

Triage 4.0: On Death Algorithms and Technological Selection. Is Today's Data- Driven Medical System Still Compatible with the Constitution?

Helbing, Dirk; Beschorner, Thomas; Frey, Bruno; Diekmann, Andres; Hagendorff, Thilo; Seele, Peter; Spiekermann-Hoff, Sarah; van den Hoven, Jeroen; Zwitter, Andrej

Published in:
Journal of European Continuing Medical Education

DOI:
[10.1080/21614083.2021.1989243](https://doi.org/10.1080/21614083.2021.1989243)

Published: 01/01/2021

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):
Helbing, D., Beschorner, T., Frey, B., Diekmann, A., Hagendorff, T., Seele, P., Spiekermann-Hoff, S., van den Hoven, J., & Zwitter, A. (2021). Triage 4.0: On Death Algorithms and Technological Selection. Is Today's Data-Driven Medical System Still Compatible with the Constitution? *Journal of European Continuing Medical Education*. <https://doi.org/10.1080/21614083.2021.1989243>



Triage 4.0: On Death Algorithms and Technological Selection. Is Today's Data- Driven Medical System Still Compatible with the Constitution?

Dirk Helbing, Thomas Beschorner, Bruno Frey, Andreas Diekmann, Thilo Hagendorff, Peter Seele, Sarah Spiekermann-Hoff, Jeroen van den Hoven & Andrej Zwitter

To cite this article: Dirk Helbing, Thomas Beschorner, Bruno Frey, Andreas Diekmann, Thilo Hagendorff, Peter Seele, Sarah Spiekermann-Hoff, Jeroen van den Hoven & Andrej Zwitter (2021) Triage 4.0: On Death Algorithms and Technological Selection. Is Today's Data- Driven Medical System Still Compatible with the Constitution?, *Journal of European CME*, 10:1, 1989243, DOI: [10.1080/21614083.2021.1989243](https://doi.org/10.1080/21614083.2021.1989243)

To link to this article: <https://doi.org/10.1080/21614083.2021.1989243>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 17 Nov 2021.



Submit your article to this journal [↗](#)



Article views: 394



View related articles [↗](#)



View Crossmark data [↗](#)

Triage 4.0: On Death Algorithms and Technological Selection. Is Today's Data-Driven Medical System Still Compatible with the Constitution?

Dirk Helbing ^a, Thomas Beschorner ^b, Bruno Frey ^c, Andreas Diekmann ^d, Thilo Hagendorff ^e, Peter Seele ^f, Sarah Spiekermann-Hoff ^g, Jeroen van den Hoven ^h and Andrej Zwitter ⁱ

^aComputational Social Science, ETH Zurich, Zurich, Switzerland; ^bInstitute for Business Ethics, University of St. Gallen, St.Gallen, Switzerland; ^cCrema, Center for Research in Economics, Management and the Arts, Zürich, Switzerland; ^dInstitut für Soziologie, Universität Leipzig, Leipzig, Germany; ^eAI Research Building, University of Tübingen, Tübingen, Germany; ^fEthics and Communication Law Center, USI Lugano, Lugano, Switzerland; ^gInstitute for Information Systems & Society, Vienna University of Economics and Business, Wien, Austria; ^hDepartment of Values, Technology & Innovation, TU Delft, Delft, The Netherlands; ⁱGovernance and Innovation, University of Groningen, Leeuwarden, The Netherlands

ABSTRACT

Health data bear great promises for a healthier and happier life, but they also make us vulnerable. Making use of millions or billions of data points, Machine Learning (ML) and Artificial Intelligence (AI) are now creating new benefits. For sure, harvesting Big Data can have great potentials for the health system, too. It can support accurate diagnoses, better treatments and greater cost effectiveness. However, it can also have undesirable implications, often in the sense of undesired side effects, which may in fact be terrible. Examples for this, as discussed in this article, are discrimination, the mechanisation of death, and genetic, social, behavioural or technological selection, which may imply eugenic effects or social Darwinism. As many unintended effects become visible only after years, we still lack sufficient criteria, long-term experience and advanced methods to reliably exclude that things may go terribly wrong. Handing over decision-making, responsibility or control to machines, could be dangerous and irresponsible. It would also be in serious conflict with human rights and our constitution.

