

论产品责任在人工智能辅助医疗诊断系统侵权责任认定中的应用

杨 杰

北京物资学院法学院, 北京

收稿日期: 2026年5月12日; 录用日期: 2026年6月8日; 发布日期: 2026年6月18日

摘 要

人工智能辅助医疗诊断系统的广泛应用对现行侵权责任规则提出了挑战。由于“算法黑箱”的存在, 医务人员的审查能力被实质性削弱, 继续沿用传统医疗损害责任有违公平, 而产品责任的适用亦面临缺陷认定难与因果关系断裂的困境。本文立足于风险社会理论, 论证了对此类系统适用产品责任的法理正当性。在规则构建上, 建议重构产品缺陷认定标准, 引入事实与法律结合的双重因果关系判定规则。最后, 基于人机协作的不同模式, 将侵权形态类型化为“隐秘型缺陷”、“常识型缺陷”及“人机耦合型缺陷”, 并据此构建差异化的责任分担机制, 以期在救济受害人与鼓励技术创新之间寻求平衡。

关键词

人工智能辅助医疗诊断, 产品责任, 算法黑箱, 责任承担, 医疗损害

On the Application of Product Liability in Determining Tort Liability for AI-Assisted Medical Diagnostic Systems

Jie Yang

Law School, Beijing Wuzi University, Beijing

Received: May 12, 2026; accepted: June 8, 2026; published: June 18, 2026

Abstract

The widespread deployment of AI-assisted medical diagnostic systems poses challenges to existing tort liability rules. Due to the existence of the “algorithmic black box”, the reviewing capacity of medical personnel is substantially weakened, making it inequitable to continue applying traditional

rules of medical malpractice liability. Moreover, the application of product liability also faces difficulties in defect identification and the breakdown of causation. Drawing upon risk society theory, this article demonstrates the jurisprudential legitimacy of applying product liability to such systems. With respect to rule construction, it proposes reconstituting the standard for determining product defects and introducing a dual causation test that integrates factual and legal causation. Finally, based on different modes of human-machine collaboration, it typologizes infringing acts into “latent defects”, “common-sense defects”, and “human-machine coupling defects”, and accordingly constructs differentiated liability allocation mechanisms, with a view to striking a balance between victim compensation and the encouragement of technological innovation.

Keywords

AI-Assisted Medical Diagnosis, Product Liability, Algorithmic Black Box, Liability Allocation, Medical Malpractice

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 问题的提出

人工智能辅助医疗诊断系统是人工智能在医疗领域的应用，近年来，人工智能辅助医疗诊断系统广泛应用于临床实践。美国 2018 年正式批准全球首款人工智能辅助诊断系统 IDx-DR，日本的 CAD 系统通过分析 CT 照片检查肺结核等疾病。沃森肿瘤系统已经能筛查出 13 种器官中的肿瘤，谷歌公司开发的检测转移性乳腺癌的人工智能辅助医疗诊断系统的准确率比人类病理学家高出 61% ([1], p. 48)。目前，我国对人工智能应用于医疗领域持积极肯定态度，《国务院关于深入实施“人工智能+”行动的意见》¹以及《关于促进和规范“人工智能+医疗卫生”应用发展的实施意见》²，指出要规范人工智能在医疗领域的应用，不断丰富应用场景，推广医学影像智能诊断服务，更好满足人民群众日益增长的健康服务需要。我国继 2020 年以来批准了一系列人工智能辅助医疗诊断系统，比如糖尿病视网膜病变眼底图像辅助诊断软件。北京某医院引入人工智能系统辅助医学影像识别和早期筛查³，某区 9 家医院应用心肺联筛人工智能系统实现冠心病和肺癌风险评估⁴。温州某区部署人工智能系统构建覆盖诊前至诊后的慢病管理闭环⁵。截至 2023 年底，27 个试点县的基层机构通过辅助诊疗系统提供诊断建议 2600 余万次。除这些试点地区以外，安徽、江西、宁夏等地也在全省(自治区)范围内推广辅助诊疗系统基层应用⁶。人工智能算法介入传统的医疗诊断，使得传统医患二元结构演变为三元结构，同时，人工智能辅助医疗诊断系统涉及多方主体，包括设计者、生产者和数据提供者等，设计者和数据提供者也可能获得独立的责任主体地位，承担侵权责任，这使得问题更加复杂⁷。

