Drivers of Well-Calibrated Trust in Human-AI Collaboration: A Scoping Review of Task Dynamics and Regulatory Frameworks

Dogtooth – Deliverable 1 – Final version

Abstract

The increasing use of artificial intelligence in healthcare requires workers to decide when and how to rely on these systems, making trust a key factor. This scoping review wants to identify the main factors influencing workers' trust in AI and how it can be well-calibrated. A literature search led to the selection of 48 studies, from which a framework was developed and divided into four paragraphs: AI system characteristics, human factors, task characteristics, and organizational context. A work use case regarding dentistry illustrates how these factors interact in a more practical way. The conclusion shows that trust in AI is a dynamic and context-dependent process that requires continuous human evaluation rather than a simple decision.

1 Introduction

Artificial intelligence is increasingly being integrated into healthcare practice, where it is used to support diagnosis, clinical decision-making, and treatment planning. As these systems become more used, healthcare professionals are required to interact with AI and decide whether and when to rely on its outputs.

However, trust in AI is not a straightforward process. Evidence from recent literature shows that AI systems are often characterized by variability in performance, uncertainty, and limited explainability, which can affect how their outputs are interpreted. For example, issues such as inaccurate predictions, lack of transparency (the "black box" problem, explained later on in the review), and the presence of bias can reduce confidence in AI systems. At the same time, excessive reliance on AI can lead to unsafe decision-making.

That is why it is important to introduce the concept of trust calibration, defined as the ability to appropriately adjust their trust on AI based on its capabilities and limitations. Trust calibration is influenced by the technical characteristics of AI systems, such as accuracy and reliability, but mostly by human factors, such as user expertise, training level, and familiarity with the technology. In addition, elements such as task complexity, risk level, and responsibility for outcomes play a crucial role in shaping trust.

Healthcare professionals must learn how to balance the benefits AI can give, such as increased efficiency and diagnostic support, with bigger concerns related to accountability, ethical implications, and system limitations.

While existing studies have explored the performance and ethical implications of AI in medicine, it has been given minor attention to the identification of the factors that determine whether healthcare workers trust AI systems and how that trust is calibrated in practice. Many studies focus either on the AI performance or on patient perspectives, leaving a gap in understanding the decision-making processes of workers dealing with AI. This study addresses the following research question: *"What factors determine whether a worker trusts AI to perform a task, and when is that trust well calibrated?"*
To answer this, we conducted a scoping review of the literature on AI in healthcare, identifying and organizing the key factors influencing trust and trust calibration. A specific application in dentistry is used as a worked example to show how these factors can apply to a more specific healthcare sector.

2 Method

We adopted a scoping review approach and followed PRISMA-ScR to identify and map the factors influencing workers' trust in AI in professional contexts.

The search was mainly conducted on Google Scholar, Scopus, ResearchGate, WebofScience and PubMed, selected for their broad academic coverage and access to scientifically reliable sources [Appendix A]. A time restriction was applied to only include studies that could ensure focus on recent developments in the use of AI in healthcare [Appendix B].

The search was conducted using combinations of keywords related to artificial intelligence, medicine, dentistry, and trust-related decision-making [Appendix C].

The selection process was systematically divided into steps: title screening, abstract screening, and full-text review. We evaluated articles based on inclusion and exclusion criteria regarding their relevance to AI in professional fields and their contribution to understanding trust calibration [Appendix D].

Some studies were partially included as only specific sections provided generalizable insights relevant to the research question [the screening process is summarized in Appendix F, Appendix G and Appendix H]. The included studies were then analyzed to extract relevant factors, including accuracy, reliability, uncertainty, human oversight, task-related risk, data quality, and usefulness, which were grouped into broader themes to support the analytical framework [Appendix G].

AI tools, such as ChatGPT, Gemini, and Copilot, were used to support the summarization of articles, the first screening process, and the organization of the appendix. To ensure academic integrity and mitigate the risk of AI "hallucinations," the team implemented a strict verification protocol [Appendix E].

While AI supported the initial skimming of records, all inclusion decisions, interpretations, and final coding were performed and verified by the group [Appendix H and Appendix G]. Every citation and interpreted finding was manually cross-referenced against the original full-text source to ensure accuracy and prevent the inclusion of fabricated data.

2.1 Inclusion/Exclusion Criteria

Regarding the screening process, a large number of articles initially appeared relevant to the topic under investigation. However, a set of inclusion and exclusion criteria was applied to ensure consistency and relevance in the selection. The primary criteria included: publication date, context and target, topic, system scope, and source type. This systematic screening process enabled the identification of the most pertinent studies for the proposed framework, while excluding those that were not sufficiently aligned with the research objectives.

Table 1: Inclusion and Exclusion Criteria

Criterion	Inclusion	Exclusion
Date	After 2016	Before 2016
Context and Target	Medical professional contexts	Case-specific without general insight
Topic	Factors influencing trust, reliance and decision-making processes	Does not focus on the research question
System Scope	AI or algorithmic systems	No involvement with AI
Source Type	Credible and reliable sources	Unreliable sources

3 Framework

3.1 AI systems characteristics

An important factor in determining whether workers place trust in AI systems is their understanding of how these tools operate and perform. When AI systems are perceived as accurate and reliable [14] [18], workers are more likely to trust them and integrate them into their daily tasks.

However, the level of trust may vary depending on the specific work the AI system has to perform; for example, studies show that workers may trust AI differently when it is used for image recognition, risk prediction, or professional decision support [21] [44].

On the other hand, when AI tools are inconsistent or poorly functioning [7] [16], trust and confidence are reduced, particularly in professional contexts where decisions may directly impact significant outcomes.

One important aspect that workers should consider is that, even when AI systems appear accurate and reliable, their output may still be affected by uncertainty and unpredictability [19] [20]. In some cases, AI systems can generate incorrect or misleading information [17] [19] while presenting it with high confidence. This phenomenon is known as *AI hallucination* [12]. Such errors are often difficult to detect and may increase risks for stakeholders.

