Machine learning for prognosis of oral cancer: What are the ethical challenges?

Long paper

Rasheed Omobolaji Alabi[1] [0000-0001-7655-5924], Tero Vartiainen[2] [0000-0003-3843-8561], Mohammed Elmusrati[1] [0000-0001-9304-6590]

[1] Department of Industrial Digitalization, School of Technology and Innovations, University of Vaasa, Vaasa, Finland
[2] Department of Computer Science, School of Technology and Innovations, University of Vaasa, Vaasa, Finland
rasheed.alabi@student.uwasa.fi

Abstract.
Background: Machine learning models have shown high performance, particularly in the diagnosis and prognosis of oral cancer. However, in actual everyday clinical practice, the diagnosis and prognosis using these models remain limited. This is due to the fact that these models have raised several ethical and morally laden dilemmas. **Purpose:** This study aims to provide a systematic state-of-the-art review of the ethical and social implications of machine learning models in oral cancer management. **Methods:** We searched the OvidMedline, PubMed, Scopus, Web of Science and Institute of Electrical and Electronics Engineers databases for articles examining the ethical issues of machine learning or artificial intelligence in medicine, healthcare or care providers. The Preferred Reporting Items for Systematic Review and Meta-Analysis was used in the searching and screening processes. **Findings:** A total of 33 studies examined the ethical challenges of machine learning models or artificial intelligence in medicine, healthcare or diagnostic analytics. Some ethical concerns were data privacy and confidentiality, peer disagreement (contradictory diagnostic or prognostic opinion between the model and the clinician), patient's liberty to decide the type of treatment to follow may be violated, patients–clinicians' relationship may change and the need for ethical and legal frameworks. **Conclusion:** Government, ethicists, clinicians, legal experts, patients' representatives, data scientists and machine learning experts need to be involved in the development of internationally standardised and structured ethical review guidelines for the machine learning model to be beneficial in daily clinical practice.

Keywords: Ethics, machine learning, oral tongue cancer, systematic review

1 Introduction

Cancer is the second leading cause of death, with an estimated 9.6 million deaths worldwide in 2018 (Bray et al., 2018). From this estimation, oral cancer accounts for 354,864 new cases and 177,384 deaths (Bray et al., 2018), making it one of the most common cancers and thus a source of significant health concern. Notably, oral squamous cell carcinoma is the most frequent of all cases of oral cancer (Ng et al., 2017). It represents about 90% of all the reported cases of oral cancer (Le Campion et al., 2017; Neville et al., 2009). Oral tongue cancer has been reported to have a worse prognosis than squamous cell carcinoma arising from other subsites of the oral cavity (Rusthoven et al., 2008). Therefore, an accurate tool for the effective prognostication of oral cancer is necessary.

Artificial intelligence (AI), or its subfield machine learning (ML), holds great promise in effective oral cancer diagnosis and prognosis (Amato et al., 2013), clinical decision making (Bennett & Hauser, 2013; Esteva et al., 2019; Topol, 2019) and personalised medicine (Dilsizian & Siegel, 2014) because of the improved availability of large datasets (big data), increased computational power and advances in ML training algorithms. In the era of unprecedented technological advancements, AI or ML is recognised as one of the most important application areas. It is currently positioned at the apex of the hype curve and is touted to facilitate improved diagnostics, prognostics, workflow and treatment planning and monitoring of oral cancer patients.

Several studies have been published emphasising the importance of ML techniques in prediction outcomes, such as recurrence (Alabi, Elmusrati, Sawazaki-Calone, et al., 2019; Alabi, Elmusrati, Sawazaki-Calone, et al., 2019), occult node metastasis (Bur et al., 2019) or five-year overall survival in oral cancer patients (Karadaghy et al., 2019). Despite the reported high accuracy in the application of ML techniques in head and neck cancer studies, there is also some trepidation among clinicians regarding its uncertain effect on the demand and training of the current and future workforce. Some clinicians have considered the introduction of ML to daily routine medical practice as a transformative improvement in the ability to diagnose the disease early enough and more accurately, and others have expressed concerns about the assessment of and consensus on possible ethical pitfalls. Interestingly, this is usually the case with most disruptive technologies.

The adoption of AI technology in actual daily medical practice has been argued to threaten patients' preference, safety and privacy (Michael, 2019). Considering the progress made by AI technology and ML-based models in cancer management, the current policy and ethical guidelines are lagging (Michael, 2019). Although there are some efforts to engage in these ethical discussions (Luxton, 2014, 2016; Peek et al., 2015), the medical community needs to be informed about the complexities surrounding the application of AI technology and ML-based models in actual clinical practice (Michael, 2019).

Studies have examined the ethical challenges in the implementation of AI, or its subfield ML, in healthcare or medicine. As this approach seems general, few published works have focused on the ethical challenges in AI or ML in oral cancer. Therefore, our study aims to systematically review the research on the ethics of AI in medicine. This study mainly focuses on ML models. These ethical dilemmas are

adapted to when these ML models are used in oral cancer management. To this end, this systematic review addresses the following research questions (RQ):

RQ. What are the ethical challenges in the integration of the ML model into the daily clinical practice of oral cancer management?

RQ. What are the generic approaches to addressing these ethical challenges?

This paper is organised as follows. Section 2 describes the methodology. Section 3 examines the results obtained from the systematic review. Section 4 discusses the results and the implications for daily clinical practices.

2 Materials and methods

2.1 Search protocol

In this study, we systematically retrieved all studies that examined ethics in ML or AI. The systematic search included the databases of OvidMedline, PubMed, Scopus, Institute of Electrical and Electronics Engineers, Web of Science and Cochrane Library from their inception until 17 March 2020. The search approach was developed by combining the following search keywords: [('machine learning OR artificial intelligence') AND ('ethics')]. The retrieved hits were further analysed for possible duplicates and irrelevant studies. To further minimise the omission of any study, the reference lists of all eligible articles were manually searched to ensure that all the relevant studies were duly included. In addition, the Preferred Reporting Items for Systematic Review and Meta-Analysis was used in the searching and screening processes (Figure 1).

