

Brief Report

Narrative Medicine and AI in Anesthesiology Training: Teaching Empathy in End-of-Life Care

Anna La Palma ^{*,†} , Giuliana Scarpati [†] , Giulia Savarese [†]  and Ornella Piazza [†] 

Department of Medicine and Surgery, School of Medicine “Scuola Medica Salernitana”, University of Salerno, 84081 Baronissi, SA, Italy; gscarpati@unisa.it (G.S.); gsavarese@unisa.it (G.S.); opiazza@unisa.it (O.P.)

* Correspondence: alapalma@unisa.it; Tel.: +39-345-585-7156

† These authors contributed equally to this work.

Abstract

Teaching empathy remains a challenge in medical education, particularly in anesthesiology, where physicians frequently care for patients at the end of life. Narrative Medicine, centered on communicative competence and patients’ lived experience, offers a framework for cultivating reflective and relational skills. Meanwhile, artificial intelligence (AI) systems can generate expressions of empathy, raising questions about authentic moral engagement. To explore how narrative-based education, combined with AI-generated texts, may stimulate reflection, we implemented an exploratory narrative-based intervention involving 25 anesthesiology residents, supported by three tutors, integrating literature, film, and AI-generated narratives. After an introduction session, participants engaged with excerpts from the book *What Are You Going Through* and the film *The Room Next Door*, followed by reflective writing based on five prompts. The same prompts were submitted to ChatGPT (OpenAI, GPT-4o) for comparative analysis, discussed during a debriefing session. Reflective writings were assessed using an adapted REFLECT rubric, alongside qualitative lexical and semantic analyses. Most participants did not reach the highest levels of reflective capacity, while ChatGPT texts achieved higher REFLECT scores, primarily due to linguistic coherence. These findings suggest that empathic competence is neither automatically acquired through medical training nor reducible to verbal fluency. Rather, it requires structured training grounded in meaningful engagement with patients.

Keywords: narrative medicine; end-of-life; artificial intelligence



Academic Editors: Brian M. Ross and Hideki Kasuya

Received: 18 January 2026

Revised: 20 February 2026

Accepted: 4 March 2026

Published: 10 March 2026

Copyright: © 2026 by the authors.

Published by MDPI on behalf of the Academic Society for International Medical Education. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The role of patients in healthcare and their relationships with professionals has evolved significantly, with greater emphasis on shared decision-making and patient-centered care. In this context, illness narratives illuminate experiences of vulnerability and dependence on others not only for care but also for communication [1], bringing into focus the relational and ethical dimensions of clinical practice. Despite growing recognition of the importance of empathy and communication in clinical practice, medical education continues to prioritize technical skills over relational and reflective capacities [2]. This tension is particularly evident in Anesthesiology, Intensive Care, and Pain Medicine, where physicians frequently care for critically ill and dying patients. Listening, effective communication, and empathic engagement are particularly crucial in the care of critically ill and dying patients, yet the affective and moral dimensions of end-of-life care are often left to informal learning rather than addressed through structured educational programs. Physicians are expected to

accompany patients and families with compassion and emotional composure, but they frequently receive limited guidance on how to process, interpret, and integrate the emotional impact of death into their professional identity. As a result, young physicians working in high-intensity settings such as anesthesiology and intensive care commonly experience emotional burden and moral distress, often in the absence of protected educational spaces for reflection and meaning-making [3].

In this context, Narrative Medicine, as conceptualized by Charon, offers a pedagogical framework for cultivating reflective and relational skills through the integration of literary texts, films, and reflective writing [4]. It is grounded in the premise that clinical practice involves not only biomedical reasoning but also the capacity to recognize, interpret, and be moved by patients' stories. Described as a structured methodology based on specific communicative competencies, Narrative Medicine uses storytelling to acquire, understand, and integrate the multiple perspectives involved in illness and care. By facilitating narrative exchange, it creates a space of encounter between healthcare professionals and patients' lived experience, fostering reflection, emotional awareness, and professional self-consciousness [5]. This approach encourages clinicians to engage with patients' experiences and the language of suffering not merely as technical professionals, but as morally and emotionally involved participants in care, even within highly technical disciplines.

More recently, the emergence of generative artificial intelligence (AI) tools, such as ChatGPT, has stimulated interest in their potential applications in medical education [6]. By producing linguistically coherent texts that simulate an empathic tone [7], these systems may serve as stimuli for reflection and discussion. However, their use also raises critical questions, particularly the risk that linguistic fluency may be conflated with genuine empathic or moral engagement. Empathy cannot be reduced to a rhetorical performance; rather, it involves human engagement rooted in lived experience and moral responsibility.