¹http://www.scio.gov.cn/zdgz/jj/202509/t20250901_928364.html

²<https://www.nhc.gov.cn/guihuaxxs/c100133/202511/d1a42ae835c743b9b3e83ac0253c9ef.shtml>

³参见《北京儿童医院嘉禾儿科诊疗中心落户丰台》，载环京津新闻网。

<https://baijiahao.baidu.com/s?id=1844131152547501226&wfr=spider&for=pc>

⁴参见《AI 医生基层助筛心肺疾病，昌平群众看病用上“黑科技”》，载北京商报。

<https://baijiahao.baidu.com/s?id=1804179557442708985&wfr=spider&for=pc>

⁵参见《智慧健康服务领域新成果！人工智能辅助诊断系统助力鹿城区基层医疗能力提升》，载微信公众号“健康鹿城”2024 年 12 月 31 日。<https://mp.weixin.qq.com/s/DXgVY2dfIcmhlM2FXBzkwQ>

⁶参见《国家卫生健康委：推进医学人工智能辅助诊断技术在基层的应用》，载四川省卫生信息学会网。

<http://www.schia.cn/show.jsp?id=2802>

⁷责任主体这一问题十分复杂，囿于篇幅限制，本文不进行深入讨论。为了本文论述的连贯性和简洁性，在论述过程中将设计者、生产者和数据提供者这些处于生产端的主体整体称为“生产者”。

在这样的背景下,人工智能辅助医疗诊断系统侵权致人损害,若适用医疗损害责任将有违公平,但适用产品责任同样面临困难和阻碍。通过北大法宝检索关键词“人工智能”“侵权”,选择“医疗卫生”专题后找到一案例,马某某等与河北省人民医院等医疗损害责任纠纷案⁸。在本案中涉及的软件春雨医生符合本文讨论的人工智能辅助诊断系统,通过试用发现,其主要包括快速问诊、买药、寻找医院等功能,其中“春雨慧问”是帮助解答医疗专业问题的人工智能系统。这一类型软件应用于医疗诊断的基本运行模式为,医务人员输入患者病情,系统根据患者病情,结合自身数据库中的医学知识作出初步的诊疗判断。在这样的诊断模式下,人工智能辅助医疗诊断系统在作出诊断决策的过程中若出现算法错误,但得出的错误结论具有临床欺骗性,医务人员无法识别,此时医务人员的诊断地位被实质性架空,审查能力也被实质性削弱,适用医疗损害责任将违背公平原则。而适用产品责任又面临如何判断人工智能辅助医疗诊断系统存在产品缺陷,产品缺陷与损害结果之间的因果关系是否会因为医务人员介入的诊断行为而中断等法律适用壁垒。至于人工智能辅助医疗诊断系统软件是否属于产品这一问题,学界通说与国内外司法实践已经逐渐形成共识,嵌入式或独立的医疗软件具有产品属性,属于产品([2], p. 28)。

法律的滞后性决定其必须要根据社会实践的发展不断作出调整,在新的法律关系出现时能够为受害人提供明确可操作的规则指引。通过对比《中华人民共和国民法典》(后文简称《民法典》)1202条和1218条,在人工智能辅助诊断的模式下,适用产品责任更公平。第一,二者的归责原则不同。产品责任采用无过错责任原则,医疗损害责任采用过错责任原则。第二,在举证责任方面,产品责任中,受害者需要证明产品存在缺陷、发生损害结果以及产品缺陷与损害结果之间存在因果关系。医疗损害责任中,受害人需要证明自己在诊疗过程中受到损害,医疗机构存在过错和损害行为,损害行为和损害结果之间存在因果关系。在产品责任框架下,生产者不能依据算法“黑箱”的不可解释性而逃避责任,这样既便于患者维护合法权益,又在医疗机构和人工智能辅助医疗诊断系统生产者之间合理分配了责任,相较于医疗损害责任,产品责任更具有合理性,因此本文基于传统产品责任框架,论证产品责任在人工智能辅助医疗诊断中的应用,突破人工智能辅助医疗诊断系统适用产品责任的理论障碍,重构符合人工智能辅助医疗诊断系统侵权责任的分配规则。