ARTICLE HISTORY

Received 22 September 2021
Revised 29 September 2021
Accepted 30 September 2021

KEYWORDS

Triage; trolley problem; big data; artificial intelligence; technological selection; eugenics; ethics

“AS A MEMBER OF THE MEDICAL PROFESSION:
I SOLEMNLY PLEDGE to dedicate my life to the
service of humanity;

THE HEALTH AND WELL-BEING OF MY
PATIENT will be my first consideration;

I WILL RESPECT the autonomy and dignity of my
patient;

I WILL MAINTAIN the utmost respect for human life;

I WILL NOT PERMIT considerations of age, dis-
ease or disability, creed, ethnic origin, gender, nation-
ality, political affiliation, race, sexual orientation, social
standing or any other factor to intervene between my
duty and my patient ... ”

So demands “The Physician’s Pledge” (the
“Declaration of Geneva”¹) in its new 2017 version,
also known as the Hippocratic Oath. But are the pro-
cedures used in medicine still compatible with this – or
is our society now threatened by technical selection?

Given the recent use of various forms of triage meth-
ods, among them algorithm-based and data-driven
ones, serious concerns arise.

Illustration of the Problem

For thousands of years, in fact, for the entire history of
humanity, machines were never allowed to make
autonomous, unsupervised life-and-death decisions.
This taboo may soon be broken, or this has happened
already. Not only military machines, but also more
mundane robots like autonomous vehicles, may decide
about lives in split seconds. There are also attempts to
give this an ethical foundation, as the recent work on
the “trolley problem” [1–6] and the “moral machine
experiment” [7]² illustrate. These new approaches,

CONTACT Dirk Helbing  dirk.helbing@gess.ethz.ch  ETH Zürich, Computational Social Science, Stampfenbachstrasse 48, 8092 Zurich, Switzerland
¹<https://www.wma.net/policies-post/wma-declaration-of-geneva/> (last accessed on 18 August 2021).

²A reflection of this research paper has been provided in this and many other newspaper articles: https://www.finanzen100.de/finanznachrichten/wirtschaft/ethische-probleme-im-auto-oma-oder-obdachlosen-ueberfahren-makabrer-mit-test-zeigt-dilemma-selbstfahrender-autos_H439792768_308609/. Note that the Moral Machine experiment has recently been seriously called into question by scientists: [26, 27]. Ethicists have also come to different conclusions, see for example (last accessed August 18, 2021): <https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethik-kommission.pdf>; <https://www.researchgate.net/publication/318340461>; <https://www.zeit.de/gesellschaft/zeitgeschehen/2020-03/deutscher-ethikrat-coronavirus-behandlungsreihenfolge-infizierte>; <https://www.ethikrat.org/fileadmin/Publikationen/Ad-hoc-Empfehlungen/deutsch/ad-hoc-empfehlung-corona-krise.pdf>

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

which may fundamentally change the practices in hospitals and emergency medicine, must undergo constitutional scrutiny.

In cases of scarce medical resources, the decisive criterion used for allocating these resources is typically the probability of the prospective patient to survive. Risk stratification normally comprises three groups: individuals who are unlikely to benefit from a treatment (e.g. because they are anyway expected to die); individuals who can wait for treatment (and will get it, as soon as capacity becomes available – but perhaps never); and individuals who urgently require treatment in order to survive. But how to assess the chance of surviving or whether someone belongs to a certain risk group? Besides using frameworks such as the sequential organ failure assessment (SOFA) or the acute physiology and chronic health evaluation II (APACHE II) [8,9], where the task of risk assessment is currently performed by humans, it may also be performed by machines at some point in time.

Machines, in particular AI systems based on machine learning (ML), are believed to augment or even surpass human capabilities in predicting the outcome of diseases or the likelihood of survival [10–12]. Based on this premise, medical AI tools are increasingly being heralded as a paradigm change in clinical practices. During the Covid-19 pandemic, AI tools became especially attractive, not just for purposes of early warning, tracking, and diagnosis, but also for prognosis [13].