Within this framework, the aim of the study was to foster authentic perspective-taking among anesthesiology residents in end-of-life contexts by inviting them to think from the patient's point of view. Consistent with the theoretical foundations of Narrative Medicine, which conceptualizes empathy as the capacity to recognize and engage with the experiential dimension of illness, AI-generated texts were used to make visible the distinction between linguistic coherence and empathy anchored in human experience.

2. Materials and Methods

2.1. Study Design

This study was designed as an educational intervention combining quantitative assessment of reflective capacity with qualitative lexical and semantic analyses, conducted within a residency training program in Anesthesiology, Intensive Care, and Pain Medicine.

Twenty-five anesthesiology residents participated in this educational intervention, supported by three clinical tutors. Participants included residents from different years of training (first to third year); participation was voluntary and anonymous. No incentives were provided, and participation or non-participation had no impact on formal evaluation. Tutors fulfilled a primarily facilitative role, aimed at creating a psychologically safe learning environment, guiding group discussion, and supporting reflective dialogue, rather than demonstrating superior reflective performance. This approach was intended to minimize hierarchical influence and to encourage authentic expression. The educational activity began with a short introductory lecture addressing the principles of Narrative Medicine, the doctor–patient relationship, and communication at the end of life. Participants were then invited to engage with two narrative stimuli: selected excerpts from Sigrid Nunez's novel *What Are You Going Through* [8] and the film *The Room Next Door* [9]. Both works were

chosen for their exploration of dying, accompaniment, autonomy, and meaning from the patient's perspective.

Following exposure to the narrative materials, participants completed an online reflective questionnaire consisting of five open-ended prompts, each corresponding to a specific narrative domain: emotional resonance, place of death, companionship in death, timing of death, and manner of dying. The five prompts were intentionally designed to reflect central dimensions of end-of-life experience identified in the literature on quality of dying and end-of-life care. These domains have been identified as significant contributors to perceptions of a "good death" among patients and caregivers [10]. The prompts were designed to encourage personal, non-judgmental reflection and to facilitate connections between emotional responses, ethical considerations, and professional identity.

To extend the reflective process, the same prompts were submitted to ChatGPT (OpenAI, San Francisco, CA, USA, version GPT-4o) in zero-shot mode, without contextual framing or guidance, in order to generate AI-produced narratives.

A final debriefing session was conducted involving residents and faculty tutors, with the participation of instructors from the field of psychology, whose role was to facilitate reflective discussion and to ensure emotional safety when addressing potentially distressing themes related to dying and professional vulnerability.

During this session, aggregated results from the lexical and semantic analyses, including word clouds and semantic maps, were discussed collectively. Selected student responses achieving higher REFLECT scores were presented alongside AI-generated texts produced by ChatGPT. This comparative discussion was designed to stimulate reflective dialogue on the physician's role in end-of-life care, the meaning of empathy, and the educational implications of using artificial intelligence in medical training.

This narrative-based approach was intentionally chosen because engagement with literary and cinematic representations of illness offers an experiential pathway to perspective-taking that abstract ethical instruction alone may not provide.

2.2. Analysis

Participants' reflective writings were analyzed using an adapted version of the REFLECT rubric, which assesses reflective capacity across four levels: habitual reflection (score 1), thoughtful reflection (score 2), analytical reflection (score 3), and critical reflection (score 4). This rubric evaluates depth of reflection, self-awareness, and integration of experience into professional identity [11,12].

Reflective writings were scored by members of the research team familiar with the REFLECT rubric. Scores were assigned following collective discussion, and consensus was reached for each text to promote consistency in rubric application.

Given the exploratory nature of the intervention, and the small sample size, quantitative analysis was limited to descriptive statistics. Qualitative lexical and semantic analyses were conducted to explore thematic and conceptual patterns within the narratives. Mean REFLECT scores and standard deviations were calculated for residents and tutors.

Application of the REFLECT rubric to AI-generated texts was exploratory and not intended to establish performance equivalence or superiority; rather, its inclusion was pedagogically intentional and aimed at stimulating critical comparison within the educational setting, acknowledging that the rubric was originally developed for human reflective writing.

In addition to rubric-based assessment, qualitative lexical and semantic analyses were conducted using T-LAB software (version 10, Franco Lancia, Rosasecca, Italy) on residents' reflective writings only. Word frequency analysis identified the most recurrent lexical units, and co-occurrence analysis examined associations between words within the corpus. The

resulting semantic maps and word clouds were used as visual tools during the debriefing session to support reflective discussion.