Another important factor that influences trust in AI systems is their level of transparency. Many AI tools operate as “black boxes” [27], meaning that they provide results without clearly explaining how those results were obtained. This lack of transparency can make it difficult for workers to interpret and verify AI outputs [4] [21]. Humans have a very high need to understand how things work to be able to trust them, and this lack of clarity may discourage them from relying on Artificial Intelligence, especially when high-stakes consequences are at risk.

Recent literature explains how AI systems that are easy to understand may help overcome this issue by making the decision-making process more interpretable and understandable for workers [29] [30].

Data quality and the presence of bias also play an extremely important role [14] [16] in the gain of trust in AI systems. AI models that have been trained on limited or biased datasets may produce unreliable or non-generalizable results [18], especially if applied to new contexts to them unknown. These limitations increase uncertainty and may reduce workers’ confidence in AI-supported decision-making.

The "black box" nature of AI manifests when a system provides high-confidence recommendations without exposing its underlying logic or mathematical weighting. This lack of transparency becomes dangerous when combined with algorithmic bias; for instance, a tool trained primarily on adult male datasets often exhibits decreased accuracy when applied to outlier groups, such as pediatric patients or underrepresented ethnicities. Without human validation to bridge these transparency gaps, such biases can go undetected, resulting in significant professional errors [27].
For example, in fields involving complex analysis and critical decisions, where AI could achieve important results, human validation remains necessary to ensure safe professional use [31].

Overall, our findings suggest that trust in AI systems is strongly influenced by perceived reliability, transparency, and consistency [21]. However, the presence of uncertainty, potential errors, and system limitations emphasizes the need for constant human oversight and validation, making it clear that trust in AI must be well-calibrated.

3.2 Human factors

Workers’ trust in AI is also strongly influenced by human-related factors. The level of expertise, knowledge and ability to objectively evaluate the AI system play an important role in how workers decide whether to trust AI systems [9] [15] [26].

When workers have enough knowledge to verify the results given by AI, they are more likely to use Automative tools appropriately.

On the other hand, when the task goes beyond their level of knowledge, it becomes more difficult to verify the outputs, increasing the risk in relying on the system too much without fully understanding whether the result is correct or not [12] [19] [20].
Becoming familiar with AI systems also has a significant impact on trust. Workers who are more exposed to these tools and have received proper training [15] [26] tend to feel more confident in using them. A lack of knowledge and experience often leads to hesitation on relying on AI, even when technology could be useful.

Repeated exposure to AI tools during the education period and daily practice may increase familiarity and confidence, shaping trust over time [26].
Another important aspect is how the AI systems are perceived by the professional users. Some factors that could greatly damage trust-building could be concerns about errors, ethical issues, or even the possibility of being replaced [4] [5]. This shows that trust is not only based on how well the system works, but also on how users feel about it and how comfortable they are interacting with it [8] [9].

On the other hand, too much trust can also become a problem. If workers rely too heavily on AI without being able to fully verify the results, they may accept results without questioning the answers enough. This can lead to incorrect decisions and reduce their critical thinking during the decision-making process, leading to significant risks for stakeholders [28] [46].

3.3 Task characteristics

In this paragraph we want to analyze the nature of different tasks that play an important role in determining whether trust in AI should be given or not. Rather than replacing the workforce, AI is starting to be seen as a collaborative tool that can help human expertise. [34] [45].

In particular, the level of risk associated with a task [14] [21] influence how much workers are willing to rely on AI. In high-risk situations, such as critical decision-making, workers tend to be more cautious and are less likely to fully trust AI without additional proof.

Another important component is time [13]. In time-sensitive situations, workers may be more inclined to rely on AI systems to support faster decision-making, even if this involves a certain level of uncertainty. This can increase trust in AI, especially when immediate action is required.

The complexity of the task can also affect trust [7] [14]. When tasks are complex or go beyond what the worker knows, AI systems may be perceived as more useful, which can increase reliance. However, in these situations, the ability to verify the output is often overlooked, increasing the risk of over-trust. AI performance may vary depending on the specific professional task, meaning that trust should be calibrated differently based on the application [44] [48].

One final element to take into consideration is the possibility of verifying AI outputs. When workers can cross-check the results, they are more likely to trust AI systems. When verification is difficult [12] [19] or not possible at all, trust becomes harder to give out.

3.4 Organizational and contextual factors

One final factor we would like to consider is the organizational and contextual environment in which the AI tool would have to be used.

Issues related to responsibility and accountability [4] [5] can strongly influence how comfortable workers feel when relying on these systems.

In most professional environments, workers are ultimately responsible for their decisions, which can make them more cautious in trusting AI, especially when the consequences of errors are very significant.

Regulation also contributes to trust [4] [6]. The presence of clear guidelines, rules and standards can increase confidence in AI systems by making sure that they bring safety and quality.

Uncertainty regarding legal responsibility and lack of regulations may discourage workers from relying on AI, even when the technology is effective.

Another important aspect is how well AI systems integrate into how people are already used to working [6] [13]. When AI tools are incorporated smoothly into professional practice and support decision-making, workers are more likely to trust them. Well-integrated digital workflows may facilitate AI inclusion in the work process and increase trust by reducing disturbance to already established professional routines. [46] [47].

Overall, these findings suggest that trust in AI is not only an individual or technical issue but also depends on the larger environment in which technology is implemented. This brings to light the importance of considering regulations, laws, and leaving no space for legal grey areas in any specific field of work.

3.5 Interaction of Factors

Trust is not a static response to individual dimensions but a dynamic state emerging from their interaction. A critical synthesis of the literature reveals that the Over-Trust Risk is most acute when High Task Complexity intersects with Low User Expertise. In these instances, the worker's inability to verify the "black box" logic results in a statistical shift toward uncritical acceptance of algorithmic outputs.