In August 2019, Wynants et al. (2020) [14] listed an impressive number of 107 prognostic models for predicting progression to severe disease, intensive care unit admission, ventilation, intubation, length of hospital stay, and mortality risk (39 models). The sobering conclusion of the study is that all models – except for one [15] – were rated as having a high risk of bias, overfitting due to small or modest sample sizes, and exaggerating reported predictive performance. While the predictive performances of the models were reported to be moderate to excellent, the high risk of bias implies that model performance in out-of-sample detection will be lower, i.e. the generalisability to real-life settings is rather questionable. Hence, the researchers concluded: “We cannot yet recommend any of the identified prediction models for widespread use in clinical practice” [16].

Other studies have yielded similar daunting results [17,18].

A general conclusion from the studies is that predictive models for assessing mortality risk may perform well in the patients that the studies look at, but will not perform well in other patients. Therefore, models that have only been assessed based on calibration performance, while a validation with independent data samples (for different locations, population compositions etc.) is lacking, should not be considered for use. Moreover, machine learning (ML) models for Covid-19 prognostic, similar to ML models in general, make a trade-off between accuracy and fairness. Take for example the role of age in ML models. One can either use age-sensitive models (since age is a significant predictor of Covid-19 outcomes), and thus be relatively accurate, or one can use fair models, which are not age-sensitive, but also less accurate.

Models incorporating race face a similar conflict between accuracy and discrimination. For example, risk prediction models for lung disease, kidney disease, breast cancer, death after heart failure, and other illnesses assign lower scores to patients of colour (possibly because the current health system is biased against them). This may result in less or worse ML-based treatment recommendations as compared to white patients [19]. In other words: discrimination effects in the past, which are reflected by treatment performance data used to train ML/AI systems, may perpetuate discrimination in the future. Concretely, this could mean that algorithm-based triage would put a person of colour into the no treatment group, where a white person would be treated, or in a delayed treatment group, where a white person may get treated immediately, despite comparable health conditions.

Discussion of Some ML-Related Issues in the Health Care System

In view of the possible or even actual use of algorithms for life-or-death decisions³ and the serious issues with the underlying predictive models, there looms a dangerous diffusion of responsibility with unforeseeable consequences for our society. To illustrate the problem, we describe below – in a simplified way – some crucial elements of the process chain in the health care system and how they interact with each other.

³Triage-Software in Notaufnahmen: Der nächste Schnellschuss aus dem Hause Spahn, Netzpolitik.org (23 March 2021) <https://netzpolitik.org/2021/triage-software-in-notaufnahmen-der-naechste-schnellschuss-aus-dem-hause-spahn/> (last accessed on 18 August 2021).

- The reality in hospitals includes capacity constraints, as the COVID-19 crisis has revealed again. In such circumstances, prioritisation decisions are implicitly or explicitly made. In the worst case, these can be triage decisions resulting in the death of people, who would otherwise be saved under normal circumstances.
- Prioritisation decisions are often serious and difficult. It cannot be assumed that all people involved in the related decisions do always fully understand all their ethical and legal implications, so mistakes can easily happen.
- For such and other reasons, medical personnel is increasingly supported in their decisions by measured data and expert systems or even AI systems. This is often very helpful. However, for certain decisions, such decision support can be problematic, because such tools are often created by software developers, who have limited knowledge of medical, legal, ethical, and societal issues.
- The algorithms are typically subject to business secrecy and are frequently updated, such that their effects are partially unknown and, in some cases, are not verifiable due to a lack of algorithm transparency. Accordingly, their results may change unpredictably. Nevertheless, algorithms may exert “epistemic authority”: the compliance with algorithmic recommendations often seems to be advised. It may also appear to be an advantage, in case a lawsuit later occurs.
- Business secrecy is not the only reason why expert or AI systems are often lacking transparency. Machine learning systems have often been characterised as “black boxes” [20], because in contrast to what is common in science and in medical research, ML/AI-based systems are typically not based on transparent, validated, and reproducible causal relations. Moreover, as machine learning progresses, ML-based decision-making and its outcomes may change, without any explanation given. Although machine learning programs have high potential for specific tasks such as skin cancer detection [21], they can also sometimes fail dramatically. “Black-box medical algorithms” should, therefore, be monitored closely; the accuracy of predictions is not granted and requires continuous external validation.
- In concrete applications, medical decisions are often based on criteria such as the expected effectiveness of the measures taken. In this context, decisions are recommended by algorithms that judge between different treatment options – including the possibility of no treatment. The underlying machine decisions are typically based on mean values, such as those obtained in extensive medical studies. However, such mean values are not suited to make precise forecasts on the individual level, because the variation in the data is often large. In other words, a treatment, which is better on average, may actually be worse for certain individuals.
- Triage decisions may also be made on the basis of data that was never collected and intended for triage decisions. As a result, data quality can be insufficient or the application context of the data can be inappropriate. Therefore, the application in a Triage context is likely to be problematic.
- *Data-driven decisions are often very sensitive to details of the algorithm used or the dataset evaluated.* In other words, taking a data-driven decision using a different dataset or a different algorithm (as other research teams might provide them) may result in pretty different priorities. Consequently, the selection of patients who are disadvantaged by triage decisions or medically prioritised can greatly vary with the procedures, algorithms, or datasets used. Hence, the approach of data-driven decisions may suffer from an undesirable, but hidden degree of randomness and arbitrariness, which is particularly worrying when it comes to life-and-death decisions. Therefore, “data-driven” or “evidence-based” does not automatically mean that the resulting decisions are scientifically objective and sound at the level one would require for grave decisions. Despite these methodological shortcomings, vital decisions are increasingly taken on the basis of algorithms.
- The legitimacy and meaningfulness of the algorithms and personal data used in such contexts may be questioned as well. For example, some of these algorithms seem to be based on life expectancy⁴, even though this may vary a lot.⁵ There are at least two undesirable consequences: (1) While life expectancy