The integration of quantitative REFLECT scoring and qualitative lexical–semantic analysis was intended to provide complementary perspectives on residents’ reflective processes; while the REFLECT rubric assessed the structural depth of reflective writing, lexical and semantic analysis explored thematic patterns and conceptual associations within the texts.

3. Results

3.1. Participant Characteristics and REFLECT Scores

A total of 28 participants were included in the educational intervention: 25 anesthesiology residents (89.3%) and 3 clinical tutors (10.7%). The cohort consisted of 19 males (67.9%), and 9 were female (32.1%). Participants represented different stages of postgraduate training. All participants completed the reflective questionnaire. Each prompt referred to a selected excerpt from Chapter 8 of the novel *What Are You Going Through* by Sigrid Nunez. Chapter 8 was selected because it explicitly addresses the protagonist’s confrontation with terminal illness and assisted dying, offering a concentrated narrative exploration of autonomy, vulnerability, and relational responsibility at the end of life (Table 1).

Table 1. Narrative reflection questionnaire: open-ended prompts by thematic domain.

	Narrative Domain	Prompt
1	Emotional Resonance	What emotions did the film and book evoke in you?
2	Place of Death	After reading the text from “In the end, it’s a choice between two places. . .” to “you don’t know the day or the hour”, what do you think is the protagonist’s attitude in choosing the place of death?
3	Companionship in Death	After reading the text from “I need someone who understands all this” to “people have to work”, what do you think is the protagonist’s attitude in choosing who will accompany her at the moment of death?
4	Timing of Death	After reading the text, from “It’s not like she picked a specific date” to “I want a little peace”, what do you think is the protagonist’s attitude in choosing the timing of her death?
5	Manner of Death	After reading the text from “I won’t go out in a state of crushing anguish” to “I want everything to happen in an orderly way”, what do you think is the protagonist’s attitude in choosing the manner of her death?

Legend: Five open-ended reflective prompts used for the narrative medicine tool “close reading”, each corresponding to a narrative domain: emotional resonance, place, companionship, timing, and manner of death. Each prompt refers to a selected excerpt from Chapter 8 of *What Are You Going Through* by Sigrid Nunez, encouraging participants to explore the protagonist’s end-of-life choices from different narrative perspectives.

Residents achieved a mean REFLECT score of 2.48 (SD 1.05), corresponding predominantly to the level of thoughtful reflection. Tutors obtained a mean score of 2.00 (SD 0.85). Female residents demonstrated higher mean REFLECT scores compared to male residents (3.06 vs. 2.09). Tutors did not consistently achieve higher reflective scores than residents. Given the small and unbalanced sample, no inferential statistical analysis was performed, and these findings are reported descriptively. Descriptive characteristics of the sample and corresponding REFLECT scores are summarized in Table 2.

Table 2. Demographic characteristics and REFLECT scores of participants.

Characteristic	N (%)	Mean REFLECT Score ± Standard Deviation
Total Participants	28 (100%)	-
Role		

Table 2. Cont.

Characteristic	N (%)	Mean REFLECT Score $\pm$ Standard Deviation
Residents	25 (89.3%)	2.48 $\pm$ 1.05
Tutors	3 (10.7%)	2.00 $\pm$ 0.85
Gender		
Male	19 (67.9%)	2.09 $\pm$ 0.74
Female	9 (32.1%)	3.06 $\pm$ 0.98
Training Level		
First-year residents	14	2.55 $\pm$ 1.00
Second-year residents	8	2.00 $\pm$ 1.19
Third-year residents	3	2.66 $\pm$ 0.48

Legend: Participants' roles, gender, and training levels, alongside their corresponding mean REFLECT scores and standard deviations.

3.2. Qualitative Analyses and Comparison with AI-Generated Narratives

In addition to quantitative assessment, qualitative analyses were conducted to explore linguistic and thematic patterns. Lexical and semantic analyses were performed using T-LAB software. These results were visually represented through word clouds and semantic maps, which were later used as pedagogical tools during debriefing sessions. Word clouds revealed recurrent terms such as sadness, choice, and control, indicating that participants frequently intertwined emotional and ethical dimensions in their reflections on dying care (Figure 1). Semantic mapping further illustrated conceptual bridges between affective and moral language (Figure 2).