On the other hand, the Under-Trust Barrier creates a paradoxical situation where objectively high-performing systems are rejected. This typically occurs when Persistent Legal Ambiguities and a Lack of Transparency align, forcing professionals to prioritize risk mitigation and personal accountability over the potential benefits of AI-supported efficiency. This suggests that "accuracy" alone is insufficient for calibration if organizational and system transparency factors are not simultaneously addressed.

Finally, Trust Calibration is fundamentally a developmental process of Familiarization. Beyond simple training, it requires repeated exposure during professional education to empower workers to identify "anomalous" cases. This evidence suggests that the most effective way to prevent both over-reliance and under-reliance is through a balanced integration of high system transparency and advanced user domain knowledge.

3.6 Trust Calibration Checklist

To put into practice the analysis of the framework, the following checklist provides direct questions that workers should ask before delegating a task to an AI system [28] [46].

Table 2: Questions to ask before delegating a task

Dimension	Question	High-risk indicator
System	"Can I interpret how the AI reached this result?"	The system operates as a "black box" with no transparency or explainability
Human	"Do I have the specific expertise to spot a subtle AI mistake in this task?"	The task exceeds the user's current knowledge or ability to verify the output
Task	"What are the consequences if this output is wrong?"	The task involves high-risk decision-making or irreversible outcomes
Organization	"Is there a clear protocol for accountability in case an error occurs?"	There is uncertainty regarding legal responsibility

Beyond the checklist, professionals should adopt a "double-check" protocol for any AI recommendation involving irreversible outcomes. This involves a secondary manual review or a peer consultation when the AI output contradicts clinical intuition, effectively bridging the gap between ethical theory and day-to-day operational reality.

4 Worked Use Case: AI in Dentistry

4.1 Real-world relevance

The factors previously identified can be observed by considering the use of AI in dentistry, in areas such as diagnosis, treatment planning, and clinical decision support. AI systems are increasingly being used to analyze clinical images, assist in identifying dental conditions, and support decisions such as specific treatments [10] [24]. Some AI tools have already been FDA-cleared (cleared by the Food and Drug Administration) for applications such as dental imaging analysis, supporting the practical relevance of this use case [25]. These applications vary from automated diagnostics and radiographic interpretation to treatment planning, workflow optimization, and decision support during clinical practice [6] [24].

4.2 AI system characteristics in dentistry

In this specific field, trust is influenced by the accuracy, reliability, and ability of verification of these tools. For example, while AI can support the analysis of radiographic images or intraoral scans, issues such as limited generalizability, lack of transparency, and potential bias may reduce confidence in the results [16]. The "black box" scheme of many AI systems makes it difficult for practitioners to fully understand how decisions are made, demonstrating the need of human validation [27].

Within this use case, the risk of bias is evident in tools used for radiographic interpretation. If an AI model is trained primarily on adult dental scans, its ability to identify conditions in pediatric patients or those with rare orthodontic anomalies may be compromised. Because the dentist cannot always "see" the math behind the AI's detection (the black box), they must manually verify each highlighted anomaly against their own clinical expertise to ensure the system is not hallucinating a diagnosis or missing a critical factor due to biased training data [27] [32].

4.3 Human factors

Human factors also play an important role. Dentists must have sufficient knowledge and training to evaluate AI outputs and include them in their decision-making process. A lack of familiarity with AI tools can lead to setbacks, while limited expertise may increase the risk of over-trust. In contrast, proper education and exposure to AI technologies can support more informed and balanced use.

4.4 Task characteristics

The characteristics of the task also influence trust. In high-risk clinical decisions, such as treatment planning, dentists are more likely to carefully verify AI recommendations before fully relying on them [14] [21]. At the same time, in situations where tasks are complex or time-consuming, AI may be perceived as particularly useful, increasing trust despite the presence of uncertainty.

Overall, the application of our research in the dentistry field illustrates how trust in AI is not determined by a single factor, but it emerges from the interaction between system, human capabilities, tasks, and organizational context. This confirms the relevance of our framework in understanding how workers consider and use AI in real-world situations.

5 Gaps And Future Of Work

Looking at the future of work, the integration of AI is likely to transform professional roles rather than fully replace them. Beyond healthcare, AI is expected to reshape the work dimension, creating both opportunities and concerns related to job opportunities and displacements [36]. In healthcare, AI is expected to act only as an assistance tool that supports human decision-making rather than fully replacing it.

In contrast with the growing number of studies on AI in healthcare, a lot of gaps still stand in understanding how workers develop and calibrate trust in these systems.

A large amount of the existing articles on this area of study focus either on the technical performance of AI or on ethical considerations, while not that many studies address the decision-making process of workers and how the decision of giving (or not giving) trust evolves in real-world scenarios. In particular, there is limited empirical evidence on how different factors interact in the real work field and how workers regulate their trust over time when exposed to AI systems.
Another important gap concerns the role of training and education. While several studies bring to light the importance of AI studies, there is still a lack of clear paths on how to actually prepare healthcare professionals to evaluate and use AI tools. On the same note, issues related to accountability and law regulation remain only partially addressed, creating uncertainty that may influence trust.

As for healthcare, AI is expected to act as a support tool that expands human decision-making, rather than fully automating and replacing it. This will require workers to develop new skills, such as the ability to interpret AI outputs, identify potential errors, and understand when the trust is to be given or when to trust their own instincts.

Future research should also focus on how trust can change in rapidly evolving diagnostic applications in fields such as pathology, endoscopy, and clinical decision support [21] [45] [46].

Overall, we believe that future research should focus on understanding how trust in AI develops over time, how it can be effectively calibrated in different contexts, and how training, regulation, and better design can support safer and more effective collaboration between workers and AI.

6 Conclusion

This scoping review explored the factors that determine whether workers trust AI and when that trust can be considered well calibrated. By analyzing the existing literature on this issue, it was possible to identify four main dimensions influencing trust: AI system characteristics, human-related factors, task characteristics, and organizational or contextual factors.

