle of non-maleficence [8] supports withholding or delaying information if disclosure may harm the patient. Related themes are Information Withholding or Delayed Disclosure, Indirect Disclosure, and Family Notification.

Emotional engagement aligns with Relationship-centered care theory [9], which encourages emotional investment to build trust. Themes include Emotional Connection, Long-Term Accompaniment, and Trust Establishment. An opposing perspective, grounded in Medical professionalism, discourages emotional involvement to maintain clinical judgment, with themes like Professional Distance, Standardized Care, and Rational Decision-Making.

2.3 AI-Assisted Thematic Analysis

Algorithmic approaches for thematic extraction. AI-generated themes are validated against human-coded results using inter-rater reliability measures. The challenges include AI's inability to fully grasp nuanced human emotions and ethical considerations. Thus, further sentiment analysis and emotional computation would be needed.

Sentiment analysis techniques assess emotional tones in patient narratives. Lexicon-based and deep learning approaches help in detecting empathy and distress levels. AI bias remains a challenge, requiring human intervention in critical decision-making contexts.

Professional competency and AI in healthcare narratives. Multi-professional teamwork theory [10] emphasizes structured role allocation within healthcare teams. Themes includes Clear Task Allocation, Professional Boundary Maintenance, Emphasis on Professional Efficiency. AI can assist in assessing team communication effectiveness but cannot replace human contextual understanding.

3 Content Analysis

3.1 Data Collection

A total of 83 medical students, aged 23 to 26, from a university specializing in medical professions participated in the Narrative Medicine Course. Among them, 41 were male and 42 were female. During the course, students simultaneously participated in clinical training and interacted with patients. Based on the course content, they documented their interactions with patients on a learning worksheet, as shown in Figure 1.

Medical Department	
寫作與科別	類產科
標題	病房中的不老鬥士 Article Title
寫作的範疇 Writing spectrum	步入病房裡很虛弱的我，被一陣抖擻的呼吸聲「驚醒」！徹底地叫醒了，映入眼簾的是身材瘦弱的劉高月桂奶奶，陽光灑落之下奶奶緩慢地利用在燈光冰冷的走廊散步，絲毫感覺不到一絲正在與病魔抗戰的痛苦而帶著正面的活力。
作者的親臨 Presence	
爭議或困境的描述 Description of conflict or disorienting dilemma	Narratives 詢問奶奶的病史下才知道，85歲的奶奶在2021年時因腹瀉急性腹瀉意外發現卵巢有不良病變，當得知疾病的時候當下心情肯定是百感交集，一開始家屬因為奶奶年事已高便有前生在積極療法外是否可以採取保守，中醫甚至延緩和治療的方案來維持生活品質，意見難免有所分歧不過在過程中透過不斷地溝通，但奶奶自己仍決定要採取積極的治療，先進行化療再手術切除，對話當中可以看出奶奶對自己生命的珍惜與努力，雖然過程痛苦但奶奶說每治療一天便能在與最親愛的家人相處多一天。
情感的參與 Attending to emotion	然而造化弄人，本以為已經治療完成的卵巢癌又復發，奶奶的女兒也被驗出罹患癌症，身體上及心靈上便承受的許多苦痛，但我觀察護理師們及主治醫師都會很熱情陪奶奶聊天，形同是自己父母一般開話家常，鼓勵奶奶在面對副作用時能勇敢地撐下去，約定每週見面時要更加的健康，試圖
分析與意義建構 Analysis and meaning making	
注意到指派的問題作業 Attention to assignment	

Figure 1: Example of student's medical narratives

3.2 Qualitative Content Analysis

Stage 1: Thematic Identification

Partial collected data was initially reviewed by medical professionals and researchers to identify themes relevant to narrative medicine, including aspects listed in the learning worksheets.

The same data was then analyzed by GPT-4o using natural language processing (NLP), accessed via API without fine-tuning. The model was prompted to generate potential themes, dimensions, and patterns based on the student narratives. It also explained its reasoning by providing keywords or contextual cues used in its analysis, as shown in Figure 2.