Figure 1. Word clouds of the most frequently used words by the participants. Legend: This word cloud was generated from the written responses of 28 participants. The size and color of each word indicates its frequency; larger and warmer-colored words (ranging from light blue to purple) were mentioned more often in the students' narratives.

The same reflective prompts were submitted to ChatGPT; application of the REFLECT rubric to these outputs was exploratory. ChatGPT achieved a mean score of 3.0, surpassing 63% of the participants. ChatGPT consistently produced coherent, logically structured, and apparently empathic narratives, achieving intermediate-to-high REFLECT scores in all domains (Table 3). AI-generated responses provided a comparative reference for discussion.

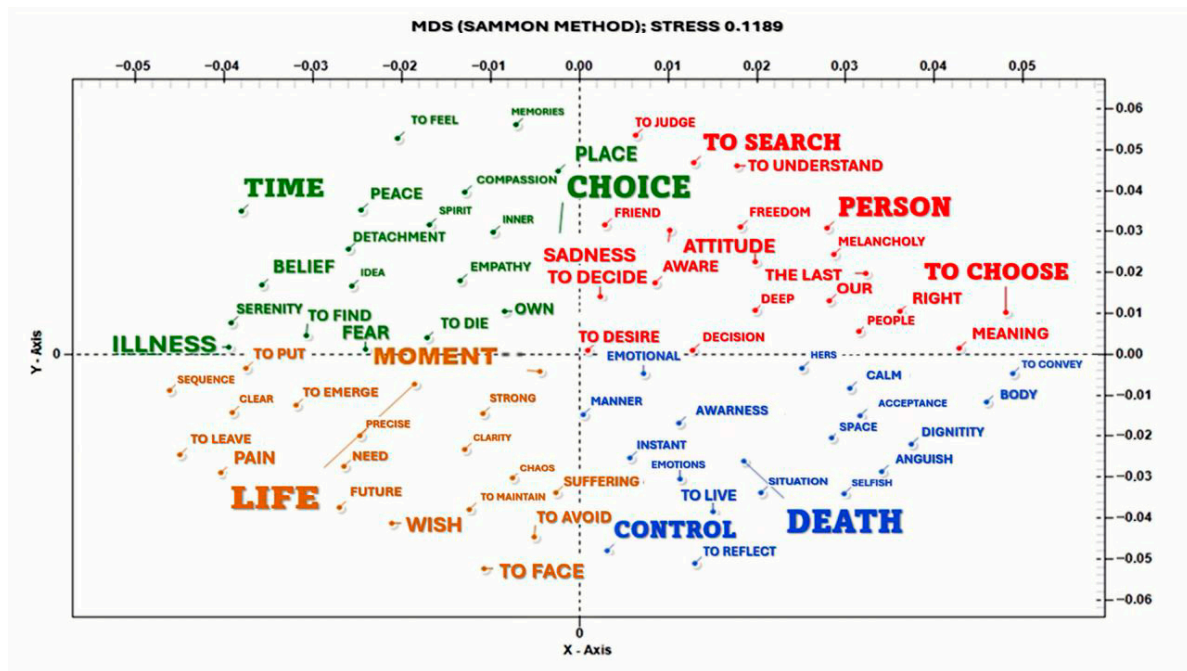


Figure 2. Multidimensional scaling map. Legend: This semantic map, generated using T-lab software through co-occurrence analysis, represents relationships among the most significant lemmas emerging from participants' open-ended responses. Colors represent clusters identified through hierarchical clustering.

Table 3. Comparison between ChatGPT's and some students' responses (score 4).

ChatGPT's Response	Anesthesiology Resident's Response (Score 4)
It evoked deep, layered emotions: tenderness for the delicate care between the two friends; sadness at the calm, lucid preparation for a final goodbye; admiration for the sick woman's quiet strength in refusing a heroic narrative; and fear, facing the profound question of how we would want to accompany or be accompanied at the end of life.	I empathized deeply with the protagonist. Her lucid, dignified approach to death mirrored my experience: a moment that, if faced with awareness, can become liberating. I share her wish to stay in control and avoid unnecessary suffering. It made me reflect on how painful—but necessary—it is to accept that letting go can bring peace. I felt both sadness and relief in her serene, realistic way of facing illness.
The protagonist chooses the place of her death with a surprising clarity, almost detached: she's not driven by emotions or sentimentality, but seeks a simple, quiet place to live her final days with dignity. Her choice is an act of self-care, a final gesture of autonomy, where death is not forced but embraced with serenity, far from nostalgia and memories that might trap her.	The choice of place is both practical and symbolic: she wants a peaceful, neutral space free from emotional chaos and nostalgic memories. She creates a sanctuary for her last days, protecting both herself and those around her. I think this detachment helps her stay firm in her decisions: to die with dignity, which I believe should be the right of every person.
The protagonist chooses her companion for the final moment with rationality and protection: she wants someone who deeply understands her choice without judging her, but who is also not so emotionally involved that it turns into a dramatic experience.	I saw it as a choice balanced between clarity, trust, and emotional distance. She looks for someone who truly understands her decision without being overwhelmed by it, exactly what I would want too. I share the fact that being someone outside her daily life feels like a way to protect them both from excessive emotional weight