The research suggests that trust in AI is not a fixed or automatic response, but a dynamic process, modified by the interaction of all these dimensions. Factors such as accuracy, transparency, data quality, user expertise, task complexity, and accountability all contribute to how workers evaluate AI outputs and decide whether to trust them.

The worked use case in dentistry demonstrates how these factors interact in a real-world healthcare setting, confirming the hypothesis stated in the framework. At the same time, the literature shows that trust must be calibrated really carefully. Both under-reliance and over-reliance on AI may lead to unsafe outcomes, highlighting the need for constant human oversight and evaluation.

Overall, as AI becomes increasingly integrated into professional healthcare practice, understanding and supporting well-calibrated trust will be essential to maintain a safe, effective, and responsible human–AI collaboration.

References

[1] Vivek Kaul, Sarah Enslin, and Seth A. Gross. 2020. History of artificial intelligence in medicine. Gastrointestinal Endoscopy 92, 4 (10 2020), 807–812. doi:10.1016/j.gie.2020.06.040

[2] Oleg O. Varlamov, Dmitry A. Chuvikov, Larisa E. Adamova, Maxim A. Petrov, Irina K. Zabolotskaya, and Tatyana N. Zhilina. 2019. Logical, philosophical and ethical aspects of AI in medicine. International Journal of Machine Learning and Computing 9, 6 (12 2019), 881-886. doi:10.18178/ijmlc.2019.9.6.885

[3] Alexander S. Doyal, David Sender, Monika Nanda, and Ricardo A. Serrano. 2023. ChatGPT and artificial intelligence in medical writing: concerns and ethical considerations. Cureus 15, 11 (10 2023), e49071. doi:10.7759/cureus.43292

[4] Felix Busch, Lisa C. Adams, and Keno K. Bressem. 2023. Biomedical ethical aspects towards the implementation of artificial intelligence in medical education. Medical Science Educator 33, 4 (6 2023), 1007–1012. doi:10.1007/s40670-023-01815-x

[5] Farhud and Shaghayegh Zokaei. 2021. Ethical issues of artificial intelligence in medicine and healthcare. Iranian Journal of Public Health 50, 11 (11 2021), 2153–2165. doi: 10.18502/ijph.v50i11.7600

[6] Tao-Yuan Liu, Kun-Hua Lee, Arvind Mukundan, Riya Karmakar, Hardik Dhiman, and Hsiang-Chen Wang. 2025. AI in dentistry: innovations, ethical considerations, and integration barriers. Bioengineering 12, 9 (8 2025), 928. doi:10.3390/bioengineering12090928

[7] Amal Alfaraj, Toshiki Nagai, Hawra AlQallaf, and Wei-Shao Lin. 2025. Race to the moon or the bottom? Applications, performance, and ethical considerations of artificial intelligence in prosthodontics and implant dentistry. Dental Journal 13, 1 (1 2025), 13. doi:10.3390/dj13010013

[8] Galvin Sim Siang Lin, Jia Yee Foo, Shu Meng Goh, and Mohammad Khursheed Alam. 2024. Exploring the ethical dimensions of artificial intelligence and robotics in dental education. Bangladesh Journal of Medical Science 23, 4 (10 2024), 999–1007. doi:10.3329/bjms.v23i4.76509

[9] Lukas Weidener and Michael Fischer. 2024. Artificial intelligence in medicine: cross-sectional study among medical students on application, education, and ethical aspects. JMIR Medical Education 10, (1 2024), e51247. doi:10.2196/51247

[10] N.F. Nordblom, M. Büttner, and F. Schwendicke. 2024. Artificial intelligence in orthodontics: critical review. Journal of Dental Research 103, 6 (4 2024), 573-581. doi:10.1177/00220345241235606

[11] Natalia Kazimierczak, Wojciech Kazimierczak, Zbigniew Serafin, Paweł Nowicki, Jakub Nożewski, and Joanna Janiszewska-Olszowska. 2024. AI in orthodontics: revolutionizing diagnostics and treatment planning—a comprehensive review. Journal of Clinical Medicine 13, 2 (1 2024), 344. doi:10.3390/jcm13020344

[12] Arun James Thirunavukarasu, Darren Shu Jeng Ting, Kabilan Elangovan, Laura Gutierrez, Ting Fang Tan, and Daniel Shu Wei Ting. 2023. Large language models in medicine. Nature Medicine 29, 8 (8 2023), 1930–1940. doi:10.1038/s41591-023-02448-8

[13] Shuroug A. Alowais, Sahar S. Alghamdi, Nada Alsuhebany, Tariq Alqahtani, Abdulrahman I. Alshaya, Sattam N. Almutairi, Sahar H. Al-Nojaidi, Raghad S. Al-Shehri, Sarah N. Al-Ghamdi, Ahmed S. Al-Haji, Ahmed S. Al-Harbi, Sarah S. Al-Habsi, and Sarah M. Al-Garni. 2023. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Medical Education 23, 1 (9 2023), 689. doi:10.1186/s12909-023-04698-z

[14] Pranav Rajpurkar, Emma Chen, Oishi Banerjee, and Eric J. Topol. 2022. AI in health and medicine. Nature Medicine 28, 1 (1 2022), 31–38. doi:10.1038/s41591-021-01614-0

[15] Faiza Ali, Fizzah Ikram, Zahra Masood, Dinaz Ghandhi, and Rashna Hoshang Sukhia. 2026. Knowledge and awareness of dental residents regarding the applications of artificial intelligence for diagnosis and treatment planning- a cross-sectional survey. BMC Oral Health 26, 1 (1 2026), 143. doi:10.1186/s12903-026-07712-0

[16] Esmaeil Mehraeen, Niloofar Attarian, Arghavan Afra, Setareh Parsakian, SeyedAhmad SeyedAlinaghi, and Omid Dadras. 2026. The current state of artificial intelligence in healthcare: a narrative review of opportunities and challenges. Shiraz E Medical Journal 27, 2 (2 2026), e161280. doi:10.5812/semj-161280