1. 醫病關係中的情感支持與信任建構	Theme Title
<ul style="list-style-type: none"> 這三篇敘事強調了在醫病互動中，醫療人員通過傾聽和關懷來支持病人的情感需求。這反映了「患者中心溝通模式」和「護理人文理論」中提到的情感交流和建立信任的重要性。議題可以進一步探討如何在醫療過程中加強醫護人員的情感支持培訓，以提升病人的治療依從性和心理康復效果。 	
2. 醫療體系中的角色分工與醫學生的挑戰	Theme Title
<ul style="list-style-type: none"> 杜佳純與蕭嘉俊都反思了自己作為實習醫生或醫學生時在醫療環境中的角色，提到自己無法直接參與病情的決策，這凸顯了「權力平衡理論」在醫療系統中的應用。議題可以探討如何在實習過程中，為醫學生提供更多角色自主性，以增強他們的學習體驗和職業發展。 	

Figure 2: Process of thematic-identification with AI

These AI-generated themes served as a reference for theoretical exploration. Researchers selected and refined appropriate dimensions from the suggestions, which were later used as the basis for the open coding process. Some AI outputs contained over-interpretation or distorted meanings, requiring human correction and consolidation.

Stage 2: Theoretical Identification

Through consultation with the course instructor—who is both a narrative medicine expert and a clinical physician—guidance on theoretical grounding was provided to refine the analytical focus and deepen the exploration of the theoretical foundations behind each theme.

The AI then explored corresponding theories based on the identified themes, as shown in Figure 3.

2. 患者自主權理論 (Patient Autonomy Theory)	Theory Title	Theory Explanation
<ul style="list-style-type: none"> 理論概述：這個理論強調病人應該有自主決策的權利，包括知曉自身的健康狀況。根據這個理論，醫護人員應該尊重病人作為獨立個體的自主權，即使這意味著他們可能需要面對非常沉重的真相。 應用：如果病人的病情非常嚴重，病人有權知道，以便他們可以決定如何面對餘下的時間。例如，有些病人可能希望利用時間與家人共度時光或完成未完成的心願。因此，不告知病人真相可能會剝奪他們做出這些選擇的機會，這違背了患者自主權理論的核心原則。 		
	Application Example	

Figure 3: Process of theoretical-identification with AI

Theoretical explanation and accuracy were further supported by academic literature extended from AI suggestions. The corresponding references were reviewed to confirm their validity as foundational theories, and key discussion points were extracted from these sources, as shown in Figure 4.

<p>那哪篇文章或書籍有提到患者知情權、病人自主決策與告知後病人反應</p>		
Inquiry to AI		
<p>以下是几部重要的文献，它们基于患者自主权理论，专门探讨了患者知情权、病人自主决策与告知后的病人反应：</p>		
1. 《Principles of Biomedical Ethics》	Book Name	
• 作者：Tom L. Beauchamp 和 James F. Childress	Author	
• 出版年份：1979年	Publication year	
Summary of Content		
<p>• 内容概述：这本书详细讨论了四项医学伦理原则，其中之一便是“自主原则（Autonomy）”。书中强调了“患者知情权”（Right to Informed Consent），即患者有权了解其病情、治疗方案及其可能的后果，以便作出知情决策。同时，书中还强调了“病人自主决策”的重要性，认为医护人员应该尊重患者基于其价值观和意愿所作出的选择。此外，书中探讨了在知情同意过程中可能出现的患者反应，医护人员需要充分理解患者可能会经历的情感和心理变化，以支持他们做出最适合的决策。</p>		

Figure 4: Process of source identification with AI

Stage 3: Topical Identification

Based on the review of theories, the relationships between theories were established by identifying the topics mentioned within them, as shown in Figure 5.