Legend: This table presents selected responses generated by ChatGPT-4 alongside students' responses that scored 4 on the modified REFLECT rubric. All responses have been translated into English, preserving their original tone and content.

4. Discussion

The results of this report describe a narrative-based intervention aimed at fostering reflective capacity and empathic awareness among anesthesiology residents in the context of dying care. When considered together, quantitative REFLECT scores and qualitative lexical–semantic findings provide complementary insights into residents’ reflective engagement. Quantitative assessment through the REFLECT rubric offered an evaluation of the structural depth of reflection, whereas qualitative lexical–semantic analysis explored thematic content and conceptual associations emerging from participants’ narratives.

The findings suggest that structured narrative approaches integrating literature and cinema can provide a meaningful educational space in which young doctors engage with the emotional and ethical dimensions of end-of-life care, dimensions that are often insufficiently addressed within technically oriented curricula.

The analysis of participants’ reflective writings revealed recurrent emotional and ethical themes, including sadness, choice, control, and accompaniment. The application of the REFLECT rubric indicated that most participants did not reach the highest levels of reflective capacity. Reflective writings frequently revealed difficulties in articulating deep empathic engagement, highlighting the need for structured educational interventions to support the development of reflective and relational skills.

Although subgroup comparisons were not planned *a priori* and the study was not powered for inferential analysis, some descriptive patterns deserve cautious reflection. Female participants demonstrated higher mean REFLECT scores compared to male participants. While derived from a small and unbalanced sample, this pattern is consistent with research suggesting gender-related differences in empathic engagement and reflective expression, and should therefore be interpreted as hypothesis-generating rather than confirmatory. Similarly, the non-linear distribution of REFLECT scores across training years may reflect fluctuations in reflective engagement during transitional phases of professional development, possibly influenced by increasing clinical responsibility and workload. Larger longitudinal studies would be required to clarify these potential trajectories.

A distinctive element of this intervention was the inclusion of generative artificial intelligence as a comparative pedagogical stimulus. AI-generated narratives achieved higher REFLECT scores than a majority of participants, largely due to their linguistic coherence and structured organization. Rather than indicating superior empathic capacity, this apparent superiority highlights the risk of equating empathy with linguistic fluency or stylistic coherence. Algorithms can recognize and reproduce patterns, thematic clusters, and lexical associations, generating narratives that appear nuanced and emotionally insightful, qualities traditionally associated with human experience, yet they remain detached from lived experience and lack genuine moral sensitivity.

This raises a compelling question: can AI enhance our comprehension of individuals? As AI becomes increasingly significant in healthcare, should we explore its capacity to understand patient emotions? And if we begin to outsource the reading of emotional cues, do we risk surrendering the very part of care that defines us—the human presence?

This study has several limitations. The intervention was exploratory and conducted within a single residency program, with a small and unbalanced sample. Quantitative findings were therefore descriptive and not intended for inferential generalization. Future studies with larger and more balanced samples may further clarify how reflective capacity develops across training stages.

5. Conclusions

This brief report suggests that integrating narrative medicine approaches, such as literature and cinema, into anesthesiology training can support the development of reflective capacity and empathic awareness in end-of-life care.

The use of generative artificial intelligence as a pedagogical mirror may further enhance critical reflection by prompting trainees to critically examine what constitutes meaningful empathic engagement in clinical practice.

In an era in which algorithms can convincingly sound human, medical education bears the responsibility of ensuring that learners continue to cultivate what cannot be automated: the capacity for presence, moral engagement, and human connection when accompanying patients at the end of life.

Empathy should therefore not be treated as an incidental byproduct of clinical exposure, but as a professional competency requiring deliberate and structured cultivation. Relational skills are not optional supplements to technical expertise, but foundational elements of clinical competence.

