# Psychotherapy, artificial intelligence and adolescents: ethical aspects

LINDA ALFANO[1], IVANO MALCOTTI[2], ROSAGEMMA CILIBERTI[1]
[1] Department of Health Sciences, University of Genoa, Genoa, Italy; [2] Società Genovese di Cremazione, SO. CREM., Italy

## Keywords

Mental health app • Artificial intelligence • Mental disorder • Ethics • Psychotherapy

## Summary

*Artificial intelligence (AI) has rapidly advanced in various domains, including its application in psychotherapy. AI-powered psychotherapy tools present promising solutions for increasing accessibility to mental health care. However, the integration of AI in psychotherapy raises significant ethical concerns that require thorough consideration and regulation to ensure ethical practice, patient safety, and data privacy. This article discusses the ethical considerations surrounding the utilization of AI in psychotherapy, emphasizing the need for responsible implementation, patient privacy, and the human-AI interaction.*

*The challenge raised by the use of artificial intelligence requires a comprehensive approach. Schools, in particular, are crucial in providing both knowledge and ethical guidance, helping young individuals decipher the complexities of online content. Additionally, parental support is essential, requiring the provision of time, fostering relationships, encouraging dialogue, and creating a safe environment to share experiences amidst the intricacies of adolescence. Reimagining social and healthcare services tailored for adolescents is equally crucial, taking into account recent societal changes. The integration of AI in psychotherapy has vast potential to transform mental healthcare. However, ensuring its accuracy and effectiveness demands a proactive approach to address associated ethical considerations. By adopting responsible practices, preserving patient autonomy, and continually refining AI systems, the field can leverage the benefits of AI in psychotherapy while maintaining high ethical standards.*

## Introduction

The introduction of artificial intelligence (AI) in the treatment of mental distress is challenging a range of assumptions and certainties within the professional and epistemological landscape of clinical psychology, raising ethical questions about its impact in psychotherapy, as well as the effects in patients and psychotherapists [1, 2].

The integration of chatbots and other AI apps into psychotherapy doesn't just impact the traditional biomedical ethical concerns of autonomy, beneficence, non-maleficence, and justice; it significantly influences the trust dynamics and relationships between patients and psychotherapists [3, 4].

The increasingly significant proliferation of companies providing digitally-mediated psychotherapy services has begun to gain prominence, particularly in conjunction with the recent pandemic situation [5].

While official statistics providing information about the actual number of patients utilizing these services are not yet available, it is noteworthy that there is a growing number of digital agencies offering such support. Unlike the past, these companies are providing new and valuable insights into the extent of mental distress in Italy, as they are no longer exclusively reliant on data provided by the public healthcare system.

AI-driven tools have proven highly effective in delivering cost-effective and accessible therapeutic solutions. However, these new healthcare services prompt multiple reflections regarding the specific offerings of these applications, the characteristics of the patients accessing these platforms, the interactions created with the intervention recipients, the therapy vision conveyed by marketing strategies, and the effectiveness of the treatment path [6]. In addition, these advancements in AI for mental health, introduce ethical challenges around its uses in psychotherapy that require in-depth consideration on the topics of respect for autonomy, access to treatment, non-discrimination, fulfillment of people's needs, protection of dignity, data privacy and much more [3].

From the limited and still incomplete data provided by digital agencies, it can be inferred that the most prominent users seeking these services primarily consist of adolescents and post-adolescents, aged between 16 and 25 years [7, 8].

These are young individuals who were born and raised in a world deeply immersed in digital technology, a generation accustomed to mediating all forms of relationships and communication through smartphones, social networks, messaging apps; a generation that experienced the most significant moments of their growth within their homes due to the pandemic [9].

These young people have a different perspective on psychotherapy compared to previous generations, even to adults only slightly older (30-40 years old) who grew up with the image of a patient lying on a couch and talking to a psychotherapist physically present in the room.

With the advent of new technologies and the introduction of digital forms of psychotherapy, the very techniques, tools of the discipline, or the structural rituals of the clinical encounter concerning the physical environment, space, time, and the centrality of the setting have indeed been questioned [10].

Historically, all approaches to mental health have recognized the importance of two fundamental aspects of care: restoring meaning to the dimensions of time and space, often compromised by illness (Eugene Minkowski - Ludwig Binswanger), and externally reorganizing the individual's relationship with these coordinates of reality when the individual themselves was temporarily unable to do so [11].

The tools used for this purpose were precisely represented by those rituals (regularly attending the therapist's office, respecting session times, scheduling appointments in advance, interrupting during the vacation period, and so on), which AI has profoundly altered, giving rise to heated debates.