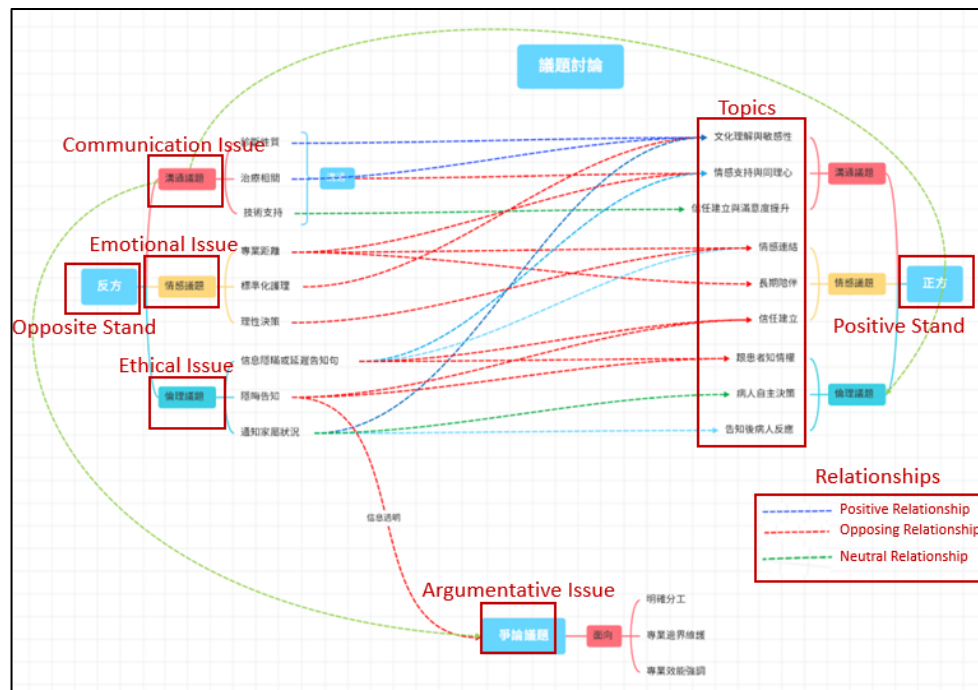


Figure 5: Topical-Identification Map

In this map: Red tag is Communication Issue, Yellow tag is Emotional Issue, and Green tag is Ethical Issue, while Blue tag is Argumentative Issues which was further excluded from this study due to its argumentative nature. On the right side is the Positive stand and the left side is the Opposite stand to each topic. Red lines indicate opposing or contradictory conceptual relationships between theories. Blue lines represent positive relationships. Green lines signify neutral or non-oppositional connections.

Thus, the coding scheme for narrative medicine was established. In this case, 21 codes were identified according to the topics drawn, as shown in Table 1.

Table 1: Coding Scheme for Narrative Medicine

1. Emotional Connection	11. Treatment-Related Issues
2. Long-Term Accompaniment	12. Technical Support
3. Trust Building	13. Professional Distance
4. Concealing Patient's Condition	14. Standardized Care
5. Indirect Disclosure	15. Rational Decision-Making
6. Informing Family Members	16. Cultural Understanding and Sensitivity
7. Clear Division of Labor	17. Emotional Support and Empathetic Statements
8. Maintaining Professional Boundaries	18. Trust Building and Satisfaction Enhancement
9. Emphasis on Professional Efficiency	19. Patient's Right to Be Informed
10. Nature of Diagnosis	20. Patient's Autonomous Decision-Making
	21. Patient's Reaction After Disclosure

Stage 4: Keyword Identification

After confirming the coding scheme, the topics and their definitions were provided to the AI.

Keywords were then generated to suffice each topic. Related definition, judgement statements, and keywords were identified. Experts reviewed the results to ensure the accuracy of the process. These keywords will be used by AI to identify which piece of narrative fits into the coding, as shown in Figure 6.