Author Contributions: Conceptualization, methodology, investigation, data analysis, writing—original draft, writing—review and editing, and supervision: A.L.P., G.S. (Giuliana Scarpati), G.S. (Giulia Savarese) and O.P. All authors made the same contribution. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This educational study involved voluntary participation of healthcare professionals and did not involve patients or sensitive personal data. All participants provided informed consent prior to participation. The survey was conducted anonymously, in accordance with the principles of the Declaration of Helsinki, and followed the code of ethics and practices established by the American Association of Public Opinion Research (AAPOR). According to Italian national regulations, Ethics Committee approval is required for clinical trials and interventional studies involving patients or sensitive personal data, whereas anonymous, non-interventional educational research conducted among healthcare professionals does not require Ethics Committee approval. This is consistent with the scope of Ethics Committees as defined by Ministerial Decree of 8 February 2013 and subsequent amendments, which defines the organization and functioning of Ethics Committees within the national healthcare system. Accordingly, this study did not require IRB/Ethics Committee approval prior to data collection.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Aggregated data supporting the findings of this study are available from the corresponding author upon request.

Acknowledgments: Generative artificial intelligence was used exclusively as part of the educational intervention and for comparative pedagogical purposes. ChatGPT (OpenAI) was employed to generate narrative responses to the same reflective prompts administered to participants, without contextual input or prompt engineering. Generative AI was not used for data analysis, interpretation of participants' responses, or manuscript writing.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Karlsen, M.M.W.; Heyn, L.G.; Heggdal, K. Being a patient in the intensive care unit: A narrative approach to understanding patients' experiences of being awake and on mechanical ventilation. *Int. J. Qual. Stud. Health Well-Being* **2024**, *19*, 2322174. [[CrossRef](#)] [[PubMed](#)]
2. Costa-Drolon, E.; Verneuil, L.; Manolios, E.; Revah-Levy, A.; Sibeoni, J. Medical Students' Perspectives on Empathy: A Systematic Review and Metasynthesis. *Acad. Med.* **2021**, *96*, 142–154. [[CrossRef](#)] [[PubMed](#)]

3. Rosenwohl-Mack, S.; Dohan, D.; Matthews, T.; Batten, J.N.; Dzung, E. Understanding Experiences of Moral Distress in End-of-Life Care Among US and UK Physician Trainees: A Comparative Qualitative Study. *J. Gen. Intern. Med.* **2021**, *36*, 1890–1897. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
4. Charon, R.; Dasgupta, S.; Hermann, N.; Irvine, C.; Marcus, E.R.; Colón, E.R.; Spencer, D.; Spiegel, M. *The Principles and Practice of Narrative Medicine*; Oxford Academic: New York, NY, USA, 2016. [[CrossRef](#)]
5. Palla, I.; Turchetti, G.; Polvani, S. Narrative Medicine: Theory, clinical practice and education—A scoping review. *BMC Health Serv. Res.* **2024**, *24*, 1116. [[CrossRef](#)] [[PubMed](#)]
6. Komasa, N.; Yokohira, M. Generative Artificial Intelligence (AI) in Medical Education: A Narrative Review of the Challenges and Possibilities for Future Professionalism. *Cureus* **2025**, *17*, e86316. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
7. Sorin, V.; Brin, D.; Barash, Y.; Konen, E.; Charney, A.; Nadkarni, G.; Klang, E. Large Language Models and Empathy: Systematic Review. *J. Med. Internet Res.* **2024**, *26*, e52597. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
8. Nunez, S. *What Are You Going Through*; Riverhead Books: New York, NY, USA, 2020.
9. Almodóvar, P. *The Room Next Door*. Film. 2024. Available online: <https://www.imdb.com/title/tt29439114/> (accessed on 18 December 2025).
10. Pokpalagon, P.; Chaiviboontham, S.; Siripitayakunkit, A.; Junda, T.; Jaiboon, P.; Khunpinit, K. Quality of end-of-life care, quality of dying and death, and grief in bereaved family caregivers. *BMC Nurs.* **2025**, *24*, 1382. [[CrossRef](#)]
11. Leijenaar, E.; Eijkelboom, C.; Milota, M. “An invitation to think differently”: A narrative medicine intervention using books and films to stimulate medical students’ reflection and patient-centeredness. *BMC Med. Educ.* **2023**, *23*, 568. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
12. Wald, H.S.; Borkan, J.M.; Taylor, J.S.; Anthony, D.; Reis, S.P. Fostering and evaluating reflective capacity in medical education: Developing the REFLECT rubric for assessing reflective writing. *Acad. Med.* **2012**, *87*, 41–50. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.