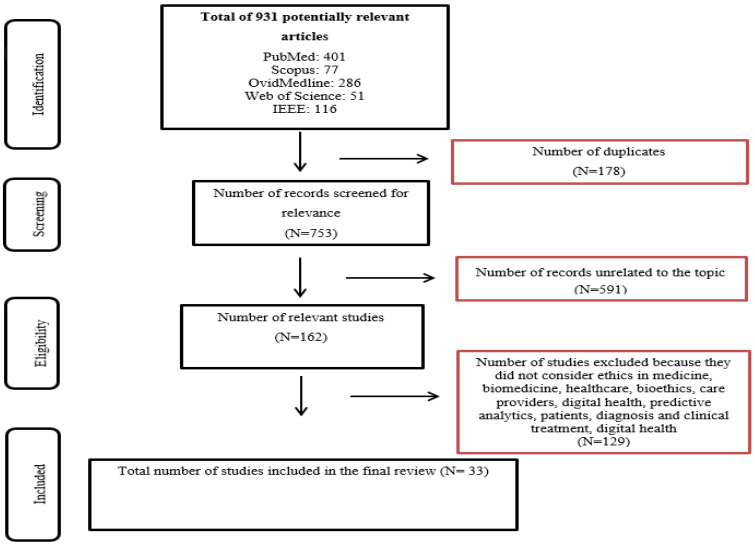

Figure 1. The number of articles included that examined the ethical concerns of ML models in medicine.

2.2 Inclusion and exclusion criteria

All original articles that considered the ethics of ML or AI in medicine or healthcare were included in this study. The eligible studies must have evaluated the ethical considerations or concerns of ML or AI in medicine. Studies that examined privacy issues, ethics of data practice and stewardship were also deemed eligible. Owing to the nature of the research questions in this study, perspectives, editorials and reviews were included. However, studies on animals, abstracts and conference papers were omitted. Articles in languages other than English were also excluded (Figure 1).

2.3 Screening

A data extraction sheet was used to minimise errors due to the omission of eligible studies.

2.4 Data extraction

The extracted parameters from each study included the author's/authors' name, year of publication, country of authors, title of studies and summary of the ethical issues mentioned in the study (Supplementary Table 1). Other important parameters, such as how to address such ethical challenges, were noted and discussed collectively in the discussion section.

3 Results

3.1 Results of the search strategy

The flow chart (Figure 1) describes the study selection process. A total of 931 hits were retrieved. Among them, 178 studies were found to be duplicate studies, and 591 were found to be irrelevant to the research questions in this review. Additionally, 129 studies did not consider ethics in medicine, biomedicine, healthcare, predictive analytics, digital health or patients. Thus, they were all excluded. Overall, 33 studies were found eligible for this systematic review (Figure 1, Supplementary Table 1). The findings of these studies indicated the ethical consideration of AI or ML in medicine. They were examined on how they relate to the implementation of ML models in oral cancer management. The ethical concerns discussed in these studies were privacy and confidentiality of patients' data, bias in the data used to develop the model, peer disagreement (Grote & Berens, 2020), responsibility or accountability gap (Grote & Berens, 2020; Jaremko et al., 2019; Kwiatkowski, 2018), fiduciary relationship between physicians and patients may change (Char et al., 2018; Nabi, 2018; Reddy et al., 2020) and patients' autonomy may be violated (Arambula & Bur, 2020; Boers et al., 2020; Grote & Berens, 2020; Johnson, 2019). These ethical concerns, brief definitions and corresponding structural aspects (what and how to address these concerns) are presented in Table 1.

Table 1. Ethical concerns of ML models in oral cancer prognostication.

Ethical concerns	**Meaning**	**The structural aspect of the ethical concerns** Ethical and moral concerns	
Privacy and confidentiality of patients' data	Approval of patients' consent and the concerned authority to use patients' data	*Concern 1:* Will the ML model developer use the extracted patients' information from the hospital registry without their consent?	*Concern VII*: How can the developer seek informed consent from the patient, hospital authority and national agency?
Bias in the data used to develop the model	Data may tend towards a particular race, geographical location, sexual orientation and so on	*Concern II:* Will the developed ML model be biased due to the imbalance in the data?	*Concern VIII*: How can the developer handle the possible data imbalance in the developed ML model?
Peer disagreement	Contradictory diagnostic or prognostic opinion between the model and the clinician	*Concern III:* Will the clinician follow his/her own diagnostic decision in cases in which the ML model gives a contrary opinion?	*Concern IX:* How can I find balance between conflicting diagnostic opinions? Is there an ethical guideline or standard that guides the use of a ML model in cancer management?
Responsibility gap	Assignment of responsibility when the ML models gave a wrong prediction	*Concern IV:* Will the clinician be held responsible when the ML model gives a wrong prediction?	*Concern X:* How should the clinicians interpret the hospital guidelines on the use of ML models? What does medical ethics stipulate? What are ethical guidelines or standards that guide the use of ML models in cancer management?
Clinician–patient relationship	Fiduciary interaction between the physicians and patients may change	*Concern V:* Will the patient feel comfortable and confident about the diagnostic decision made by a machine/computer?	*Concern X1:* How will I explain to the patient that the ML model is capable of making an accurate decision? How can I further justify the decision made by the model? How can I uphold clinician–patient relationship?
Patients' autonomy	Ability of the patient to determine the best treatment and take part in a shared decision-making process	*Concern VI:* Will the patient be allowed to choose the treatment approach that suits him/her when the model gives a different treatment plan?	*Concern XII:* How can the clinician take into consideration the treatment plan that best considers the daily activities of the patient?