2. 人工智能辅助医疗诊断系统适用产品责任的困境

(一) 人工智能辅助医疗诊断系统产品缺陷标准不明

产品缺陷可分为设计缺陷、制造缺陷、警示缺陷与跟踪缺陷([3], pp. 348-353),制造缺陷是指因产品原材料或配件存在缺陷或者在装配最终产品的过程中出现错误,从而导致产品具有不合理危险性([3], p. 350)。人工智能辅助医疗诊断系统制造缺陷判断较为容易,只要最终呈现的系统不符合原始设计,即可判断存在制造缺陷([4], p. 215)。因此判断人工智能辅助医疗诊断系统存在产品缺陷的困难集中在设计缺陷、警示缺陷与跟踪缺陷上。

第一,在设计缺陷上,人工智能难以消除算法偏见,人工智能的运算逻辑尚未实现因果推理,在个案中容易产生方向性偏差,里奇·卡鲁安纳的研究发现,某肺炎死亡率预测模型错误地将哮喘患者评为低风险,与临床经验明显背离,若被采信可能造成严重误判([5], p. 135)。判断是否存在设计缺陷,学理上存在消费者期待标准和风险效用两个标准。消费者期待标准是指产品不合理危险超过了购买该产品的消费者的合理期待,强调将产品的安全水平和消费者的合理期待进行比较。由于消费者对于人工智能辅助医疗诊断的认知水平有限,消费者本身难以产生合理期待,从而该标准难以适用。风险效用标准是指从客观上进行成本效应分析,判断产品在进入流通之前是否能够被设计得更加安全,是否存在一个更加合理的替代设计。一方面算法“黑箱”导致目前人工智能的发展处于“不可解释”的阶段,系统作出决策的过程复杂且不透明,

⁸北京市朝阳区人民法院:马某某等与河北省人民医院等医疗损害责任纠纷,(2016)京0105民初34091号。

难以解释结果的得出过程,此时无法找到替代设计。另外一方面,人工智能诊断的结果不强制医务人员采纳,医务人员的决策行为的介入使得成本效应的评估更加复杂([4], p. 205)。第二,警示缺陷是指生产者没有提供适当的警示与说明,致使产品存在不合理危险。人工智能辅助医疗诊断系统最直接的法律挑战之一就是医疗说明义务的冲击。医疗领域存在“专业中间人”规则,如果产品是在医务人员的指导下使用的,医务人员具有专业知识,本质上是对系统本身的监管,将人工智能辅助医疗诊断系统给出的决策与自身专业知识结合作出最终判断。如果生产者已经向医务人员提供了完全的信息,此时就应当认为医务人员能够作出安全且正确有效的诊断。但从现实来看,医务人员也同样面临算法“黑箱”,对其内部的运行逻辑难以完全掌握,如果让其承担过重的责任,会导致责任承担失衡,转嫁风险([5], p. 135)。第三,产品的生产者对于投入流通的产品负有质量跟踪义务,即使投入流通时不能发现缺陷存在,但通过跟踪能够发现缺陷,则应当及时采取警示、召回等补救措施([3], p. 353)。若未能及时采取补救措施,造成侵权损害结果,那么则认为存在跟踪缺陷。人工智能辅助医疗诊断系统因科技发展水平的局限,大多在刚投入流通时无法发现缺陷的存在,需要生产者持续跟踪监测该系统在运行过程中是否存在缺陷。但算法“黑箱”会让追踪变得困难,若生产者履行了追踪义务但因人工智能自主学习和不可解释的特征,导致生产者未能及时发现缺陷,此时生产者能否进行抗辩,若能够进行抗辩,边界又在哪里,都是需要解决的问题。