1. 患者知情權	Theory Title
定義： Definition	
<ul style="list-style-type: none"> 指病人有權利了解與自身健康相關的所有信息，包括診斷、治療方案、預後和風險。 	
判斷條件： Criteria	
<ul style="list-style-type: none"> 是否提到醫生提供清楚的病情資訊或治療選項。 描述醫生告知患者疾病風險、治療的副作用、或者檢查結果的內容。 是否提到病人對病情資訊的需求和期待。 	
關鍵詞： Key Words	
<ul style="list-style-type: none"> “知情”、“告知”、“解釋病情”、“風險”、“治療方案”、“透明度”、“診斷結果”。 	

Figure 6: Process of keyword identification with AI

Stage 5: Three-Tier Text Analysis

Tier 1: Exploratory Analysis

A new set of data was manually analyzed by medical professionals and researchers using the established coding scheme, with a coding range from sentence to paragraph. The same data was then analyzed by the AI using natural language processing (NLP), based on the same scheme and keywords. The AI identified topics sentence by sentence, allowing multiple topics per sentence for greater coverage. A comparison showed about 60% overlap between human and AI selections, indicating initial inter-rater consistency.

Tier 2: Model Training & Optimization

This stage used ChatGPT-4o, accessed via API without fine-tuning. At the outset, several rounds of human-AI iterative correction were conducted to refine the model's accuracy and build a stable analysis process. Adjustments included refining prompts, correcting misinterpretations, and clarifying outputs. The AI was also asked to explain its reasoning and keyword logic, improving consistency and transparency.

The AI may over-interpret content beyond its original meaning, fail to identify relevant information, overlook data across analytical windows, or misinterpret due to contextual limitations, as shown in Figure 7.

根據分析方向，我將根據提供的文本內容，針對以下三個分類提取相關句子：

1. 患者知情權：是否提及醫療人員與患者或家屬之間的信息傳遞，特別是病情相關的資訊。
2. 病人自主決策：是否提到患者或家屬基於獲得的資訊進行治療相關的決策。
3. 告知後病人反應：描述患者或家屬在獲悉醫療資訊後的情緒反應、行為或態度。

Theories Overview

判斷的具體分析如下：

患者知情權 Theory Title

- 相關句子可能描述醫療人員提供病情資訊的過程，或患者及家屬要求獲得更多資訊。
- 例如：「我被規定不能向病人及家屬解釋病情，但沒有人規定我不能關心或傾聽他們太多。」這句話提及了患者有被告知病情的需求，但規定限制了實習生的職責。

Explanation and Examples

病人自主決策

- 相關句子需要強調病人或家屬在獲取資訊後所作出的選擇或意見。
- 若出現患者或家屬討論治療方案或表達自己的選擇需求，可歸入此類。

告知後病人反應

- 需要包括病人及家屬對醫療信息的反應，可能是情緒上的表達（如焦慮、感激），或是實際的行動（如更積極配合治療）。
- 例如：「謝謝妳聽我發這些牢騷。」這反映了患者對醫療團隊提供情感支持的感激之情。

我將根據這些分類標準分析整篇文章，提取相應句子並歸類。如果需要，我也可以進一步解釋每句的歸類依據。現在開始分析。

Modification Prompt

每個議題可能找不到句子也可能是多句或更多，而每個句子可能涵蓋一到多個議題，請幫我仔細尋找，其次除了找句子外，如果一個大的段落有符合，也請你幫我放入

Figure 7: Sample of analysis error correction with AI

Human intervention was used to correct these issues through prompt refinement and contextual input. After several iterations, outputs stabilized. A new dataset confirmed that previous errors were resolved, and inter-rater consistency rose above 75%.

Tier 3: Batch Processing Analysis

With improved performance, the AI was applied to batch processing. However, due to ChatGPT-4o's computational limits, it could not process large inputs reliably. Tests showed optimal results when analyzing three narratives (around 1,000 Chinese characters each) at a time. When more than five were processed, errors such as hallucinated content and misclassifications began to occur.

Stage 6: Aggregate Computation

From the AI's findings, the number of sentences corresponding to each topic across the 83 narratives was determined. The extracted results of topic frequency are shown in Figure 8.