The title of each concern (Table 1) addresses the core ethical challenge: in the case of ethical and moral concerns, 'Will the clinician, ML developer or the corresponding model perform the unethical action?' and in the case of morally acceptable actions, 'How can the clinician, ML developer or the corresponding model resolve the ethical concerns'? From these findings, it is important for the ML model to be trustworthy before it can be considered in actual medical practice. To ensure the trustworthiness of the model, the five trustworthiness principles of transparency, credibility, auditability,

reliability and recoverability should be incorporated (Figure 2) (Keskinbora, 2019; Rossi, 2016). Moreover, an ethics board has been proposed to discuss ethics in ML models from the perspective of experts and patients (Mamzer et al., 2017) (Figure 3).

3.2 Characteristics of the study

In terms of language, all the studies included were conducted in English. Out of the 33 included studies, 16 (48.5%) emphasised the privacy and confidentiality of patients' data (Bali et al., 2019a; Balthazar et al., 2018; Boers et al., 2020; Geis et al., 2019; Grote & Berens, 2020; Jaremko et al., 2019; Kluge, 1999; Kohli & Geis, 2018; Ma et al., 2019; Nabi, 2018; Nebeker et al., 2019; Reddy et al., 2020; Seddon, 1996; Sethi & Theodos, 2009; Vayena et al., 2018; Yuste et al., 2017), 13 (39.4%) examined the significance of informed consent, data protection, access, usability, sharing and regulatory schemes or rules prior to the use of patients' data (Balthazar et al., 2018; Gruson et al., 2019; Jaremko et al., 2019; Kluge, 1999; Kohli & Geis, 2018, 2018; Ma et al., 2019; Nabi, 2018; Nebeker et al., 2019; Reddy et al., 2020; Sethi & Theodos, 2009; Vayena et al., 2018; Yuste et al., 2017), 12 (36.4%) discussed the possibility bias in the data used for ML applications (Boers et al., 2020; Cahan et al., 2019; Char et al., 2018; Geis et al., 2019; Grote & Berens, 2020; Gruson et al., 2019; Kohli & Geis, 2018; Nabi, 2018; Reddy et al., 2020; Vayena et al., 2018; Wiens et al., 2019; Yuste et al., 2017), 4 (12.1%) suggested that the integration of ML models in clinical settings could assist clinicians to make informed decisions (Berner, 2002; Boers et al., 2020; Grote & Berens, 2020; Kwiatkowski, 2018) and 13 (39.4%) reported the need for ethical principles, guidelines and legal frameworks before ML models could be integrated into medical practice (Arambula & Bur, 2020; Cahan et al., 2019; Char et al., 2018; Gruson et al., 2019; Jian, 2019; Johnson, 2019; Keskinbora, 2019; Mamzer et al., 2017; Morley & Floridi, 2020; Nebeker et al., 2019; Rajkomar et al., 2018; Reddy et al., 2020; Robles Carrillo, 2020).

4 Discussion

This systematic review examined the ethical challenges in ML models in clinical practice. These challenges were examined on how they relate to the integration of ML models in oral cancer management. These ethical challenges carry significant implications in terms of integrating the ML model for daily routine in oral cancer management. The following highlights these ethical challenges and suggests a generic approach to addressing them.

Data privacy and confidentiality: the patient's consent should be sought

The first of these ethical concerns is healthcare data privacy (Arambula & Bur, 2020; Nabi, 2018). Developing ML models involves the substantial usage of healthcare data of patients. Therefore, it raises privacy and patient confidentiality concerns (Ma et al., 2019; Nabi, 2018). To arrest this concern, the patients, or their respective subjects, need to be informed about the collection and usage of their data (Geis et al., 2019; Powles & Hodson, 2017) to ensure informed consent and avoid illegal proprietary exploitation of the data and data privacy breaches (Bali et al., 2019b; Balthazar et al.,

2018; Char et al., 2018; Nabi, 2018; Powles & Hodson, 2017; Yuste et al., 2017). Nevertheless, it is important that data use agreements should be reviewed and approved by the appropriate quarters (Kohli & Geis, 2018). Moreover, a scheme (i.e., privacy-preserving clinical decision with cloud support) that preserves the privacy of the patient in terms of their data can be introduced (Geis et al., 2019; Liu et al., 2017; Ma et al., 2019; Vayena et al., 2018; Wang et al., 2015; Zhang et al., 2018). However, the discussion about the ownership of the data is beyond the scope of this study.

Trustworthy AI: the model should be trustworthy

It is important for the model to work as expected. Therefore, the model should have minimal errors in the training phase. Any form of error/malfunctioning of the model should be mentioned and defined (England & Cheng, 2019; Park & Kressel, 2018; Vayena et al., 2018; Zou & Schiebinger, 2018) to give transparency to the model and consequently, the results from these models (Geis et al., 2019; Park et al., 2019). Therefore, a possible imbalance in the data should be considered when developing the model to ensure the trustworthiness of the model. To address this challenge, related guidelines can be followed for transparent reporting (Bossuyt et al., 2015; Collins et al., 2015; England & Cheng, 2019). With these guidelines, the ML model deployed will be trustworthy and uphold the fundamental pillars of medical ethics (autonomy, beneficence, nonmaleficence and justice) (Arambula & Bur, 2020) and the ethical principles of transparency, credibility, audibility, reliability and recoverability (Keskinbora, 2019) (Figure 2).

Figure 2. The trustworthiness principles expected from a ML model

In this way, an inherently biased model is avoided (Arambula & Bur, 2020; Collins & Moon, 2018; Reddy et al., 2020; Wiens et al., 2019). Trustworthiness should not only concern the properties of the ML or AI inherent model but also the socio-technical systems involving the ML or AI applications (European Commission, 2019), that is, the expected trustworthiness of all actors and processes that constitute the socio-technical context in the application of AI for the prognostication of oral tongue cancer. Thus, for trustworthiness in AI, the essential components of trust in design, development, law compliance, ethics and robustness must be present (European Commission, 2019). In addition, the key requirements for a trustworthy AI include human regulatory agency, technical robustness and safety, privacy and data governance, transparency, non-discrimination and fairness, environmental friendliness and compliance, and accountability (European Commission, 2019).