(二) 三元医患结构导致产品缺陷致人损害因果关系认定困难

在传统产品责任中,作为构成要件之一的因果关系是指产品的缺陷与受害人遭受损害之间的相互关系。它具有一般侵权责任中因果关系的共性,也具有其特殊性([6], pp. 211-212)。人工智能辅助医疗诊断系统侵权致人损害的因果关系认定困难主要体现在,第一,人工智能辅助医疗诊断系统决策的过程是算法利用大数据发现事物之间的相关性而非因果性联系([7], p. 174),算法“黑箱”导致可解释性难题和可追溯性难题([7], p. 174)。由于算法模型越复杂,可解释性就越低,目前我们无法理解人工智能为什么作出这样的决策,即使依照透明性要求能够发现人工智能的不合法行为或瑕疵设计,也无法解释完整的决策作出过程,很难准确找到是哪一个数据或信息的错误导致了最终决策的错误,这使得因果关系判断过程更加复杂。第二,难以认定医务人员诊疗行为的介入是否中断了人工智能辅助医疗诊断系统缺陷和损害结果之间的因果关系。在运用人工智能辅助医疗诊断系统的过程中会介入医务人员的诊疗行为,医务人员会结合系统给出的决策意见作出最终的诊断,即人工智能辅助医疗诊断系统并非直接作用于人身,医务人员会在判断系统决策正确与否的基础上结合自身专业知识给出最后的诊断。由此可知,如果医务人员与人工智能辅助医疗诊断系统共同诊断导致侵权结果发生的情况下,需要分别判断因果关系。一方面要判断是否因系统自身存在产品缺陷导致了决策错误,造成损害结果。另一方面要判断医务人员是否尽到了合理注意义务,其诊断地位是否被实质性替代,其诊断行为是否中断了人工智能辅助医疗诊断系统缺陷与损害结果之间的因果关系。

3. 人工智能辅助医疗诊断系统侵权适用产品责任的法理基础与边界

(一) 风险社会理论决定人工智能辅助医疗诊断系统生产者应承担安全风险

“风险”的概念起源于现代社会,是指在社会现代化过程中产生的各种不可预测且难以控制的潜在危害。人类面临着威胁其生存的由社会所制造的风险([8], p. 241)。德国社会学家乌尔里希·贝克将“风险”的概念结合社会学研究领域,结合自反性现代化、系统脆弱性、全球性等理论,演绎成为风险社会理论。自反性现代化是风险社会理论的核心概念,自反性是风险社会具备的特性。科学和技术的发展不仅带来了便利,同时带来了新的问题和挑战。以上问题的解决又反过来依赖新知识和新技术加以解决([9], p. 129-130)。人工智能辅助医疗诊断系统是科技发展过程中人工智能应用于医疗的产物,在该系统应用于临床的过程中,可能会因为算法错误造成患者身体健康权受损,这不仅威胁了人类的身体健康,还引发了一

系列法律问题，需要通过更新规则解决问题。

在风险社会中，风险的分配具有整体性和不平等性。整体性是指在风险的增长和分配中将打破阶级，威胁社会全体成员的安全。但不平等性决定了处于优势地位的社会群体能够主动参与到风险分配的过程中，促成对自己更有利的风险分配方案，在占据大部分财富的同时承担更少的风险。在人工智能辅助医疗诊断的过程中，生成的错误诊断结果造成的损害可能会作用于每一个人，我们都是潜在的受害人，即使是系统的生产者也无法逃避。然而人工智能辅助医疗诊断系统的生产者在占据主要财富的同时如果能够基于算法具有不可解释性就能转嫁风险，规避责任，使处于弱势地位的其他群体承担大部分的安全风险显然有失公平，处于优势地位的生产者理应承担起更多的责任和义务，防范风险^[9], p. 130)。