⁴Soliman T. Der Todesalgorithmus: Computer berechnet Lebenserwartung, Das Erste; 2017 Dec 14; Available from: <https://daserste.ndr.de/panorama/archiv/2017/Der-Todesalgorithmus-Computer-berechnet-Lebenserwartung,todesalgorithmus112.html>

⁵In particular, the mean life expectancy varies considerably across countries. Thus, if such an algorithm is fed with international data, there is a risk of life-shortening decisions in countries, where the mean life expectancy would actually be higher.

is actually quite variable [22],⁶ the application of algorithms to decide the level of life support may *determine* the lifespan pretty much in the sense of a self-fulfilling prophecy. (Thereby, “good” and “bad luck” would be largely eliminated – and hope as well.) (2) While the expectation of a long life may be rewarded, the expectation of a short life would be almost like a “death sentence”, particularly when the prediction is incorrect.⁷

- If algorithms were developed with the aim to meet limited budgets or reduce costs,⁸ this could imply that AI-based technological selection might shorten lifespans systematically. This is a serious threat since, after all, Big Data-driven medical decisions are often also supposed to ensure the *efficient* use of available funds. Such a development can shift the focus away from human dignity and equality to economic profit, and it may obfuscate the rationing of health-care resources.
- If, in addition, individual insurance coverage (the treatments reimbursed by the respective health insurance) is taken into account, social selection could happen.⁹ Furthermore, as known from recent AI research, such discriminatory effects [23] can even occur where not intended at all, namely as a result of the opacity of the algorithms or the training data used.¹⁰
- Despite all this, in the context of the current pandemic, triage decisions involving life-and-death decisions are increasingly being made.¹¹ They are often

presented as inevitable in view of insufficient medical capacities.¹² Nevertheless, triage decisions are considered to be extremely problematic and acceptable only in war times or extreme disasters.

In conclusion, the application of algorithms for life-and-death decisions can imply technological selection and a *mechanisation of death*¹³ [6], particularly if these are allowed to act autonomously.