Peer disagreement: the model and clinician should act to protect the patient from harm

As the ML model is viewed as an expert system/model, peer disagreement and its possible resolution guidelines are another important ethical issue (Christensen, 2007; Kelly T, 2010). What happens when the model and the clinicians disagree on the output of a proposition (diagnosis or prognosis) (Frances & Matheson, 2018)? It is impossible to have a dialogical engagement with the model, as proposed by Mercier and Sperber in the argumentative theory of reasoning (Mercier & Sperber, 2017). Should the clinician follow the proposition of the ML model (Christensen, 2007) or adhere to her own proposition (Enoch, 2010)? Therefore, there is a standoff in terms of the possible decision to make by the clinician. In this case, ethical guidelines and legal frameworks become imperative (Figure 3).

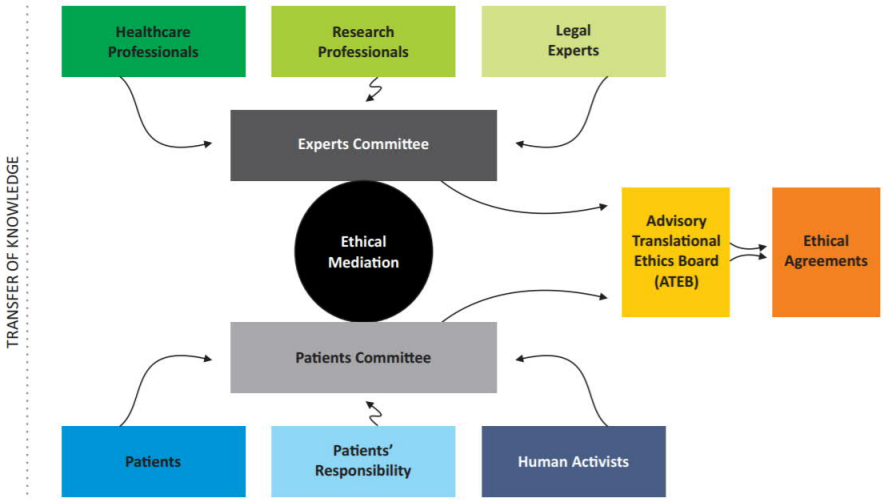

Figure 3. Ethical and legal frameworks for ethical agreements.

The ethical guidelines in this case ensure that clinicians make a decision to protect the safety and improve the overall health condition of the patient. The hospital and ethical

guidelines should also address the possible errors that may arise from using the model (responsibility gap).

Patients' autonomy: shared decision making

The ethical question of patients' autonomy also comes to fore (Grote & Berens, 2020). For example, an ML model that predicts the type of treatment for an oral cancer patient should eschew the preferred treatment that could minimise the suffering of the patient. Instead, it should maximise the lifespan and overall survival of the patient, thereby making this model paternalistic in nature. This raises the ethical question of a shared decision making between the clinician and the patient to ensure that the autonomy and dignity of the patient are not violated (McDougall, 2019). Therefore, it is important to establish relevant standards to determine which information from the ML model is essential to be explained to the patient to be regarded as informed consent so that the patient can make an informed decision (Grote & Berens, 2020; McDougall, 2019; Mittelstadt & Floridi, 2016).

Humanness: Empathy and trust from the clinician–patient relationship

Another ethical concern is the 'humanness' of clinicians and the role of cognitive empathy, trust, responsibility and confidentiality among clinicians (Boers et al., 2020). This seems to be a source of concern, as the integration of ML models in oral cancer management may lead to a paradigm shift from the current face-to-face or direct interaction between patients and clinicians (two-way diagnostic procedure) to a triangular relationship of patients–models–clinicians (three-way diagnostic procedure). This concern becomes pronounced especially when the models are publicly available, as the patients may engage in self-medication and self-management. Thus, the fundamental aspects of patients' care may be undermined (Boers et al., 2020). To mitigate this, these models should be integrated in such a way that restricts patients' access. In this way, the patient–clinician relationship can still be maintained, as this type of relationship has been reported to influence how patients respond to their illnesses and treatments (Kelley et al., 2014).

Ethics is one of the essential components to achieve a trustworthy AI. It is important to have a model that ensures compliance to ethical norms and principles, including fundamental human rights, moral entitlements and acceptable moral values (European Commission, 2019). As mentioned previously, some of these principles include respect for human autonomy, prevention of harm, fairness and explicability (European Commission, 2019). To this end, we tend to agree with the suggestion of setting up a dedicated ethical research agenda (Boers et al., 2020). This ethical research agenda is expected to form the required premise for the development of internationally standardised and structured ethical review guidelines (Arambula & Bur, 2020; Gruson et al., 2019; Johnson, 2019, 2019). These guidelines should emphasise the fundamental ethical rules of honesty, truthfulness, transparency, benevolence, non-malevolence and respect for autonomy (Keskinbora, 2019) and address other criticisms surrounding the application of ML-based models in actual clinical practice (Figure 4).

Figure 4. The fundamental ethical principles expected from the clinician and the ML model.

Aside from these ethical guidelines, corresponding laws (internal framework and international sphere) should be enacted by the government to ensure the legal (e.g., the European General Data Protection Regulations) (Flaumenhaft & Ben-Assuli, 2018; Vayena et al., 2018) and jurisdictional mechanisms for their enforcement (Robles Carrillo, 2020).