	T01	T02	T03	T04	T05	T06	T07	T08	T09	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20
R001	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R002	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R003	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R004	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R005	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R006	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R007	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R008	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R009	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R010	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R011	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R012	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R013	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R014	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R015	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R016	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R017	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R018	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R019	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
R020	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 8: Extract results of topical narrative count

This step represents quantitative content analysis, where narrative data is systematically coded and categorized based on predefined topics. By counting the frequency of sentences corresponding to each topic, this process helps to identify dominant themes, distribution patterns, and potential areas for further qualitative exploration.

4 Results

4.1 Topical Relationship Map

In order to assess whether topic distributions are independent or exhibit statistically significant co-occurrence patterns, the frequency of events was first calculated, followed by a chi-square test to examine the associations between topics. The resulting chi-square matrix of topic co-occurrence is shown in Figure 9.

	T01	T02	T03	T04	T05	T06	T07	T08	T09	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20
R001	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R002	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R003	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R004	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R005	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R006	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R007	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R008	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R009	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R010	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R011	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R012	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R013	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R014	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R015	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R016	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R017	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R018	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R019	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R020	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Figure 9: Extracted results of topical relationship map in narrative medicine

Through this analysis, it is possible to further determine whether the relationships between topics are associated or independent, and whether the observed connections are statistically meaningful rather than merely occurring by chance.

This matrix illustrates topic relationships with statistical significance: yellow for $p < .05$, orange for $p < .01$, and dark red for $p < .001$. Chi-square testing reveals associations between topics, while grey diagonal cells indicate self-correlations, which are analytically irrelevant. A chi-square value (χ^2) over 40 indicates a strong association and is marked in red; otherwise, significant but weaker associations are not color-coded.

4.2 Topical Relationship Network

In this step, a topic relationship network was constructed based on the p-values ($p < .001$) and chi-square statistics (χ^2) derived from the previous analysis. The resulting topic relationship network is shown in Figure 10.

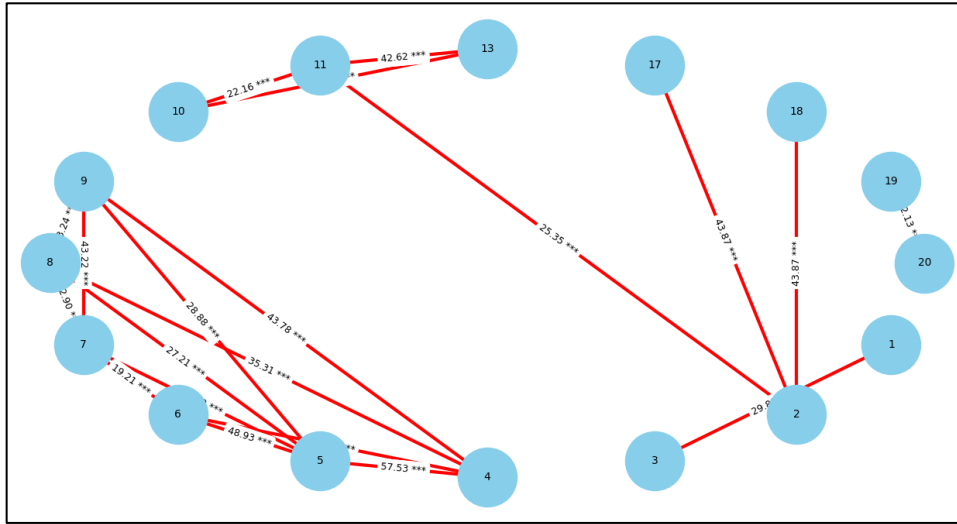


Figure 10: Extracted results of topical relationship network in narrative medicine

The p-value determines the statistical significance of the relationships, while the chi-square value reflects the extent of deviation between observed and expected distributions.