我国产品责任的制度设计，依据《民法典》第 1202、1203 条的规定可知，责任主体是生产者和销售者，在责任承担方式上，生产者责任和销售者责任是相互独立且对立的，如果生产者承担了责任，那么销售者不再承担责任，反之亦然。如果生产者或销售者一方代替另一方承担了赔偿责任，那么可以向另一方追偿⁹。产品责任是基于生产者和销售者的优势地位，为了保护消费者合法权益设置的无过错责任归责方式，与风险社会分配逻辑所契合。

医务人员在利用人工智能医疗辅助诊断系统提高决策效率，弥补知识空白的同时，还存在潜在的风险，这个风险来源于该系统的生产者，直接作用于患者，因此，只有生产者处于优势地位，医务人员和患者均处于劣势地位，同时承担着更多的风险。根据风险社会理论，风险应当在不同主体之间公平分配，生产者也应当承担责任，履行义务，因此应当根据产品责任规则，在侵权损害结果发生时，由生产者承担侵权责任。

(二) 生产者基于算法“黑箱”处于优势地位

“算法黑箱这一概念表达了由于算法不透明而引发的人类失去对决策过程控制这一风险的担忧”¹⁰。根据人工智能的可解释性，人工智能可以类型化为“白箱”“灰箱”和“黑箱”三种状态。应用于辅助医疗诊断的人工智能属于“黑箱”人工智能。“黑箱”人工智能模型结构中的神经网络层数和每层神经元数量等均不可见，人工智能的整个学习决策过程均不可解释、不可理解¹¹。算法黑箱会激化自反性安全风险。自反性安全风险是指将现代化“自我对抗”的性质引入安全领域后催生的风险形态¹²，在第二现代性的风险社会中，高度发展的技术导致的问题可能演变为大规模不可控的人为灾难。

“黑箱”人工智能应用于医疗领域，由于运作过程的不透明，决策的不可解释性，可能会得出与患者实际病情严重不符的诊断结果，但无法解释该决策是如何作出的，医务人员也无法理解该诊断作出的过程。即使不存在算法黑箱，生产者也会因为考虑商业秘密保护等缺乏公开的动力。从传统的医疗诊断过程来看，医务人员作为专业中间人应发现人工智能辅助医疗诊断系统作出诊断决策的过程或结果存在错误，此时可切断系统产品缺陷与人身损害之间的因果关系。但算法不可解释，医生的审查能力被架空，无法履行审查义务，人工智能辅助医疗诊断系统本质上构成了对医生诊断的替代。此时，具备医学专业知识的医务人员，因缺乏对人工智能决策逻辑的认识而处于不利地位，生产者不能因算法“不可解释”就逃避责任，医务人员和医疗机构不能承担自身义务以外的过重的责任。因此，必须参照产品责任由生产者承担责任。

4. 责任分配与归责路径

(一) 构建人工智能辅助医疗诊断系统的产品缺陷标准

⁹参见张新宝：《侵权责任法》(第5版)，中国人民大学出版社2020年版，第213-214页。

¹⁰汪庆华：《算法透明的多重维度和算法问责》，载《比较法研究》2020年11月，第165页。

¹¹参见刘成杰：《类型化视角下人工智能侵权责任的归结思路》，载《法律适用》2025年第3期，第120-122页。

¹²参见陈嘉鑫、李宝诚：《风险社会理论视域下生成式人工智能安全风险检视与应对》，载《情报杂志》2025年1月，第130页。

从传统理论出发，重构设计缺陷、警示缺陷和跟踪缺陷以判断人工智能辅助医疗诊断过程中是否存在产品缺陷。人工智能辅助医疗诊断系统在设计过程中应当追求的是更安全、更透明的数据模型，使诊断决策结果足够令人信赖，应注重提升诊断的准确性。因此设计缺陷主要考虑算法模型本身是否存在缺陷，通过试用该系统并进行对比实验，判断是否存在前文所提及的性别、年龄、地域等方面的算法偏见，该算法模型是否能够保持理性和中立，数据来源是否权威，数量是否足够庞大，生产者是否对其进行了足够的训练。美国 FDA 对 SaMD¹³的监管中存在“实质等同”标准，通过 510(k)程序证明新产品与已获批器械在安全性和有效性上“实质等同”即可¹⁴。在判断人工智能辅助医疗诊断系统是否存在算法偏见的过程中可以参照这一标准，对比已经获批使用其他系统，判断该系统是否具有同等的安全性和有效性，得出的结论是否存在不合理的偏见和危险。如果人工智能辅助医疗诊断系统存在算法偏差，模型不够强大，诊断的准确率较低，相较于其他已经获批使用的系统存在偏差和更高的风险，那么应当认定存在设计缺陷。