In conclusion, the development of ML models should take the ethical and legal framework into consideration from the data collection to the ML process and to the integration into clinical practice. A strong and proactive role is expected from the government, clinical experts, patients' representatives, data scientists, ML experts and legal and human rights activists in defining these ethical guidelines. Through this, ML models can achieve the touted benefits of optimising health systems and decision support for professionals and improve the overall health of patients. As oral tongue cancer was considered in this study, the ethical concerns mentioned and the proposed solution are peculiar to other cancer types.

References

Alabi, R. O., Elmusrati, M., Sawazaki-Calone, I., Kowalski, L. P., Haglund, C., Coletta, R. D., Mäkitie, A. A., Salo, T., Almangush, A., & Leivo, I. (2019). Comparison of supervised machine learning classification techniques in prediction of locoregional recurrences in early oral tongue cancer. *International Journal of Medical Informatics*, 104068. https://doi.org/10.1016/j.ijmedinf.2019.104068

Alabi, R. O., Elmusrati, M., Sawazaki-Calone, I., Kowalski, L. P., Haglund, C., Coletta, R. D., Mäkitie, A. A., Salo, T., Leivo, I., & Almangush, A. (2019). Machine learning application for prediction of locoregional recurrences in early oral tongue cancer: a Web-based prognostic tool. *Virchows Archiv*, *475*(4), 489–497. https://doi.org/10.1007/s00428-019-02642-5

Amato, F., López, A., Peña-Méndez, E. M., Vaňhara, P., Hampl, A., & Havel, J. (2013). Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine*, *11*(2), 47–58. https://doi.org/10.2478/v10136-012-0031-x

Arambula, A. M., & Bur, A. M. (2020). Ethical Considerations in the Advent of Artificial Intelligence in Otolaryngology. *Otolaryngology--Head and Neck Surgery: Official Journal of American Academy of Otolaryngology-Head and Neck Surgery*, *162*(1), 38–39. https://doi.org/10.1177/0194599819889686

Bali, J., Garg, R., & Bali, R. T. (2019a). Artificial intelligence (AI) in healthcare and biomedical research: Why a strong computational/AI bioethics framework is required? *Indian Journal of Ophthalmology*, *67*(1), 3–6. https://doi.org/10.4103/ijo.IJO_1292_18

Bali, J., Garg, R., & Bali, R. T. (2019b). Artificial intelligence (AI) in healthcare and biomedical research: Why a strong computational/AI bioethics framework is required? *Indian Journal of Ophthalmology*, *67*(1), 3–6. https://doi.org/10.4103/ijo.IJO_1292_18

Balthazar, P., Harri, P., Prater, A., & Safdar, N. M. (2018). Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics. *Journal of the American College of Radiology*, *15*(3), 580–586. https://doi.org/10.1016/j.jacr.2017.11.035

Bennett, C. C., & Hauser, K. (2013). Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach. *Artificial Intelligence in Medicine*, *57*(1), 9–19. https://doi.org/10.1016/j.artmed.2012.12.003

Berner, E. S. (2002). Ethical and legal issues in the use of clinical decision support systems. *Journal of Healthcare Information Management: JHIM*, *16*(4), 34–37.

Boers, S. N., Jongsma, K. R., Lucivero, F., Aardoom, J., Büchner, F. L., de Vries, M., Honkoop, P., Houwink, E. J. F., Kasteleyn, M. J., Meijer, E., Pinnock, H., Teichert, M., van der Boog, P., van Luenen, S., van der Kleij, R. M. J. J., & Chavannes, N. H. (2020). SERIES: eHealth in primary care. Part 2: Exploring the ethical implications of its application in primary care practice. *European Journal of General Practice*, *26*(1), 26–32. https://doi.org/10.1080/13814788.2019.1678958

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J. G., Moher, D., Rennie, D., de Vet, H. C. W., Kressel, H. Y., Rifai, N., Golub, R. M.,

Altman, D. G., Hooft, L., Korevaar, D. A., & Cohen, J. F. (2015). STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*, h5527. https://doi.org/10.1136/bmj.h5527

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424. https://doi.org/10.3322/caac.21492

Bur, A. M., Holcomb, A., Goodwin, S., Woodroof, J., Karadaghy, O., Shnayder, Y., Kakarala, K., Brant, J., & Shew, M. (2019). Machine learning to predict occult nodal metastasis in early oral squamous cell carcinoma. *Oral Oncology*, *92*, 20–25. https://doi.org/10.1016/j.oraloncology.2019.03.011

Cahan, E. M., Hernandez-Boussard, T., Thadaney-Israni, S., & Rubin, D. L. (2019). Putting the data before the algorithm in big data addressing personalized healthcare. *Npj Digital Medicine*, *2*(1). https://doi.org/10.1038/s41746-019-0157-2

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing Machine Learning in Health Care - Addressing Ethical Challenges. *The New England Journal of Medicine*, *378*(11), 981–983. https://doi.org/10.1056/NEJMp1714229

Christensen, D. (2007). Epistemology of Disagreement: The Good News. *Philosophical Review*, *116*(2), 187–217. https://doi.org/10.1215/00318108-2006-035

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*, *350*(jan07 4), g7594–g7594. https://doi.org/10.1136/bmj.g7594

Collins, & Moon. (2018). Is digital medicine different? *The Lancet*, *392*(10142), 95. https://doi.org/10.1016/S0140-6736(18)31562-9

Dilsizian, S. E., & Siegel, E. L. (2014). Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Current Cardiology Reports*, *16*(1), 441. https://doi.org/10.1007/s11886-013-0441-8

England, J. R., & Cheng, P. M. (2019). Artificial Intelligence for Medical Image Analysis: A Guide for Authors and Reviewers. *American Journal of Roentgenology*, *212*(3), 513–519. https://doi.org/10.2214/AJR.18.20490

Enoch, D. (2010). Not Just a Truthometer: Taking Oneself Seriously (but not Too Seriously) in Cases of Peer Disagreement. *Mind*, *119*(476), 953–997. https://doi.org/10.1093/mind/fzq070

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

European Commission. (2019). *High-level expert group on artificial intelligence. In: Ethics Guidelines for Tustrworthy AI.* European Union.