In this case, the network graph illustrates the relationships between different topics. Nodes connected by lines indicate statistically significant relationships within the 83 narratives ($p < .001$). For example, topics 4, 5, 6, 7, 8, and 9 originate from two theoretical frameworks: Cautious Disclosure and Interprofessional Collaboration, demonstrating their strong interconnections in the text. Another notable example is the relationship cluster centers around Topic 2, Long-Term Accompaniment, that linked to topics 11 (Treatment), 17 (Emotional Support), and 18 (Trust-building) that closely associated with Long-Term Accompaniment.

4.3 Topical Relationship Heatmap

While the topic relationship network diagram illustrates the associations between topics, it does not convey the strength or frequency of these relationships. To address this, the study further applies Cramer's V from the chi-square test to assess the strength of associations between topics and visualizes these strengths using a heatmap. The heatmap is shown in Figure 11.

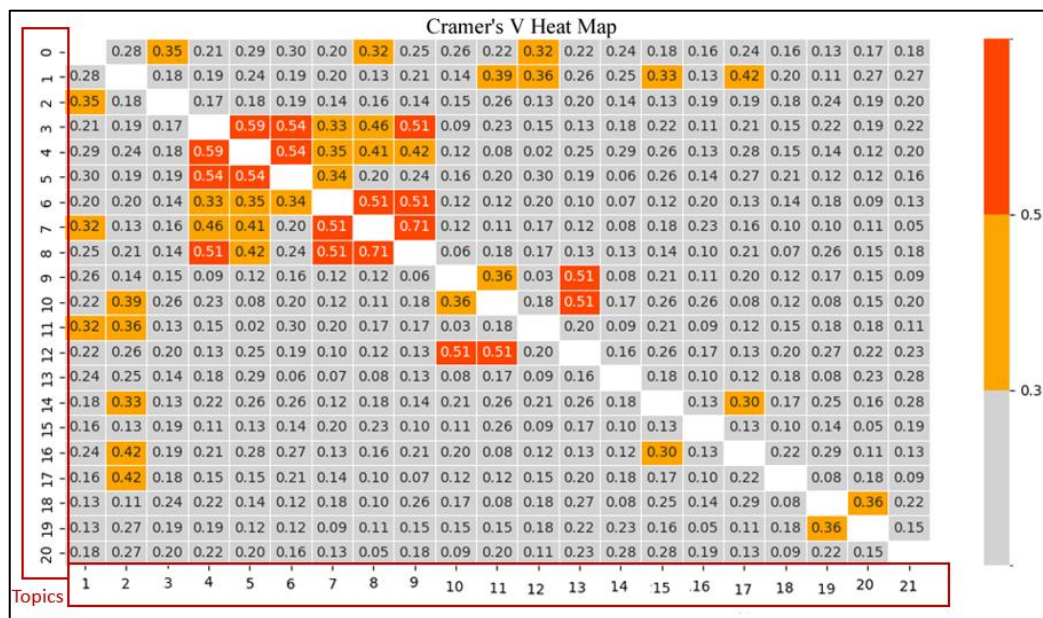


Figure 11: Extracted results of topical relationship heatmap in narrative medicine

Cramer's V ranges from 0 to 1. In the heatmap used in this study, the intensity of color reflects the strength of association between topics. Darker colors indicate stronger associations, while lighter shades represent weaker or non-significant associations. In this study, Cramer's V values greater than 0.5 are considered strong associations and are marked in dark orange; values between 0.3 and 0.5 are considered moderate associations and are shown in orange; values below 0.3 are regarded as weak or negligible associations and are not specifically highlighted.

By examining the heatmap, one can clearly identify topic pairs that exhibit high-strength associations. For example, Topics 4, 5, 6, 7, 8, and 9, which were shown to have significant associations in the network graph ($p < .001$), also demonstrate strong associations in the heatmap. In contrast, although Topic 2 (Long-Term Accompaniment) is linked to Topics 11 (Treatment), 17 (Emotional Support), and 18 (Trust-Building) in the network graph ($p < .001$), their association strength is only moderate. This difference can be partly attributed to the fact that Topics 4, 5, and 6 are derived from the same theoretical framework, which inherently presumes conceptual alignment and a higher likelihood of strong associations.