警示缺陷主要考虑生产者是否对系统作出的决策的准确性进行了说明，是否在合理范围内披露了相关信息。欧盟《人工智能法案》将人工智能的风险划分为不可接受的风险、高风险、特定透明度风险和最小风险，针对不同风险适用不同规则¹⁵。美国食品药品监督管理局(FDA)主要通过分类审查实现对 SaMD 的监管。低风险软件无需 FAD 前市场审批，采用注册备案制度，但要求算法的透明度和可解释性，生产者要确保决策过程是透明的，帮助医务人员理解人工智能的判断依据。生产者还需要提供数据验证和算法测试，确保其适应不同的临床环境，符合医疗标准¹⁶。根据以上标准，对于某些可信度低，错误率较高的高风险人工智能辅助医疗诊断系统，如果没有进行提示，或者没有对系统的运行过程、决策依据进行说明，那么认为存在警示缺陷。

跟踪缺陷的认定可以从调整认定时间上入手，将“投入市场”这一静态概念转变为动态过程。美国 SaMD 监管框架下对人工智能医疗软件实施动态监管，以应对人工智能辅助医疗诊断系统的自主学习和动态调整特性，要求生产者确保决策过程透明，提供数据验证和算法测试。同时，生产者必须实施有效的风险管理流程¹⁷。认定是否存在跟踪缺陷可以由生产者证明其内部存在有效的风险管理流程，在人工智能辅助医疗诊断系统投入流通后，是否通过系统运行记录分析系统运行状况，判断是否存在错误偏差。在发现错误后，是否及时警告、及时调整修复系统都是重要的考量因素。若生产者尽到了跟踪义务，但依然未能及时捕捉错误，此时可赋予生产者抗辩权，但必须举证证明自己已经履行了跟踪义务。

(二) 确立双重因果关系认定标准

首先需要明确，即使适用产品责任也需要证明产品缺陷和损害结果之间存在因果关系。

其次，采用双重因果关系进行认定。第一重属于事实因果关系。这一层主要判断人工智能辅助医疗诊断系统输出的结果是否存在逻辑错误或数据偏差。通过日志回溯记录和分析人工智能辅助医疗诊断系统作出决策的过程，判断其客观上是否偏离了医学常识或算法逻辑。如果存在违背医学常识的错误或偏离算法逻辑的情况，那么存在因果关系，相反，则应当认为不存在因果关系。第二重属于法律因果关系。这一层主要判断医生的介入行为是否足以中断产品缺陷与损害结果之间的因果关系。判断关键在于医务人员采信系统的诊断决策，是否是在当时的医疗水平下尽到合理注意义务的选择，如果医务人员未尽到

¹³SaMD (Software as a Medical Device): 是指不依赖于硬件的，单独用于医疗目的的软件。在 AI/ML 医疗软件的情况下，SaMD 包括那些通过 AI/ML 算法来提供诊断、预测、辅助决策等医疗功能的软件工具。

¹⁴参见《美国 FDA 关于人工智能(AI)/机器学习(ML)医疗软件监管的要求-基于 SaMD 框架》，载微信公众号“质量从业者”2025 年 2 月 6 日。<https://mp.weixin.qq.com/s/8NOJqJ4dcVQ2yHzS9PrR8A>