Flaumenhaft, Y., & Ben-Assuli, O. (2018). Personal health records, global policy and regulation review. *Health Policy*, *122*(8), 815–826. https://doi.org/10.1016/j.healthpol.2018.05.002

Frances, B., & Matheson, J. (2018). *Disagreement. In: Zalta EN,ed. The Stanford encyclopedia of philosophy. Spring.* 8. https://plato.stanford.edu/archives/spr2018/entries/disagreement/

Geis, J. R., Brady, A. P., Wu, C. C., Spencer, J., Ranschaert, E., Jaremko, J. L., Langer, S. G., Borondy Kitts, A., Birch, J., Shields, W. F., van den Hoven van Genderen, R., Kotter, E., Wawira Gichoya, J., Cook, T. S., Morgan, M. B., Tang, A., Safdar, N. M., & Kohli, M. (2019). Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement. *Radiology*, *293*(2), 436–440. https://doi.org/10.1148/radiol.2019191586

Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, *46*(3), 205–211. https://doi.org/10.1136/medethics-2019-105586

Gruson, D., Helleputte, T., Rousseau, P., & Gruson, D. (2019). Data science, artificial intelligence, and machine learning: Opportunities for laboratory medicine and the value of positive regulation. *Clinical Biochemistry*, *69*, 1–7. https://doi.org/10.1016/j.clinbiochem.2019.04.013

Jaremko, J. L., Azar, M., Bromwich, R., Lum, A., Alicia Cheong, L. H., Gibert, M., Laviolette, F., Gray, B., Reinhold, C., Cicero, M., Chong, J., Shaw, J., Rybicki, F. J., Hurrell, C., Lee, E., Tang, A., & Canadian Association of Radiologists (CAR) Artificial Intelligence Working Group. (2019). Canadian Association of Radiologists White Paper on Ethical and Legal Issues Related to Artificial Intelligence in Radiology. *Canadian Association of Radiologists Journal = Journal l'Association Canadienne Des Radiologistes*, *70*(2), 107–118. https://doi.org/10.1016/j.carj.2019.03.001

Jian, G. (2019). Artificial Intelligence in Healthcare and Medicine: Promises, Ethical Challenges, and Governance. *Chinese Medical Sciences Journal*, *0*(0), 99. https://doi.org/10.24920/003611

Johnson, S. L. J. (2019). AI, Machine Learning, and Ethics in Health Care. *Journal of Legal Medicine*, *39*(4), 427–441. https://doi.org/10.1080/01947648.2019.1690604

Karadaghy, O. A., Shew, M., New, J., & Bur, A. M. (2019). Development and Assessment of a Machine Learning Model to Help Predict Survival Among Patients With Oral Squamous Cell Carcinoma. *JAMA Otolaryngology–Head & Neck Surgery*, *145*(12), 1115. https://doi.org/10.1001/jamaoto.2019.0981

Kelley, J. M., Kraft-Todd, G., Schapira, L., Kossowsky, J., & Riess, H. (2014). The influence of the patient-clinician relationship on healthcare outcomes: a systematic review and meta-analysis of randomized controlled trials. *PloS One*, *9*(4), e94207. https://doi.org/10.1371/journal.pone.0094207

Kelly T. (2010). *Peer disagreement and higher order evidence. In: Goldman Al, Whitcomb D, eds. Social Epistemology: essential readings. Oxford University Press.* Oxford University Press.

Keskinbora, K. H. (2019). Medical ethics considerations on artificial intelligence. *Journal of Clinical Neuroscience*, *64*, 277–282. https://doi.org/10.1016/j.jocn.2019.03.001

Kluge, E. H. (1999). Medical narratives and patient analogs: the ethical implications of electronic patient records. *Methods of Information in Medicine*, *38*(4–5), 253–259.

Kohli, M., & Geis, R. (2018). Ethics, Artificial Intelligence, and Radiology. *Journal of the American College of Radiology*, *15*(9), 1317–1319. https://doi.org/10.1016/j.jacr.2018.05.020

Kwiatkowski, W. (2018). Medicine and technology. Remarks on the notion of responsibility in the technology-assisted health care. *Medicine, Health Care and Philosophy*, *21*(2), 197–205. https://doi.org/10.1007/s11019-017-9788-8

Le Campion, A. C. O. V., Ribeiro, C. M. B., Luiz, R. R., da Silva Júnior, F. F., Barros, H. C. S., dos Santos, K. de C. B., Ferreira, S. J., Gonçalves, L. S., & Ferreira, S. M. S. (2017). Low Survival Rates of Oral and Oropharyngeal Squamous Cell Carcinoma. *International Journal of Dentistry*, *2017*, 1–7. https://doi.org/10.1155/2017/5815493

Liu, Y., Li, Y., Fu, Y., Liu, T., Liu, X., Zhang, X., Fu, J., Guan, X., Chen, T., Chen, X., & Sun, Z. (2017). Quantitative prediction of oral cancer risk in patients with oral leukoplakia. *Oncotarget*, *8*(28). https://doi.org/10.18632/oncotarget.17550

Luxton, D. D. (2014). Recommendations for the ethical use and design of artificial intelligent care providers. *Artificial Intelligence in Medicine*, *62*(1), 1–10. https://doi.org/10.1016/j.artmed.2014.06.004

Luxton, D. D. (2016). *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier. https://doi.org/10.1016/C2013-0-12824-3

Ma, H., Guo, X., Ping, Y., Wang, B., Yang, Y., Zhang, Z., & Zhou, J. (2019). PPCD: Privacy-preserving clinical decision with cloud support. *PLOS ONE*, *14*(5), e0217349. https://doi.org/10.1371/journal.pone.0217349

Mamzer, M.-F., Duchange, N., Darquy, S., Marvanne, P., Rambaud, C., Marsico, G., Cerisey, C., Scotté, F., Burgun, A., Badoual, C., Laurent-Puig, P., & Hervé, C. (2017). Partnering with patients in translational oncology research: ethical approach. *Journal of Translational Medicine*, *15*(1). https://doi.org/10.1186/s12967-017-1177-9

McDougall, R. J. (2019). Computer knows best? The need for value-flexibility in medical AI. *Journal of Medical Ethics*, *45*(3), 156–160. https://doi.org/10.1136/medethics-2018-105118

Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.