- Nevertheless, there is a trend to transfer problematic decisions to data-driven algorithms and autonomous systems. In some countries, algorithm-based triage decisions seem to be already in use.¹⁴ However, this makes human beings objects of machine decisions, which is not compatible with human dignity. Besides, the transfer of human responsibility to machines undermines the principle of accountability.
- Political decisions limiting the health-care budget will inevitably determine how many people are ultimately affected by triage decisions, even if not intended.¹⁵ However, algorithm-based triage will also depend on a person’s health condition and, hence, on the person’s social conditions and genetics. Therefore, algorithm-based life-death decisions, i.e. technological selection, may also imply social and eugenic selection.
- The COVID-19 pandemic has demonstrated the threat of introducing (at least partially) automated

⁶In: Nature Public Health Emergency Collection, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7585743/>

⁷Wilson, J. 23andMe raises questions about at-home genetic testing, CNN health; 2013 Nov 26; Available from: <https://edition.cnn.com/2013/11/26/health/23andme-fda-genetic-testing/index.html>

⁸Bundesrat Berset auf Kollisionskurs mit der Verfassung? Seine Pläne für einen “Deckel” bei den Gesundheitskosten stossen auf scharfe Kritik, Neue Zürcher Zeitung (17 November 2020) <https://www.nzz.ch/schweiz/bundesrat-ber-set-auf-kollisionskurs-mit-der-verfassung-seine-plaene-fuer-einen-deckel-bei-den-gesundheitskosten-stossen-auf-scharfe-kritik-ld.1587291>

⁹Ungleichbehandlung je nach Krankenkasse: Wo man versichert ist, kann über Leben oder Tod entscheiden, Tagesanzeiger (6 March 2021) <https://www.tagesanzeiger.ch/wo-man-versichert-ist-kann-ueber-leben-oder-tod-entscheiden-752186910381>

¹⁰Also note that ethnic, social, genetic, and/or behavioural factors may sometimes correlate.

¹¹“Weiche Triage wird bereits angewendet”: Wie Ärzte entscheiden, wer behandelt wird, Focus (19 April 2021) https://www.focus.de/gesundheit/news/triage-auf-intensivstationen-dritte-welle-wenn-mediziner-entscheiden-muessen-wer-behandelt-wird_id_13195703.html

¹²even though some insiders have actually denied such capacity shortages: Corona und Krankenhäuser: “Ein Patient wird unter Umständen doppelt gezählt” WELT (16 June 2021) <https://www.welt.de/politik/deutschland/plus231872027/Corona-Krankenhaeuser-Ein-Patient-wird-unter-Umstaenden-doppelt-gezaehlt.html>

¹³Helbing D, Seele P. Death by algorithm? Project Syndicate. 2020 Nov 26. <https://www.project-syndicate.org/commentary/artificial-intelligence-resilience-covid19-climate-change-by-dirk-helbing-and-peter-seele-2020-11>.

¹⁴Hao K. Doctors are using AI to triage COVID-19 patients. The tools may be here to stay. MIT Technol Rev. 2020 Apr 23; Available from: <https://www.technologyreview.com/2020/04/23/1000410/ai-triage-covid-19-patients-health-care/>

¹⁵Apparently, emergency capacities may have been reduced by the circumstance that empty beds were profitable: RKI-Schreiben zu Intensivbetten: “Monetäre Anreize” für falsche Angaben, Tagesschau (17 June 2021) <https://www.tagesschau.de/investigativ/wdr/intensivbetten-daten-101.html>

procedures for life-death decisions into everyday clinical practice. Thereby, triage measures would no longer be limited to situations of wars and serious disasters, but become a “new normal”. We consider such uses of triage procedures as unconstitutional, which must be prevented by political and legal oversight.

Threats of Algorithmic Decisions

We have explained above why undesirable, potentially hidden, genetic and social selection effects are to be expected in a health-care system, which is increasingly operated in a data-driven way. To minimise these, we suggest that medical institutions should regularly produce statistical evaluations of any genetic, social, and behavioural selection effects. Moreover, the application of algorithms as well as the algorithms’ functioning should be regularly audited. Discrimination as well as the non-compliance with state-of-the-art quality standards should be sanctioned. Otherwise, there is an acute danger that those people, who need our medical support and solidarity most, will be particularly disadvantaged by the data-driven systems emerging today.

In our opinion, current data-based approaches do not adequately meet constitutional requirements (principle of equality, self-determination, right to life) and policy goals (such as equity) today. They are also not scientifically sound enough (given the sensitivity of many data analytics approaches). To unfold their full benefits, it must be ensured that Big Data, Artificial Intelligence and other technological innovations are used in a fair way – and that problematic applications are avoided. Otherwise, it is to be feared that the democratic principle of equality will increasingly be replaced by discriminatory scoring systems, *which would ultimately lead to a fundamentally different society*.