In digital sessions, access to the clinician's office is no longer necessary, the schedule is flexible, and distances are no longer a reason for interruption. Digital psychotherapies overcome the limits of territoriality, allowing access to treatments for individuals who traditionally had to forego them due to work or geographic distance from treatment facilities [12, 13]. Such innovations therefore allow the possibility of democratizing medicine as a perspectives of justice and care, reducing the gap in access to mental health, which particularly affects the most vulnerable individuals.

On the other hand, there is significant criticism regarding the possibility that such apps might be profit-driven, designed to generate revenue by misleading users with false promises of improving their health. Other concerns raised pertain to the lack of conclusive evidence regarding their effectiveness in diagnosing and treating mental illnesses, as well as the absence of humanistic qualities in chatbots – such as empathy, reflective ability, professionalism, and reliability – essential for delivering widespread mental health treatments [3].

## Technological innovation and mental health

This potential to receive care potentially anywhere and at any time is changing the perception of psychotherapy, enabling an increase in patient intake. However, it also gives rise to new challenges [14].

One of the most significant challenges is related to dropouts, which are associated with the ease of access and familiarity with the digital medium [15, 16]. This particularly affects younger age groups who, after a few months of sessions without observing significant changes, tend to prematurely abandon the therapy and consider it ineffective. As is known, this phenomenon also affects the individual's future willingness to seek help from another therapist. Research shows that for the subsequent five years, individuals, despite continued distress, do not re-engage with any treatment system [17, 18].

The questioning of the fundamental pillars of psychotherapy, particularly the centrality of the setting, a concept common to all clinical approaches regardless of their theoretical foundations, prompts reflection on the epistemological framework of the discipline itself.

The majority of current psychotherapists might assert that digital offerings cannot and should not be identified or recognized as genuine psychotherapy. However, there is also a significant number of scholars who advocate a more cautious approach, suggesting a suspension of judgment to further explore the subject and the cultural changes surrounding the emergence of these new forms of mental health care [19].

Currently, AI governs the majority of algorithms in the online services we interact with daily, It guides our searches on search engines and the advertisements we are continuously exposed to on the internet [20].

This raises a question: based on what criteria will patients be directed to choose a psychotherapist? What elements will the ranking system used by various platforms be based on? Could a ranking of psychotherapists be established based on their professional credentials, or should we consider that this ranking will be calibrated based on user evaluations using parameters that may not necessarily be meaningful, such as the duration of the therapeutic relationship? And what will be considered as positive? Brief therapy or long-term therapy? This presents a challenge, as some schools of thought, like strategic psychotherapy, consider the brevity of the treatment process as a strength, while others, like psychodynamic approaches, view this as a weakness.

If, as can be easily imagined, algorithms decide which intervention is qualitatively more valid, there will be consequences in terms of education as well. Education will inevitably be steered towards models of schools that better cater to the needs of the market, even if they are not based on studies and research backed by scientific evidence. This perspective should indeed prompt significant actions to protect the health of citizens, not only by professional associations and legislators but also by the entire European community.

In the United States, the situation is even more complex, as the government has already authorized several insurance companies to recognize the validity of treatments performed through digital psychotherapy [21].

Among these, some applications targeted at young individuals suffering from depression are equipped with algorithms that can identify, based on linguistic exchanges with the chatbot, certain patterns and words that may indicate suicidal risk [22, 23]. These patterns can be identified by frequency, memory of past situations, or data uploaded by the psychotherapist.

In this case, the machine proves to be more efficient than even the most experienced psychotherapist because it possesses vastly superior computational power, allowing it to analyze billions of data points in minutes and rapidly discriminate high-risk situations. However, the responsibility for treatment remains within the domain

of the therapist, both for ethical and professional reasons that emphasize the irreplaceable value of the human relationship and for legal accountability concerns.

However, the integration of digital tools within the context of a therapeutic relationship between a professional and a patient raises ethical concerns about how mental health apps will impact the therapeutic relationship. For instance, the use of this technology might diminish the therapist's monitoring role, thereby heightening therapeutic risks for patients.

AI could, therefore, play a sort of frontline role (like first aid), enabling the assessment of emergency situations and promptly notifying the clinician. Other applications are aimed at individuals with eating disorders or the treatment of initial psychotic episodes [24]. In these cases, doubts arise about the AI's actual ability to fully meet its intended purposes. While algorithms can identify patterns, they do not possess human intelligence. This highlights the need for collaboration between AI and human therapists to achieve the best results, where AI can provide valuable insights and rapid assessments, while human therapists bring their expertise, ethical considerations, and nuanced understanding of individual cases to the treatment process.