Michael, R. (2019). Ethical Dimensions of Using Artificial Intelligence in Health Care. *AMA Journal of Ethics*, *21*(2), E121-124. https://doi.org/10.1001/amajethics.2019.121

Mittelstadt, B. D., & Floridi, L. (2016). The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics*, *22*(2), 303–341. https://doi.org/10.1007/s11948-015-9652-2

Morley, J., & Floridi, L. (2020). An ethically mindful approach to AI for health care. *The Lancet*, *395*(10220), 254–255. https://doi.org/10.1016/S0140-6736(19)32975-7

Nabi, J. (2018). How Bioethics Can Shape Artificial Intelligence and Machine Learning. *The Hastings Center Report*, *48*(5), 10–13. https://doi.org/10.1002/hast.895

Nebeker, C., Torous, J., & Bartlett Ellis, R. J. (2019). Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Medicine*, *17*(1). https://doi.org/10.1186/s12916-019-1377-7

Neville, B., Damm, D. ., & Allen, C., M. (2009). *Oral and Maxillofacial Pathology* (3rd edition). Elseiver.

Ng, J. H., Iyer, N. G., Tan, M.-H., & Edgren, G. (2017). Changing epidemiology of oral squamous cell carcinoma of the tongue: A global study: Changing epidemiology of tongue cancer. *Head & Neck*, *39*(2), 297–304. https://doi.org/10.1002/hed.24589

Park, S. H., Kim, Y.-H., Lee, J. Y., Yoo, S., & Kim, C. J. (2019). Ethical challenges regarding artificial intelligence in medicine from the perspective of scientific editing and peer review. *Science Editing*, *6*(2), 91–98. https://doi.org/10.6087/kcse.164

Park, S. H., & Kressel, H. Y. (2018). Connecting Technological Innovation in Artificial Intelligence to Real-world Medical Practice through Rigorous Clinical Validation: What Peer-reviewed Medical Journals Could Do. *Journal of Korean Medical Science*, *33*(22). https://doi.org/10.3346/jkms.2018.33.e152

Peek, N., Combi, C., Marin, R., & Bellazzi, R. (2015). Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes. *Artificial Intelligence in Medicine*, *65*(1), 61–73. https://doi.org/10.1016/j.artmed.2015.07.003

Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and Technology*, *7*(4), 351–367. https://doi.org/10.1007/s12553-017-0179-1

Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G., & Chin, M. H. (2018). Ensuring Fairness in Machine Learning to Advance Health Equity. *Annals of Internal Medicine*, *169*(12), 866–872. https://doi.org/10.7326/M18-1990

Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, *27*(3), 491–497. https://doi.org/10.1093/jamia/ocz192

Robles Carrillo, M. (2020). Artificial intelligence: From ethics to law. *Telecommunications Policy*, 101937. https://doi.org/10.1016/j.telpol.2020.101937

Rossi, F. (2016). Artificial Intelligence: Potential Benefits and Ethical Considerations. *Artificial Intelligence*, 8.

Rusthoven, K., Ballonoff, A., Raben, D., & Chen, C. (2008). Poor prognosis in patients with stage I and II oral tongue squamous cell carcinoma. *Cancer*, *112*(2), 345–351. https://doi.org/10.1002/cncr.23183

Seddon, A. M. (1996). Predicting our health: ethical implications of neural networks and outcome potential predictions. *Ethics & Medicine: A Christian Perspective on Issues in Bioethics*, *12*(3), 53–54.

Sethi, P., & Theodos, K. (2009). Translational bioinformatics and healthcare informatics: computational and ethical challenges. *Perspectives in Health Information Management*, *6*, 1h.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine*, *15*(11), e1002689. https://doi.org/10.1371/journal.pmed.1002689

Wang, G., Lu, R., & Huang, C. (2015). PSLP: Privacy-preserving single-layer perceptron learning for e-Healthcare. *2015 10th International Conference on Information, Communications and Signal Processing (ICICS)*, 1–5. https://doi.org/10.1109/ICICS.2015.7459925

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P. N., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm: a roadmap for responsible machine learning for health care. *Nature Medicine*, *25*(9), 1337–1340. https://doi.org/10.1038/s41591-019-0548-6

Yuste, R., Goering, S., Arcas, B. A. Y., Bi, G., Carmena, J. M., Carter, A., Fins, J. J., Friesen, P., Gallant, J., Huggins, J. E., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., … Wolpaw, J. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*, *551*(7679), 159–163. https://doi.org/10.1038/551159a

Zhang, X., Chen, X., Wang, J., Zhan, Z., & Li, J. (2018). Verifiable privacy-preserving single-layer perceptron training scheme in cloud computing. *Soft Computing*, *22*(23), 7719–7732. https://doi.org/10.1007/s00500-018-3233-7

Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist — it's time to make it fair. *Nature*, *559*(7714), 324–326. https://doi.org/10.1038/d41586-018-05707-8

Appendix

Supplementary Table 1. Included studies and the main ethical points discussed (below)