[17] Merve Hayriye Kocaoğlu, Akif Demirel, and İlhan Kaya. 2026. Accuracy, quality, and readability analyses of responses from large language models to questions on pediatric dental sedation. BMC Oral Health 26, 1 (2 2026), 183. doi:10.1186/s12903-026-08026-x

[18] Kevin B. Johnson, Wei-Qi Wei, Dilhan Weeraratne, Mark E. Frisse, Karl Misulis, Kyu Rhee, Jing Zhao, and James L. Snow. 2021. Precision medicine, AI, and the future of personalized health care. Clinical and Translational Science 14, 1 (1 2021), 86–93. doi:10.1111/cts.12884

[19] Peter Lee, Sebastien Bubeck, and Joseph Petro. 2023. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. New England Journal of Medicine 388, 13 (3 2023), 1233–1239. doi:10.1056/NEJMsr2214184

[20] Tirth Dave, Sai Anirudh Athaluri, and Satyam Singh. 2023. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. Frontiers in Artificial Intelligence 6, (5 2023), 1169595. doi:10.3389/frai.2023.1169595

[21] Charlotte J. Haug and Jeffrey M. Drazen. 2023. Artificial intelligence and machine learning in clinical medicine, 2023. New England Journal of Medicine 388, 13 (3 2023), 1201–1208. doi:10.1056/NEJMra2302038

[22] F. Schwendicke, W. Samek, and J. Krois. 2020. Artificial intelligence in dentistry: chances and challenges. Journal of Dental Research 99, 7 (7 2020), 769–774. doi:10.1177/0022034520915714

[23] Florin Eggmann, Roland Weiger, Nicola U. Zitzmann, and Markus B. Blatz. 2023. Implications of large language models such as ChatGPT for dental medicine. Journal of Esthetic and Restorative Dentistry 35, 7 (10 2023), 1098–1102. doi:10.1111/jerd.13046

[24] Naseer Ahmed, Maria Shakoor Abbasi, Filza Zuberi, Warisha Qamar, Mohamad Syahrizal Bin Halim, Adnan Aliane, and Mohammad Khursheed Alam. 2021. Artificial intelligence techniques: analysis, application, and outcome in dentistry - a systematic review. BioMed Research International 2021, (6 2021), 9751564. doi:10.1155/2021/9751564

[25] Nighat Naved, Samira Adnan, and Fahad Umer. 2026. Development trend of artificial intelligence (AI) in dentistry: exploring FDA-cleared dental devices. BMC Oral Health 26, 1 (1 2026), 14. doi:10.1186/s12903-025-07302-6

[26] Chengchen Hu, Fang Li, Shengguo Wang, Zhi Gao, Shan Pan, and Maofeng Qing. 2025. The role of artificial intelligence in enhancing personalized learning pathways and clinical training in dental education. Journal of Oral Microbiology 17, 1 (5 2025), 2490425. doi:10.1080/2331186X.2025.2490425

[27] Warren J. von Eschenbach. 2021. Transparency and the Black Box Problem: Why We Do Not Trust AI. Philosophy & Technology 34, 4 (12 2021), 1607–1622. doi:10.1007/s13347-021-00477-0

[28] Arka Choudhury and Hanieh Shamszare. 2026. Human factors influencing trust in healthcare artificial intelligence: systematic literature review. Human Factors in Healthcare 8, (2026), 100103. doi:10.1080/24725838.2026.2625655

[29] Vanessa Tucci, Jennifer Saary, and Thomas E. Doyle. 2021. Factors influencing trust in medical artificial intelligence for healthcare professionals: a narrative review. Journal of Medical Artificial Intelligence 4, (2021), 25. doi:10.21037/jmai-21-25

[30] Megan Mertz, Kelvi Toskovich, Gavin Shields, Gregory Attema, Jennifer Dumond, and Erin Cameron. 2025. Exploring trust factors in AI-healthcare integration: a rapid review. Frontiers in Artificial Intelligence 8 (12 2025), 1658510. doi:10.3389/frai.2025.1658510

[31] Mitul Harishbhai Tilala, Pradeep Kumar Chenchala, Aasritha Choppadandi, Jaspreet Kaur, Sravya Naguri, Rahul Saoji, and Bhanu Devaguptapu. 2024. Ethical Considerations in the Use of Artificial Intelligence and Machine Learning in Health Care: A Comprehensive Review. Cureus 16, 6 (6 2024), e62443. doi:10.7759/cureus.62443

[32] Juan Manuel Durán and Karin Rolanda Jongsma. 2021. Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. Journal of Medical Ethics 47, 5 (5 2021), 329–335. doi:10.1136/medethics-2020-106820

[33] Daiju Ueda, Taichi Kakinuma, Shohei Fujita, Koji Kamagata, Yasutaka Fushimi, Rintaro Ito, Yusuke Matsui, Taiki Nozaki, Takeshi Nakaura, Noriyuki Fujima, Fuminari Tatsugami, Masahiro Yanagawa, Kenji Hirata, Akira Yamada, Takahiro Tsuboyama, Mariko Kawamura, Tomoyuki Fujioka, and Shinji Naganawa. 2024. Fairness of artificial intelligence in healthcare: review and recommendations. Japanese Journal of Radiology 42, 1 (1 2024), 3–15. doi:10.1007/s11604-023-01474-3

[34] Arka Choudhury and Osman Asan. 2020. Human factors and trust in healthcare: focus on clinicians. Journal of Medical Internet Research 22, 6 (6 2020), e15154. doi:10.2196/15154

[35] Julia Amann, Alessandro Blasimme, Effy Vayena, Dietmar Frey, and Vince I. Madai. 2020. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Medical Informatics and Decision Making 20, 1 (2020), 310. doi:10.1186/s12911-020-01332-6

[36] Erik Brynjolfsson and Tom Mitchell. 2017. What can machine learning do? Workforce implications. Science 358, 6370 (12 2017), 1530–1534. doi:10.1126/science.aap8062

[37] Nancy Sanjay Gupta, Pravir Kumar. Perspective of artificial intelligence in healthcare data management: A journey towards precision medicine. Computers in Biology and Medicine 162 (2023) 107051.