Although constitutional and fair data-based methods may be possible to develop, the methods currently in use do not meet all justified expectations. If one takes the principle of equality seriously, according to which no human life should count more than another, then the blessings of modern medicine would have to benefit all people equally. This fairness requirement would imply that health-disadvantaged people with a shorter life expectancy

should get greater medical support per year, not less. This, however, would require a completely different use of data-based methods and different goal functions.

Note that our assessment is supported by the recent WHO Guidance on “Ethics and Governance of Artificial Intelligence for Health”, which points out:¹⁶ “Use of computerized decision-support programs – AI or not – to inform or guide resource allocation and prioritization for clinical care has long raised ethical issues. They include managing conflicts between human and machine predictions, difficulty in assessing the quality and fitness for purpose of software, identifying appropriate users and the novel situation in which a decision for a patient is guided by a machine analysis of other patients’ outcomes. In some situations, well-intentioned efforts to base decisions about allocations on an algorithm that relies only on a rules-based formula produce unintended outcomes”. In its section on “The ethics of resource allocation and prioritization” it also warns that “if an AI technology is trained to ‘maximize global health’, it may do so by allocating most resources to healthy people in order to keep them healthy and not to a disadvantaged population”. Last, but not least, “Use of AI tools for triage or rationing is one of the most compelling reasons for ensuring adequate governance or oversight. Although intentional harm is not ethically controversial – it is wrong – the possibilities of unintended bias and flawed inference emphasize the need to protect ... people and processes from computational misadventure”.

Technological and Ethical Awareness Needed

AI systems will certainly play an increasing role in the medical profession. Augmented intelligence has a large potential in assisting medical doctors, medical personnel, and patients in health-related decision-making. These systems can also raise the quality and cost-effectiveness of health care. However, there are a lot of caveats concerning fairness, equality, and bias in the treatment of certain patient groups, regarding the accuracy and validity of results, unintended side effects, and the transparency of algorithms. Therefore, the use of AI systems in practice will typically raise serious ethical questions.

Health care professionals must be aware of the underlying principles, possible weaknesses, unintended consequences, and ethical problems when applying AI systems. After all, they are responsible for the results.

¹⁶World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. 2021. p. 49f. Available at <https://www.who.int/publications/i/item/9789240029200>









Thus, in a recent report, the US National Academy of Medicine demands: “Develop and deploy appropriate training and educational programs to support health care AI” [24]. These programs have to be integrated in the medical education curricula and they should become a high priority in the continued education of health care professionals.

AI courses for continued medical studies are already offered by certain medical schools (e.g. Massachusetts Medical Society, The Radiological Society of North America, Mayo Clinic, Stanford University School of Medicine) [25]. However, courses in continued medical education should, besides a focus on engineering and technical issues, place a higher emphasis on understanding the principles of algorithms and on reflection of related ethical problems. Consequently, not only medical experts and engineers should be responsible for teaching AI courses in medicine. It is highly important that AI programs in medical studies and in continued education “must be multidisciplinary and engage AI developers, implementers, health care system leadership, frontline clinical teams, ethicists, humanists, patients, and caregivers” [24].

Disclosure Statement

No potential conflict of interest was reported by the author(s).

ORCID

Dirk Helbing  <http://orcid.org/0000-0002-9898-0101>
 Thomas Beschorner  <http://orcid.org/0000-0001-7097-647X>
 Bruno Frey  <http://orcid.org/0000-0001-5380-6655>
 Andreas Diekmann  <http://orcid.org/0000-0002-7344-5591>
 Thilo Hagedorff  <http://orcid.org/0000-0002-4633-2153>
 Peter Seele  <http://orcid.org/0000-0002-6456-4003>
 Sarah Spiekermann-Hoff  <http://orcid.org/0000-0001-5481-3961>
 Jeroen van den Hoven  <http://orcid.org/0000-0003-2376-3185>
 Andrej Zwitter  <http://orcid.org/0000-0001-7391-141X>