It's therefore crucial to have a human therapist involved in the process to oversee and interpret AI-generated insights. The absence of human judgment might indeed result in misinterpretation or misdiagnosis of patient concerns, potentially causing harm. Patients should be aware of AI involvement in their therapy and its implications. They must provide informed consent and retain autonomy in decisions regarding their treatment. Ensuring transparency and comprehensibility in AI's role is vital for maintaining patient trust.

## Digital therapies for adolescents: Risks and benefits

Certainly, the introduction of AI in the field of mental health brings benefits but also risks and weaknesses. The first limitation relates to the machine's inability to grasp the emotional, symbolic, relational, and anthropological dimensions of the data it is meant to interact with, as it can only decipher the empirical reality. A word holds no evocative value for a machine; it lacks emotional meaning. However, every therapist knows that to understand the individual and their anxieties, the internal and emotional resonances of words are more relevant. For example, consider the power of certain words like "mother", "death", "love", or seemingly innocuous ones for algorithms but bearers of disturbing memories for an individual.

The lack of human supervision in patient interactions not only results in flat and uniform narratives but also poses the risk of dangerous self-diagnosis, without any opportunity for correction [25].

Ensuring that AI systems comply with data protection regulations and safeguarding the privacy of patients' information is an ethical imperative in relation to adverse consequences for individuals and society (Article 8 of the European Convention on Human Rights).

Another ethical and professional concern pertains to privacy and data protection, as well as the type of content being discussed. Where will the data from these conversations be stored? Who will be responsible for the clinical records, the conversation history stored in the cloud? Companies managing this data could go out of business or sell the data, for instance, to prospective employers seeking to acquire personality profiles of potential employees.

## Artificial intelligence and the psychological distress of the youth world

As previously mentioned, the generation turning to these new forms of digital psychotherapy is predominantly composed of adolescents [26]. These are young individuals who are more comfortable interacting with digital media now prevalent in diverse areas of social life (technology, healthcare, finance, culture and more), as they have grown up with it [27]. They often feel misunderstood by adults and are driven by loneliness to seek a non-human other that can listen and be accessed without physical, economic, and social barriers.

As Pietropolli Charmet and Lancini assert, today's youth are growing up in a fragile family environment where both parents usually work, mothers and fathers play distant roles, and relationships lack physical presence. Parents idealize, immortalize, constantly photograph, and put their children on display, burdening them with expectations of achievement and popularity [28, 29].

These images of a magnificent, grandiose self, free from limitations and projected toward every success, expose adolescents to dangerous feelings of disappointment. These emotions are related to the shame of not becoming what was expected, shame for their vulnerabilities, and shame regarding their bodies. The contemporary illnesses affecting adolescents arise from the ideal of an inflated self that collapses during adolescence and leads to the emergence of so-called shame-based pathologies that include social withdrawal, substance use, eating disorders, depression, anxiety, and severe self-harming behavior.

Adolescence, often described as the age of omnipotence and invincibility, is, in reality, the age of confronting limits and one's own vulnerability. Unlike children who fear the death of a parent, adolescents discover that they themselves can die. They uncover their own limitations and vulnerabilities.

The problem of dropping out of school, which predominantly affects males and has been on the rise in recent years, is also related to shame-based pathologies, particularly feelings of inadequacy and unpopularity. It's about the inability to cope with disappointment, frustration, and the challenges of growing up. Furthermore, today's adolescents are often unmotivated to pursue educational qualifications that may have no practical usage and that no longer offer the hope of a better life, as they did in the past.

The generation gap reflects an ever increasing distance between students and teachers, who no longer serve as strong role models for young people. While adults lose credibility and authority, peers, mass media subculture, the internet, YouTubers, and influencers gain influence. The internet (the virtual world) is one of the modern contexts in which adolescents face the challenge of identity. For them, the internet is a stage that can offer popularity and success but can also lead to dangerous setbacks [9].

## Conclusions

Online, young people are turning to new forms of psychotherapy that they perceive as more accessible and less judgmental. Robots and animals lack ulterior motives, which helps adolescents overcome shame and engage in discussions on intimate and difficult topics [30].

In reality, the exponential increase in adolescent distress, evident from the rising number of adolescents admitted to psychiatric wards, the unprecedented suicide rates, and the significant school dropout rates, requires complex, coordinated efforts among institutions, companies, and associations that are in daily contact with young people. These interventions, however, have so far been isolated and fragmented.