[38] David Shalom Liu, Jake Sawyer, Alexander Luna, Jihad Aoun, Janet Wang, Lord Boachie, Safwan Halabi, Bina Joe. Perceptions of US medical students on artificial intelligence in medicine: Mixed methods survey study. JMIR Medical Education 8, 4 (2022), e38325.

[39] Xiujing He, Xiaowei Liu, Fengli Zuo, Hubing Shi, Jing Jing. Artificial intelligence-based multi-omics analysis fuels cancer precision medicine. Seminars in Cancer Biology 88 (2023), 187–200.

[40] Hui Jin Toh, Hui Yun Chan, Thomas Ploug, Søren Holm, and Tamra Lysaght. 2025. Attitudes Towards the Use of Artificial Intelligence in Healthcare: A Conjoint Analysis Survey in Singapore. Asian Bioethics Review 17, 4 (11 2025), 371–390. doi:10.1007/s41649-025-00380-2

[41] Hein Minn Tun, Hanif Abdul Rahman, Lin Naing, and Owais Ahmed Malik. 2026. Determinants of AI Adoption in Healthcare: Insights From a Unified Theory of Acceptance and Use of Technology (UTAUT) Study Among Doctors and Nurses in a

Tertiary Care Hospital in North India. Cureus 18, 2 (2 2026), e456762. doi:10.7759/cureus.456762
[42] Codrina Mihaela Levai, Laura Alexandra Nussbaum, Adriana Cojocaru, Daian-Ionel Popa, Andrei Marius Tomescu, and Iulius Jugănaru. 2026. Health Literacy and Acceptance of AI/XR-Enabled Telemedicine Among Romanian Medical Students: A Cross-Sectional Survey. Healthcare 14, 5 (3 2026), 570. doi:10.3390/healthcare14050570
[43] Muhammad Mohsin Khan, Noman Shah, Nissar Shaikh, Abdulnasser Thabet, Talal Alrabayah, and Sirajeddin Belkhair. 2025. Towards secure and trusted AI in healthcare: A systematic review of emerging innovations and ethical challenges. International Journal of Medical Informatics 195, (3 2025), 105780. doi:10.1016/j.ijmedinf.2024.105780
[44] Alexander V. Eriksen, Sören Möller, and Jesper Ryg. 2023. Use of GPT-4 to Diagnose Complex Clinical Cases. NEJM AI 1, 1 (11 2023), AIp2300031. doi:10.1056/AIp2300031
[45] Eric J. Topol. 2019. High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine 25, 1 (1 2019), 44–56. doi:10.1038/s41591-018-0300-7
[46] Hein Minn Tun, Hanif Abdul Rahman, Lin Naing, and Owais Ahmed Malik. 2025. Trust in Artificial Intelligence–Based Clinical Decision Support Systems Among Health Care Workers: Systematic Review. Journal of Medical Internet Research 27 (7 2025), e69678. doi:10.2196/69678
[47] A.G.M. Zondag, J.C. Beinema, M.C. Trappenburg, and M.A. van den Hoven. 2024. The Effect of Artificial Intelligence on Patient-Physician Trust. BMC Medical Ethics 25, 1 (2024), 58. doi:10.1186/s12910-024-01141-3
[48] A.M. Astobiza, J. González-García, N. M. de Miguel Beriain, and A. del Pozo. 2024. Trust and AI in healthcare: a systematic review. AI and Ethics (11 2024), 1–18. doi:10.1007/s43681-024-00654-8.

Appendix

Appendix A: Databases

The search was mainly conducted on Google Scholar, Scopus, ResearchGate, WebofScience and PubMed. The reason behind this choice was to ensure good coverage of research and to provide access to scientifically reliable sources, which can be more appropriate for academic purposes compared to general search engines.

Appendix B: Dates

The search was conducted between March and April of 2026. Articles going back further than 2016 were excluded to provide focus on more recent development in the use of AI in the medical field.

Appendix C: Search Strings

Relevant studies were identified using various combinations of keywords related to AI, the healthcare world and more specifically the orthodonture world.
Keywords were also directly used on the search string to broaden the results.

Table 3: Search strings used per concept area

Concept	Search Strings
Medicine and AI	"Artificial intelligence" OR "medicine" OR "dentistry" OR "use of AI in medicine"
Ethical aspects of AI applied to dentistry	"AI" AND "orthodontist"
Ethical aspects of AI applied to medicine	"Trust and dentistry" OR "worker trust" OR "trust in AI" OR "parameters of trust" OR "trust calibration"
Well-balanced trust between doctors and AI	"Should doctors trust AI?"
The "Black Box problem"	"Black Box applied in Medicine"

Appendix D: Coding Scheme

For all the articles we decided to include, we then identified and extracted relevant factors useful to our research question. Some of the most determining factors were:

- Accuracy and reliability
- Uncertainty and unpredictability
- Task risk and responsibility
- Stakeholders affected
- Common need of validation through human oversight
- Data quality and lack of biases
- Actual usefulness of AI in the medical field