¹⁵参见丁晓东：欧盟《人工智能法案》中文翻译版(上)，载微信公众号“网络西东”2024 年 8 月 1 日。

<https://mp.weixin.qq.com/s/hE8hjWAvGL6vO87nvYjaCA>

¹⁶同 14。

¹⁷同 14。

合理注意义务，未能识别出系统错误的决策结果，则中断了产品缺陷与损害之间的因果关系，在法律上，人工智能辅助医疗诊断系统的产品缺陷和损害结果之间就不存在因果关系。反之，如果医务人员已经尽到合理注意义务，仍未能识别出错误，那么人工智能辅助医疗诊断系统的产品缺陷和损害结果之间就存在法律上的因果关系。

最后，在双重因果关系认定的模式下应合理分配举证责任。欧盟《缺陷产品责任指令》为了实现风险的公平分担，采取因果推定模式，受损害人应当就产品存在缺陷和受到损害之间存在因果关系承担举证责任。但由于受损害人处于劣势地位，因此指令针对原告获取诉讼证据提供了便利。受损害方可以提出对方披露证据的申请，法院审查申请的必要性和范围的适当性，避免收集无关信息，保护机密信息。若被告未遵守信息披露义务，则推定产品存在缺陷，如果已经确定产品确实存在缺陷，且存在同类案例证明损害通常由产品缺陷导致，那么即可推定因果关系存在。总体而言，如果受损害人在证明缺陷上存在困难，那么只要求原告证明产品很可能存在缺陷；如果受损害人在证明因果关系上存在困难，那么只要求受损害人证明产品缺陷很可能是造成损害的原因¹⁸。由于算法的不可解释性，受损害方很难获得系统的日志记录，因此在判断技术因果关系的过程中受损害方有权向法院申请人工智能辅助医疗诊断系统的生产者披露相关信息，法院决定系统生产者一方合理的披露范围，如果生产者无正当理由拒绝披露，那么推定事实因果关系成立。

(三) 构建类型化人工智能辅助医疗诊断系统侵权的责任承担方式

人工智能辅助医疗诊断系统侵权致损根据造成损害发生的主体和原因不同可以类型化为隐秘型缺陷致损、常识型缺陷致损和人机耦合型缺陷致损。

首先，针对人工智能辅助医疗诊断系统隐秘型缺陷致损的责任承担方式。隐秘型缺陷主要指算法内部运行过程中出现错误，但得出的结果符合医学常识和医学逻辑，具有临床上的欺骗性，最终在应用于人身的过程中产生了损害。由于算法“黑箱”的不可解释性，医务人员无法识别系统决策存在错误，此时医务人员的诊断地位被实质性替代，系统缺陷与损害结果之间不仅存在事实上的因果关系也存在法律上的因果关系，医务人员的行为未能中断产品缺陷与损害结果之间的因果关系，生产者需要在产品责任框架下承担全部责任。这样的规则能够防止医疗机构过度承担风险，生产者逃避责任。同时，这一规则没有苛求医务人员和医疗机构承担过重责任，因为在这种情况下，医务人员的专业知识不足以解释该系统产生的决策结果，在医务人员尽到合理注意义务的情况下，不得使其承担过重义务。这样也有利于人工智能辅助医疗系统在医疗诊断过程中的推广和适用。

其次，针对人工智能辅助医疗诊断系统常识型缺陷致损的责任承担方式。常识型缺陷是指人工智能辅助医疗系统生成的诊断决策明显违背医学常识，基础数据存在明显的错误或矛盾。医务人员在采用人工智能辅助医疗诊断系统的决策时必须尽到合理注意义务，采取审慎态度。因为目前各国均将医疗人工智能定位为辅助诊疗设备，仅为医务人员提供有限的临床决策参考，主导权仍掌握在医务人员手中([1], p. 48)。若发生常识型缺陷，生产者明确设置了警示，提醒使用者该系统在诊断决策过程中容易出现算法错误，误诊率高，但医务人员依然未尽到合理注意义务，未排除错误，盲目相信决策结果造成患者人身损害，此时产品缺陷和损害结果之间只有事实上的因果关系，医务人员的行为中断了产品缺陷和损害结果之间法律上的因果关系，应由医疗结构独立承担责任。