References

- [1] Foot P. The problem of abortion and the doctrine of double effect. *Oxford Rev.* 1967;5:5–15.
- [2] Thomson JJ. Killing, letting die, and the trolley problem. *Monist.* 1976;59:204–217.
- [3] Thomson JJ. The trolley problem. *Yale Law J.* 1985;94(6):1395–1415. Available from: <http://www.psy.vanderbilt.edu/courses/hon182/thomsonrolley.pdf>
- [4] Etzioni A, Etzioni O. Incorporating ethics into artificial intelligence. *J Ethics.* 2017;21(4):403–418.
- [5] Andrade G. Medical ethics and the trolley problem. *J Med Ethics Hist Med.* 2019;12:3.
- [6] Nagler J, van den Hoven J, Helbing D. An extension of Asimov’s robotic laws. In: Helbing D, editor. *Towards digital enlightenment.* Berlin: Springer; 2019:41–46.
- [7] Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., and Rahwan, I. The moral machine experiment. *Nature.* 2018;563:59–64.
- [8] Jones AE, Trzeciak S, Kline JA. The sequential organ failure assessment score for predicting outcome in patients with severe sepsis and evidence of hypoperfusion at the time of emergency department presentation. *Crit Care Med.* 2009;37(5):1649–1654.
- [9] Enfield K, Miller R, Rice T, et al. Limited utility of SOFA and APACHE II prediction models for ICU triage in pandemic influenza. *Chest.* 2011;140(4):913A.
- [10] Topuz K, Zengul FD, Dag A, et al. Predicting graft survival among kidney transplant recipients: a Bayesian decision support model. *Decis Support Syst.* 2018;106:97–109.
- [11] Overweg H, Popkes A-L, Ercole A, et al. Interpretable outcome prediction with sparse Bayesian neural networks in intensive care. *arXiv.* 2019. Available from: <https://arxiv.org/abs/1905.02599>
- [12] Dong R-Z, Yang X, Zhang X-Y, et al. Predicting overall survival of patients with hepatocellular carcinoma using a three-category method based on DNA methylation and machine learning. *J Cell Mol Med.* 2019;23(5):3369–3374.
- [13] Naudé W (2020) Artificial intelligence against Covid-19. An Early Review. SSRN, Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3568314
- [14] Wynants L, van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ.* 2020;369:1–16. Available from: <https://www.bmj.com/content/369/bmj.m1328>
- [15] Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO clinical characterisation protocol: development and validation of the 4C mortality score. *BMJ.* 2020;370:1–13. Available from: <https://www.bmj.com/content/370/bmj.m3339>
- [16] Wynants L, van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ.* 2020;369:8.
- [17] Chen J, See KC. Artificial intelligence for COVID-19: rapid review. *J Med Internet Res.* 2020;22(10):e21476.
- [18] Roberts M, Driggs D, Thorpe M, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Mach Intell.* 2021;3(3):199–217.
- [19] Madhusoodanan J. A troubled calculus. Researchers use race to build disease risk assessment tools. Can removing it help resolve medicine’s race crisis? *Science.* 2021;373:380–383.
- [20] Price WN. Big data and black-box medical algorithms. *Sci Transl Med.* 2018;10(471):eaao5333. Available from: <https://stm.sciencemag.org/content/10/471/eaao5333/>

- [21] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115–118.
- [22] Flick H. Letalität und Risikofaktoren für einen schweren Verlauf der COVID-19-Pneumonie. *Der Pneumol*. 2020;17:385–393. Available from: <https://link.springer.com/article/10.1007%2Fs10405-020-00349-y>
- [23] Turner Lee N. Detecting racial bias in algorithms and machine learning. *Journal of Information, Communication, and Ethics in Society* 2018;16(3):252–260.
- [24] Matheny, M.E., Whicher, D., Israni, S.T. Artificial Intelligence in Health Care. A Report From the National Academy of Medicine. *JAMA*. 2020;323(6):509-510.
- [25] Paranjape K, Schinkel M, Panday RN, et al. Introducing artificial intelligence training in medical education. *JMIR Med Educ*. 2019;5:1–7.
- [26] Dewitt B, Fischhoff B, Sahlin N-E. Policy flaw in moral machine experiment. *Nature*. 2019;567:31.
- [27] Bigman YE, Gray K. Life and death decision of autonomous vehicles. *Nature*. 2020;579:E1–E2.