Appendix E: Use Of AI Tools

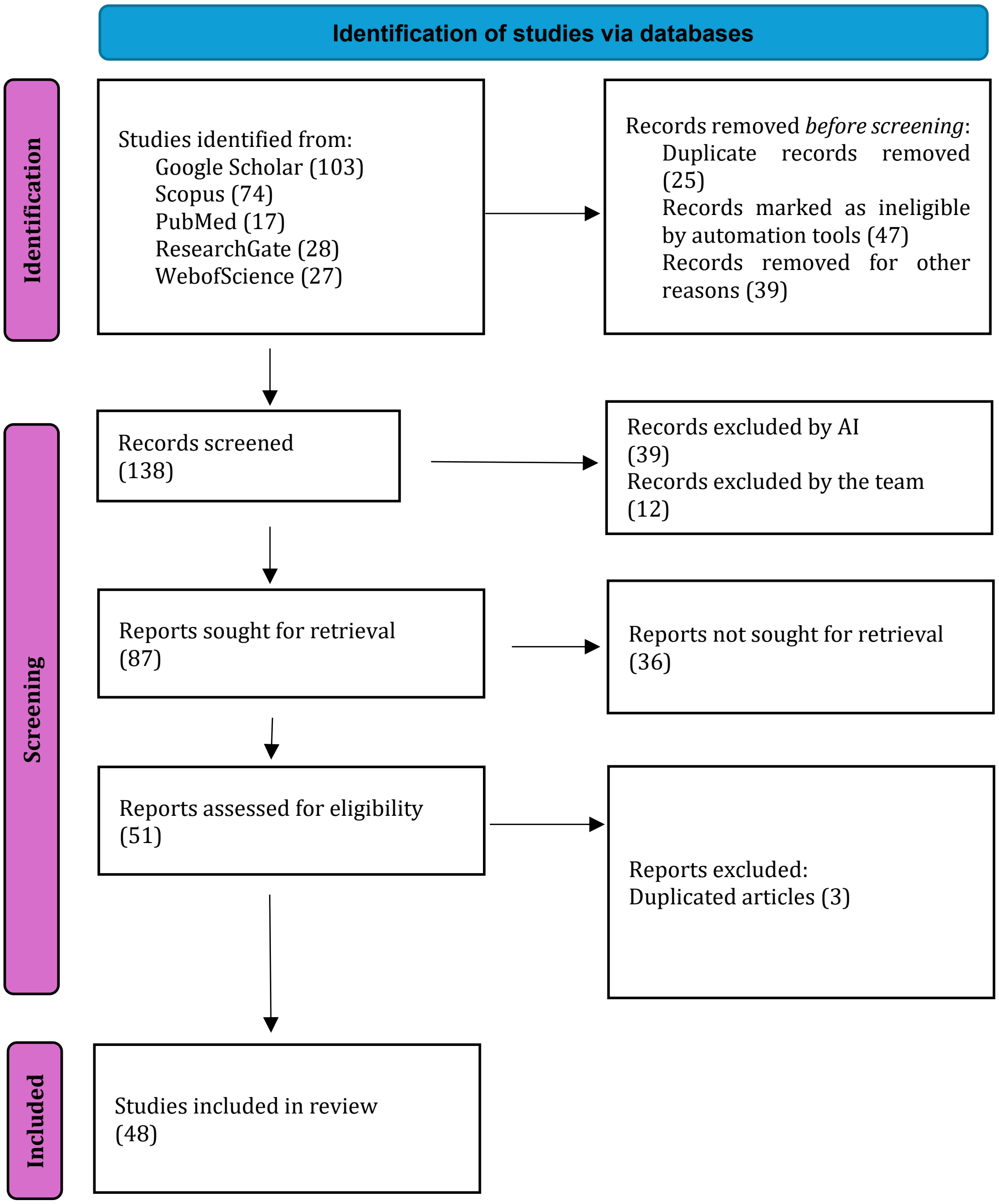
AI tools such as ChatGPT, Gemini and Copilot were used in the coding and screening processes to support the identification of key themes and summarization of articles. These tools were also used as a general support to verify completeness and righteousness of the appendix, to ensure it followed the general scoping review guidelines. It was also used during the review's correction following the peer reviews, to understand how to properly insert references in the paper.

Appendix F: Screening Process

The screening process was conducted by firstly screening titles, then screening abstracts and lastly full-text review. Groups of team members reviewed a subset of articles evaluating the relevance to our specific research question: "*What factors determine whether a worker trusts AI to perform a task, and when is that trust well calibrated?*". Some articles per partially included/dismissed when only specific paragraphs where helpful in giving generalizable insights.

Appendix G: PRISMA-ScR flow diagram

Chart 1: flow diagram regarding the screening process



1) **First elimination:** by entering the search strings and the keywords on the search engines we were able to identify a total of 249 articles based on title-screening process. A large sum of them were immediately dismissed as they did not meet the inclusion/exclusion criterion previously described, resulting in a remaining pool of 138 articles.
2) **Second elimination:** with the help of AI tools we skimmed through the abstract, leaving a total of 87 articles for further analysis.
3) **Final screening:** we finally skimmed through the entire texts (with the help of AI tools); this screening process resulted in the selection of 51 articles.
4) **Final review:** a full text review was then conducted, and we were able to retain 48 articles in our scoping review.