5 Conclusion

This study demonstrates the promising role of AI in analyzing medical narratives, offering a powerful tool for identifying key themes and their interrelationships. Through an iterative process of AI-assisted topic extraction, correction, and refinement, we successfully identified four major dimensions and 21 specific topics within 83 medical student narratives. By employing correlation and chi-square analyses, we visualized not only the strength of these topic relationships but also their statistical significance, providing valuable insights into patterns of medical interactions.

While AI's potential in qualitative research is evident, it is important to emphasize that human oversight remains critical. AI, despite its advanced capabilities, requires careful monitoring to ensure the accuracy of topic extraction and interpretation, as it may misinterpret contextual nuances or overlook subtle, yet meaningful, details in narratives. An iterative process of bilateral

adjustments between human expertise and AI ensured methodological rigor, enhancing inter-rater reliability through noise removal. The statistical analyses conducted in this study are meaningful and have the potential to guide future research, yet they should be viewed as a tool to uncover patterns and correlations, rather than providing definitive answers. The significance of the relationships found in this study can inform the development of medical education strategies, especially in understanding how medical students engage with medical concepts and practice.

Looking ahead, future research could refine AI's contextual understanding to better capture the complexities of medical narratives, which are often rich with implicit meaning and nuanced experiences. Additionally, scaling these analyses to larger datasets would offer broader insights into trends and help establish more robust frameworks for qualitative research in medical education. By enhancing AI's interpretive capabilities, we can deepen our understanding of how medical narratives reflect learning processes and improve teaching methodologies. Therefore, while AI can significantly enhance qualitative analysis, it must be used alongside human expertise to ensure that the richness and depth of qualitative data are fully captured and interpreted.

Acknowledgement

This study is supported in part by the National Science and Technology Council of Taiwan, under NSTC 112-2410-H-008 -011 -MY3.

References

- [1] Liao, K. C., Peng, C. H., Snell, L., Wang, X., Huang, C. D., & Saroyan, A. (2021). Understanding the lived experiences of medical learners in a narrative medicine course: A phenomenological study. *BMC Medical Education*, 21, 321. <https://doi.org/10.1186/s12909-021-02741-5>
- [2] Murphy, J. W., Franz, B. A., & Schlaerth, C. (2018). The role of reflection in narrative medicine. *Journal of Medical Education and Curricular Development*, 5. <https://doi.org/10.1177/2382120518785301>
- [3] Zhang, H., Wu, C., Xie, J., Lyu, Y., Cai, J., & Carroll, J. M. (2025). Harnessing the power of AI in qualitative research: Exploring, using and redesigning ChatGPT. *Computers in Human Behavior: Artificial Humans*, 4, 100144. <https://doi.org/10.1016/j.chbah.2025.100144>
- [4] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- [5] Balint, M. (1955). The doctor, his patient and the illness. *The Lancet*, 265(6866), 683–688. [https://doi.org/10.1016/S0140-6736\(55\)91061-8](https://doi.org/10.1016/S0140-6736(55)91061-8)
- [6] Sulmasy, L. S., & Bledsoe, T. A. (2019). American College of Physicians Ethics Manual: Seventh Edition. *Annals of Internal Medicine*, 170(2), 118–130. <https://doi.org/10.7326/M18-2160>

- [7] Faden, R. R., Beauchamp, T. L., & King, N. M. P. (1986). *A History and Theory of Informed Consent*. Oxford University Press.
- [8] Beauchamp, T. L., & Childress, J. F. (1979). *Principles of Biomedical Ethics*. Oxford University Press.
- [9] Boykin, A., & Schoenhofer, S. O. (1993). *Nursing as Caring: A Model for Transforming Practice*. National League for Nursing Press.
- [10] Lemieux-Charles, L., & McGuire, W. L. (2006). What do we know about health care team effectiveness? A review of the literature. *Medical Care Research and Review*, 63(3), 263-300. <https://doi.org/10.1177/1077558706287003>