Authors	Country	Year	Title	Ethical points/Summary
Seddon A.M	United Kingdom	1996	Predicting our health: ethical implications of neural networks and outcome potential predictions	Privacy concern
Kluge E.H	Canada	1999	Medical narratives and patient analogs: ethical implications of electronic patient records	Privacy Accessibility of the data
Bernie E.S	United States	2002	Ethical and legal issues in the use of clinical decision support systems	Informed decision Transparency
Sethi & Theodos	United States	2009	Translational bioinformatics and healthcare informatics: computational and ethical challenges	Sensitive nature of genetic data Privacy and confidentiality
Mamzer et al.	France	2017	Partnering with patients in translational oncology research: ethical approach	To establish a long-term partnership integrating patient's expectations Expert and Patient Cancer research and personalised medicine (CARPEM) develops translational research of precision medicine for cancer
Yuste et al.	United States	2017	Four ethical priorities for neuroethologies and AI	Privacy and consent Agency and identity Augmentation Biases
Balthazar et al.	United States	2017	Protecting patients' interest in the era of big data, artificial intelligence and predictive analytics	Privacy Confidentiality Data ownership Informed consent Epistemology

				Inequalities
Kwiatkowski W	Poland	2018	Medicine and technology. Remarks on the notion of responsibility in technology-assisted healthcare	The notion of responsibility
Rajkomar et al.	United States	2018	Ensuring fairness in machine learning to advance health equity	The principle of distributed justice Health equity Four medical ethics principles
Char et al.	United States	2018	Implementing machine learning in healthcare–Addressing ethical challenges	Biases Fiduciary relationship between physicians and patients Ethical guidelines Proposition of policy enactment, programming approaches, task force or a combination of these strategies
Nabi Junaid	United States	2018	How bioethics can shape artificial intelligence and machine learning	Biases Privacy of patients Informed consent Fairness, trust, equity and confidentiality Patient–clinician relationship might change
Vayena et al.	Switzerland & United States	2018	Machine learning in medicine: Addressing ethical challenges	Data protection Privacy preservation Biased dataset Fairness and transparency
Kohli & Geis	United States	2018	Ethics, artificial Intelligence and radiology	Informed consent Privacy Objectivity Data protection Ownership Bias

				Data use agreement Review of agreement Safety, transparency and value alignment
Geis et al.	Canada	2019	Ethics of artificial intelligence in radiology	Bias Informed consent Privacy Data protection Ownership Objectivity and transparency
Cahan et al.	United States	2019	Putting the data before the algorithm in big data addressing personalised healthcare	Biases in the data Handling the confluence between data and algorithm Generalisability of the model Introducing the quality standard for the dataset guidelines
Gruson et al.	Belgium	2019	Data science, artificial intelligence and machine learning: opportunities for laboratory medicine and the value of positive regulation	Biases Patient information and consent Ethical and legal frameworks AI human warranty Regulation of health data according to their level of sensitivity
Jaremko et al.	Canada	2019	Canadian Association of Radiologists white paper on ethical and legal issues related to artificial intelligence in radiology	Data value and ownership Data privacy Data sharing rules Reliability gap (liability)
Nebeker et al.	United States	2019	Building the case for actionable ethics in digital health research supported by artificial intelligence	Privacy Data management Risks and benefits Access and usability Ethical principles (respect for persons, beneficence, justice)

Wiens at al.	United States and Canada		Do not harm: a roadmap for responsible machine learning for healthcare	Choosing the right problems Developing a useful solution Biases Proper evaluation of the performance of the model Thoughtful reporting of the models' results Integration (making it to the market)
Guan Jian			Artificial intelligence in healthcare and medicine: promises, ethical challenges and governance	Role of government in the ethical auditing Stakeholders' responsibilities in ethical governance system
Nikola et al.		2019	Algorithm-aided prediction of patient preference: an ethics sneak peek	Protection of patients and clinicians (safety, validity, reproducibility, usability, reliability) Transparency and comprehensibility Quality control and monitoring of models
Ma et al.	China	2019	PPCD: Privacy-preserving clinical decision with cloud support	Data usage privacy concern scheme
Bali et al.	India	2019	Artificial intelligence in healthcare and biomedical research: Why a strong computational/AI bioethics framework is required	Data privacy Confidentiality Do not harm principle should be upheld
Mazuro-wski	United Kingdom	2019	Artificial intelligence in radiology: some ethical considerations for radiologists and algorithm developers	When it is unethical for a radiologist to oppose AI Conflicts of interests between radiologists and AI developers

Keskinb-ora	Turkey	2019	Medical ethics considerations on artificial intelligence	Trustworthy AI Important ethical principles should be embraced
Park et al.	Korea	2019	Ethical challenges in artificial intelligence in medicine from the perspective of science editing and peer review	Transparency in training, testing and validation dataset Clearly explains the data preparation processes
Johnson Sandra	United Kingdom	2019	AI, machine learning and ethics in healthcare	Ethical guidelines recommendation Vigilant to potential errors and biases Medical ethics
Arambu-la & Bur	United States	2019	Ethical considerations in the advent of artificial intelligence in otolaryngology	Four ethical principles (respect for patient's autonomy, beneficence, nonmaleficence, and justice)
Reddy et al.,	Australia	2019	A governance model for the application of AI in healthcare	AI biases Privacy Patient and clinician trust Regulatory guidelines Proposed governance for AI in healthcare Stages for monitoring and evaluating AI-enabled services
Boers et al.,	Netherlands and United Kingdom	2019	SERIES: eHealth in primary care. Part 2: Exploring the ethical implications of its application in primary care practice	Biased and discriminatory algorithm Patient's autonomy Shared decision making Data privacy, trust and confidentiality
Morley & Floridi	United Kingdom	2020	An ethical mindful approach to AI for healthcare	Internationally standardised and structured ethical review guidelines

Carrillo et al.,	Spain	2020	Artificial intelligence: From ethics to law	Distinguish between legal and ethical aspects Non-formalistic approach to law International law is identified as the principal legal framework for the regulation of AI models
Grote & Berens	Germany	2020	On the ethics of algorithmic decision making in healthcare	Peer disagreement Patients' autonomy Shared decision making Obscuration of accountability Biased and discriminatory algorithm Data privacy