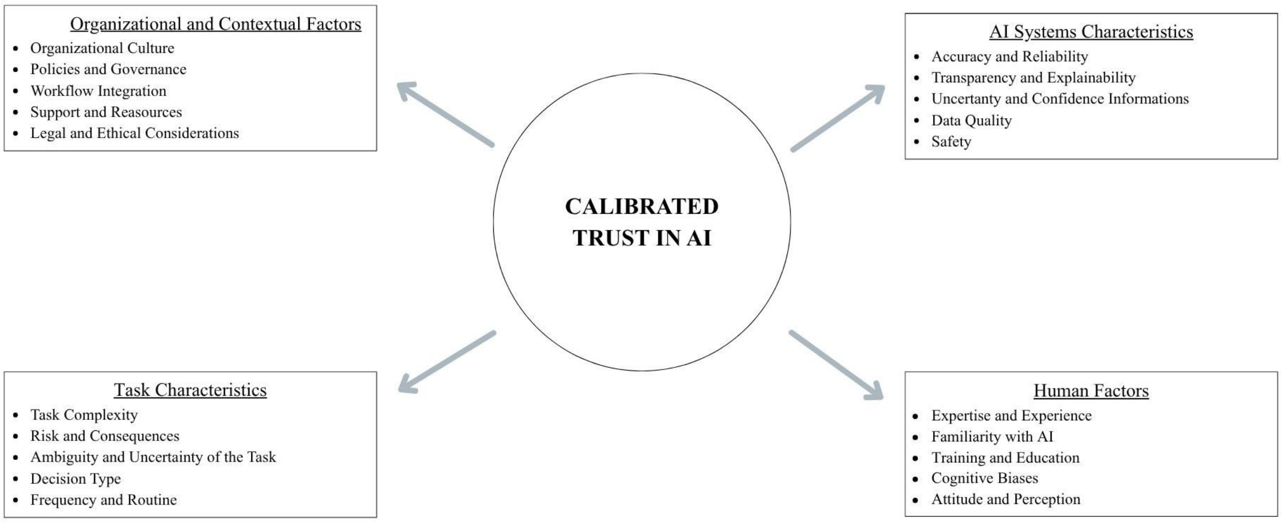
Appendix H: Coding scheme

Table 4: Coding scheme of the 48 included sources, organized by Source, Factor, Code, Theme and Evidence Type

Source	Factor	Code	Theme	Evidence Type
[1]	Development of AI in medicine	Evolution of AI	AI systems	Review
[2]	Ethical implications of AI in medicine	Ethical concerns	Organizational	Theoretical
[3]	AI in medical writing	Reliability concerns	AI systems	Review
[4]	Ethical implications in medical education	Ethical concerns	Organizational	Review
[5]	Ethical issues in healthcare AI	Ethical concerns	Organizational	Review
[6]	AI in dentistry practice	Integration barriers	Organizational	Empirical
[7]	AI performance vs limitations	Performance limitations	AI systems	Review
[8]	Ethics in dental education	Ethical concerns	Human factors	Review
[9]	AI awareness in students	User knowledge	Human factors	Empirical
[10]	AI in orthodontics	Clinical applications	Task characteristics	Empirical
[11]	AI in orthodontics diagnostics	Diagnostic support	Task characteristics	Empirical
[12]	LLMs in medicine	AI hallucinations	AI systems	Review
[13]	AI in clinical practice	Clinical integration	Organizational	Empirical
[14]	AI in health and medicine	Reliability concerns	AI systems	Review
[15]	Knowledge of AI tools	User knowledge	Human factors	Empirical
[16]	Opportunities and challenges of AI	Performance limitations	AI systems	Review
[17]	LLM responses in dentistry	Reliability concerns	AI systems	Empirical
[18]	Precision medicine and AI	System performance	AI systems	Review
[19]	GPT-4 in medicine	Use and risks	AI systems	Review
[20]	ChatGPT in medicine	Use and risks	AI systems	Review
[21]	AI in clinical medicine	Diagnostic support	AI systems	Review
[22]	AI in dentistry	Integration barriers	Organizational	Review
[23]	LLMs in dentistry	Human-AI interaction	Human factors	Review
[24]	AI in dentistry	Clinical usefulness	Task characteristics	Empirical
[25]	AI device development	Evolution of AI	AI systems	Practice
[26]	AI in education and training	User knowledge	Human factors	Review
[27]	Transparency in AI	Black box problem	AI systems	Theoretical
[28]	Human factors influencing trust in healthcare AI	Trust determinants	Human factors	Review
[29]	Factors influencing trust in medical AI	Explainability and trust	AI systems	Review
[30]	Trust factors in AI-healthcare integration	Trust determinants	Human factors	Review
[31]	Ethical considerations in AI/ML in healthcare	Human validation needed	AI systems	Review
[32]	Trust in black box algorithms	Black box problem	AI systems	Theoretical
[33]	Fairness of AI in healthcare	Biases and fairness	Organizational	Review
[34]	Human factors and clinician trust	Job displacement concerns	Human factors	Theoretical
[35]	Explainability for AI in healthcare	Explainability	AI systems	Review
[36]	Workforce implications of ML	Job displacement concerns	Organizational	Theoretical
[37]	AI in healthcare data management	Precision medicine	AI systems	Review
[38]	Medical students' perceptions of AI	User perception	Human factors	Empirical
[39]	AI-based multi-omics in cancer	Future clinical applications	Task characteristics	Review
[40]	Attitudes toward AI in healthcare	User perception	Human factors	Empirical
[41]	Determinants of AI adoption in healthcare	Adoption factors	Human factors	Empirical
[42]	AI/XR-enabled telemedicine acceptance	Clinical usefulness	Task characteristics	Empirical
[43]	Secure and trusted AI in healthcare	Human-AI collaboration	Organizational	Review
[44]	GPT-4 diagnosing clinical cases	Task-specific trust	Task characteristics	Empirical
[45]	High-performance medicine	Future clinical applications	Task characteristics	Review
[46]	Trust in AI-based CDSS among health care workers	Trust calibration	Human factors	Review
[47]	Effect of AI on patient-physician trust	Workflow integration	Organizational	Empirical
[48]	Trust and AI in healthcare	Trust determinants	Human factors	Review

Appendix I: Framework diagram

Chart 2: Diagram on the framework about factors influencing Trust in AI and Trust Calibration in Healthcare



Appendix J: ZIP Archive

Link to the ZIP archive:

https://drive.google.com/file/d/1lF8m7KO_Q5JixzCrFrpSfKkdqzBLX3Gl/view?usp=sharing

Appendix K: Dogtooth Team members

Samin Nour Azaran; s324057@studenti.polito.it
Francesco Sergio Blandino; s320524@studenti.polito.it
Giorgia Brusa; s341200@studenti.polito.it
Elisa Ciani; s340642@studenti.polito.it
Elena Corio; s342703@studenti.polito.it
Bruno Yael Zambrano Rios; s337077@studenti.polito.it
Alessandro Russo; s328917@studenti.polito.it
Sarvar Toshmurodov; s334729@studenti.polito.it
Sonia Visco; s341763@studenti.polito.it
Erika Zhou; s335819@studenti.polito.it