In this context, it is crucial and a top priority to ensure that schools once again provide knowledge and ethical tools to help young people decode the messages of the internet [31]. Additionally, parents need support to offer time, relationships, dialogue, and a space to share their children's pain and support the complexities of growing up.

It is also essential to rethink social and healthcare services for adolescents, taking into account the changes that have occurred in recent years [32]. Despite obvious needs during this phase of life, there are still significant barriers to accessing existing services in local communities. All of this is challenging to accomplish when continuous bureaucratic and political obstacles make every change and progress arduous.

Acknowledging the problem and developing appropriate, comprehensive responses is now essential.

The integration of AI in psychotherapy holds immense potential to revolutionize mental health care. However, to ensure their correctness and effectiveness, it is imperative to address the associated ethical concerns. By incorporating responsible practices, maintaining patient autonomy, and continually refining AI systems, the field can harness the benefits of AI in psychotherapy while upholding ethical standards.

## Financial support

## Conflict of interest statement

The authors declare no conflict of interest.

## Authors' contributions

All authors conceived the study and contributed to the preparation of the manuscript related to their sections and approved the final version to be submitted.

## References

[1] Yan WJ, Ruan QN, Jiang K. Challenges for Artificial Intelligence in Recognizing Mental Disorders. Diagnostics (Basel) 2022;13:2. https://doi.org/10.3390/diagnostics13010002

[2] Ciliberti R, Schiavone V, Alfano L. Artificial intelligence and the caring relationship: ethical profiles. Medicina Hist 2023;7:e2023016.

[3] Manríquez Roa T, Biller-Andorno N, Trachsel M. The Ethics of Artificial Intelligence in Psychotherapy. Oxford Handbook of Psychotherapy Ethics. Oxford: Oxford University Press 2021. https://doi.org/10.1093/oxfordhb/9780198817338.001.0001

[4] Swartz HA, Novick DM. Psychotherapy in the Digital Age: What We Can Learn From Interpersonal Psychotherapy. Am J Psychother 2020;73:15-21. https://doi.org/10.1176/appi.psychotherapy.20190040

[5] Fiorentino V, Romakkaniemi M, Harrikari T, Saraniemi S, Tiitinen L. Towards digitally mediated social work – the impact of the COVID-19 pandemic on encountering clients in social work. Qual Soc Work 2023;22:448-64. https://doi.org/10.1177/14733250221075603

[6] Stoll J, Müller JA, Trachsel M. Ethical Issues in Online Psychotherapy: A Narrative Review. Front Psychiatry 2020;10:993. https://doi.org/10.3389/fpsyt.2019.00993

[7] Orsolini L, Pompili S, Salvi V, Volpe U. A Systematic Review on TeleMental Health in Youth Mental Health: Focus on Anxiety, Depression and Obsessive-Compulsive Disorder. Medicina (Kaunas) 2021;31;57:793. https://doi.org/10.3390/medicina57080793

[8] Mortimer R, Somerville MP, Mechler J, Lindqvist K, Leibovich L, Guerrero-Tates B, Edbrooke-Childs J, Martin P, Midgley N. Connecting over the internet: Establishing the therapeutic alliance in an internet-based treatment for depressed adolescents. Clin Child Psychol Psychiatry 2022;27:549-68. https://doi.org/10.1177/13591045221081193

[9] Khalaf A M, Alubied A A, Khalaf A M, Rifaey AA. Impact of social media on the mental health of adolescents and young adults: A systematic review. Cureus 2023;15:e42990. https://doi.org/10.7759/cureus.42990

[10] Berler M, Trub L, Magaldi D. The therapist's evolving public self in a digitally disrupted marketplace. J Psychother Integr 2023;33:285-301. https://doi.org/ https://doi.org/10.1037/int0000295

[11] Binswanger L. Il problema dello spazio in psicopatologia. Macerata: Quodlibet 2022.

