

Supplementary information to:

Letter to the editor:

**ARTIFICIAL INTELLIGENCE IN HOSPITALS:
LEGAL UNCERTAINTIES AND EMERGING RISKS
FOR PATIENT SAFETY**

Meriem Gaddas^{1,2}, Mohamed Ben Dhiab³, Imen Ben Saida⁴,
Helmi Ben Saad^{1,2}

¹ University of Sousse, Faculty of Medicine 'Ibn el Jazzar' of Sousse, Farhat HACHED University Hospital, Research Laboratory LR12SP09 'Heart Failure' Sousse, Tunisia

² Department of Physiology and Functional Explorations, Farhat HACHED University Hospital, Sousse, Tunisia

³ University of Sousse, Faculty of Medicine of Sousse, Department of Forensic Medicine, EPS Farhat HACHED of Sousse, Tunisia

⁴ University of Sousse, Faculty of Medicine of Sousse, Department of Intensive care, Farhat Hached University Hospital, Sousse, Tunisia

* **Corresponding author:** Helmi Ben Saad, University of Sousse, Faculty of Medicine 'Ibn el Jazzar' of Sousse, Farhat HACHED University Hospital, Research Laboratory LR12SP09 'Heart Failure' Sousse, Tunisia. E-mail: helmi.bensaad@rns.tn

<https://dx.doi.org/10.17179/excli2025-8679>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>).

Table S1: Major documented cases of failures involving artificial intelligence (AI) tools already deployed in clinical practice

Year, references	Software type Developer	Nature of failure	Patient harm	Charge/Lawsuit Outcome of the case	Responsibility attribution gaps
2013-2021 (henricodol- fing, 2024; Luxton, 2019; Zhou et al., 2019)	WFO: diagnostic and therapeutic decision support .applied in 14 countries IBM & Memorial Sloan Kettering cancer center	High discordance rates compared to recommen- dations, varying depend- ing on the histological type (which can reach up to 88 %) Suggestions for inappro- priate and even danger- ous treatments (eg; un- justified aggressive chemotherapy)	Inappropriate treatments, management errors	No known direct legal action IBM disabled Watson in 2021 Reputation tarnished	Lack of trans- parency Black box Untraceable rec- ommendations
2015-2021 (Powles and Hodson, 2017)	DeepMind Health/Streams (Fol- low-up of kidney failure cases) Google (DeepMind) plus NHS (UK's NHS)	Illegal use of patient data without their explicit con- sent (data privacy issue) .Errors in clinical notifica- tions	Delays in the detection of kidney failures No directly proven deaths, but identified risks and doc- umented cases	UK authorities' investigation for vi- olation of the data protection act (1998) No prior approval from authorities Administrative sanctions, with- drawal of the initial version (but the emergence of other derivative applications) A public mistrust of this technol- ogy	The AI nature of the tool not rec- ognized by the developers Shared respon- sibility between the developer and the NHS
2019 (Obermeyer et al., 2019)	Population health risk stratification algorithm Optum (a subsidiary of United Health)	Structural racial bias: The algorithm prioritized pa- tients with the highest medical expenses as an indicator of the severity of their health condition The algorithm systemati- cally underestimated the needs of Black patients (with limited income)	The tool underestimated the level of risk and excluded Black patients from neces- sary intensive or preventive care programs, potentially worsening their health con- dition	No known individual legal com- plaints The developer acknowledged the bias and announced a revision of its algorithm No official public sanctions	Lack of a clear regulatory framework on the responsibility for algorithmic biases Opacity of pro- prietary algo- rithms Difficulty in demonstrating

					direct individual harm
2020 (Dufour et al., 2020)	Autonomous intensive care ventilators (Evita V300/V500) Dräger Medical	Automatic activation of a pressure-limiting algorithm without warning	Resulting in desaturation followed by asystole or coma in intubated patients , requiring resuscitation Other similar incidents have also been reported	Multiple reports were submitted to the ANSM. Due to the lack of confirmed technical malfunction, the cases were closed without resolution, no legal action was taken At the request of ANSM, the manufacturer was initially required to update the user manual and later committed to "correcting" the operating algorithms to enhance patient safety	Concerns persist regarding lack of transparency Non-traceable recommendations Presence of a "black box" system No formal investigations were conducted for the reported cases
2021 (Wong et al., 2021)	ESM: Early detection of sepsis Deployed in hundreds of hospitals across the US Epic systems corporation (largest health care software vendors in the world)	Low diagnostic performance Sensitivity of 33 % and specificity of 83 % Generates a high volume of alerts, leading to alert fatigue and potential patient safety risks Poor calibration: inconsistency between predicted scores and actual patient risk	67 % of patients with sepsis were not identified by the model 7 % of patients did not receive timely antibiotic treatment	No specific cases have been documented No legal actions have been initiated 2022: In response to growing criticism, Epic made adjustments to the model A 2023 study showed that ESM still underperformed compared to other available tools (Schertz et al., 2023) In light of ESM's limitations, new AI-based models have been developed, such as the COMPOSER model (Boussina et al., 2024) Despite its flaws, ESM continues to be used in hundreds of US hospitals (Becker Health IT, 2024)	Implemented without prior external validation

ANSM: French national agency for the safety of medicines and health products; **ESM:** Epic sepsis model; **IBM:** International business machines corporation; **NHS:** National health service. **WFO:** Watson for oncology