[12] Weightman M. Digital psychotherapy as an effective and timely treatment option for depression and anxiety disorders: Implications for rural and remote practice. J Int Med Res 2020;48:300060520928686. https://doi.org/10.1177/0300060520928686

[13] Renn BN, Hoeft TJ, Lee HS, Bauer AM, Areán PA. Preference for in-person psychotherapy versus digital psychotherapy options for depression: survey of adults in the U.S. NPJ Digit Med 2019;11;2:6. https://doi.org/10.1038/s41746-019-0077-1

[14] McDonald A, Eccles JA, Fallahkhair S, Critchley HD. Online psychotherapy: trailblazing digital healthcare. BJPsych Bull 2020;44:60-6. https://doi.org/10.1192/bjb.2019.66

[15] Fredum HG, Rost F, Ulberg R, Midgley N, Thorén A, Aker JFD, Johansen HF, Sandvand L, Tosterud L, Dahl HJ. Psychotherapy dropout: Using the Adolescent Psychotherapy Q-set to explore the early in-session process of short-term psychodynamic psychotherapy. Front Psychol 2021;22;12:708401. https://doi.org/10.3389/fpsyg.2021.708401

[16] Cohen K A, Schleider JL. Adolescent dropout from brief digital mental health interventions within and beyond randomized trials, Internet Interv 2002;27:100496. https://doi.org/10.1016/j.invent.2022.100496

[17] de Bruin EJ, Meijer AM. The impact of online therapeutic feedback on outcome measures in Internet-CBTI for adolescents with insomnia. Sleep Med 2017;29:68-75. https://doi.org/10.1016/j.sleep.2016.05.017

[18] Victor P, Krug I, Vehoff C, Lyons N, Willutzki U. Strengths-based CBT: Internet-based versus face-to-face therapy in a randomized controlled trial. J Depress Anxiety 2018;7:2. https://doi.org/10.4172/2167-1044.1000301

[19] Imondi V. iConscious: L'Inconscio connettivo: Psicopatologia delle masse nell'era di Internet e analisi dell'Io digitale - 2020.

[20] Busuioc M. Accountable artificial intelligence: Holding algorithms to account. Public Adm Rev 2021;81:25-836. https://doi.org/10.1111/puar.13293

[21] Santoro E. La digital health a supporto della professione medica. Ric&Pra 2021;37:15-22. https://doi.org/10.1707/3556.35319

[22] Khan NZ, Javed MA. Use of artificial intelligence-based strategies for assessing suicidal behavior and mental illness: A literature review. Cureus 2022;14:e27225. https://doi.org/10.7759/cureus.27225

[23] Menon V, Vijayakumar L. Artificial intelligence-based approaches for suicide prediction: Hope or hype? Asian J Psychiatr 2023;88:103728. https://doi.org/10.1016/j.ajp.2023.103728

[24] Di Leo D. Terapie digitali ed intelligenza artificiale, applicazioni pratiche e quadro normativo. Tendenze nuove 2021;2:3-11. https://doi.org/10.32032

[25] Bauer M, Glenn T, Monteith S, Bauer R, Whybrow PC, Geddes J. Ethical perspectives on recommending digital technology for patients with mental illness. Int J Bipolar Disord 2017;5:1-14. https://doi.org/10.1186/s40345-017-0073-9

[26] Lehtimaki S, Martic J, Wahl B, Foster KT, Schwalbe N. Evidence on digital mental health interventions for adolescents and young people: Systematic overview. JMIR Ment Health 2021;8:e25847. https://doi.org/10.2196/25847

[27] Licata M, Larentis O, Tesi C, Fusco R, Ciliberti R. Tourism in the time of Coronavirus. Fruition of the "minor heritage" through the development of bioarchaeological sites, our proposal. Tourism Management 2020;21:19. https://doi.org/https://doi.org/10.3390/heritage4020042

[28] Lancini M. Il ritiro sociale degli adolescenti. La solitudine di una generazione iperconnessa. Milano: Raffaello Cortina 2019.

[29] Polli Charmet P. Fragili e spavaldi: ritratto degli adolescenti di oggi. Bari: Laterza ed. 2010.

[30] Opel DJ, Kious BM, Cohen IG. AI as a mental health therapist for adolescents. JAMA Pediatr 2023;177:1253-1254. https://doi.org/10.1001/jamapediatrics.2023.4215

[31] Gulino M, Patuzzo S, Baldelli I, Gazzaniga V, Merlo DF, Maiorana L, Murialdo G, Picozzi M, Armocida G, Cattorini P, Montaguti E, Bonometti S, Grossi AA, DeStefano F, Ciliberti R. Bioethics in Italian medical and healthcare education. A pilot study. Acta Biomed 2018;89:519-31. https://doi.org/10.23750/abm.v89i4.7238

[32] Ciliberti R, Alfano L, Baldelli I, De Stefano F, Bonsignore A. Self-determination, healthcare treatment and minors in Italian clinical practice: ethical, psychological, juridical and medical–legal profiles. Acta Biomed 2018;89:34-40. https://doi.org/10.23750/abm.v89i1.6368