最后，针对人机耦合型致损的责任承担方式。当人工智能辅助医疗诊断系统和医务人员诊疗行为同时存在过错的情况下，要根据原因力的大小确定责任分担规则。生产者如果在系统使用说明中明确提示了该系统“误诊率高”“高风险”，且履行了信息披露义务，帮助医务人员了解了系统的决策运行模式，

¹⁸参见王竹、张皓翔译：欧盟 2024 年新《缺陷产品责任指令》，载微信公众号“民商法学 LexM”2025 年 8 月 1 日。
<https://mp.weixin.qq.com/s/YEMVLG-StPKjg58LNuw-KO>

但医务人员忽略了以上问题，盲目相信系统的决策结果，导致患者受损，此时医生在损害发生的过程中作用力更大，应当根据医疗损害责任由医疗机构承担主要侵权责任，生产者根据产品责任承担补充责任。如果人工智能辅助医疗诊断的生产者在推广系统的过程中涉及虚假宣传，过度夸大系统的功能和准确性，甚至伪造诊断结果正确率，导致医务人员产生误信最终导致了损害结果，此时应认定生产者的作用力更大，应根据产品责任由生产者承担主要责任，由医疗机构依据医疗损害责任承担补充责任。

5. 总结

人工智能的发展给人类社会带来了诸多便利，改变了人类看待世界的方式，提高了工作效率。同时，算法“黑箱”，人工智能辅助医疗诊断系统决策过程的不可解释性，决定了人类将失去对人工智能决策过程的控制，从而导致人工智能辅助医疗诊断产生不可控的风险，并影响整个人类社会。因此，当人工智能应用到关系公民生命健康的最关键的医疗领域，必须采取更加审慎的态度。目前人工智能与医疗的结合已经成为趋势，但相应的规则体系并未建立，在利用技术，使其更好更准确地服务于医疗的过程中还要注意规则的更新和构建。如果人工智能辅助医疗诊断的过程中发生侵权，采取何种归责方式，成为关键问题。虽然人工智能辅助医疗诊断系统对于传统的侵权责任认定有所冲击，但为了维护法律的稳定性，依然要基于现存法律框架进行规则的构建。从理论上分析，人工智能辅助医疗诊断适用产品责任存在基础。同时应当看到其特殊性，对产品缺陷和因果关系的判断标准进行重构。对于设计缺陷的判断主要采取“实质等同”标准，通过与其他已经获批投入使用的人工智能辅助医疗诊断系统进行对比，判断该系统是否存在明显的算法偏见，数据来源是否权威，诊断的准确率是否能够达到相应水平。警示缺陷主要考虑生产者是否对系统作出的决策的准确性进行了说明，是否在合理范围内履行了披露义务。跟踪缺陷的重构主要通过将“投入市场”这一静态概念转变为动态过程来实现，要求生产者在产品投入流通前就设计有效的风险管理流程，并在人工智能辅助医疗诊断的过程中实时检测，一旦发现系统漏洞应当及时修复并告知使用者。采用双重因果关系标准判断人工智能辅助医疗诊断系统的缺陷与损害结果之间是否存在因果关系。第一重事实因果关系，如果人工智能辅助医疗诊断系统输出的结果存在逻辑错误或数据偏差，得出违背医学常识或算法逻辑的结果，那么应认定产品缺陷和损害结果之间存在因果关系。第二重法律因果关系，人工智能辅助医疗诊断系统的辅助地位决定了诊断过程中存在医务人员的介入行为，因此判断人工智能辅助医疗诊断系统缺陷与损害结果之间是否存在法律上的因果关系还需要考虑医务人员的诊断行为是否中断了人工智能辅助医疗诊断系统缺陷与损害结果之间的因果关系。如果医务人员未尽到合理注意义务，未审查系统决策的准确性，能够发现错误而未发现，盲目采信系统结论，此时应当认为医务人员的诊断行为中断了人工智能辅助医疗诊断系统缺陷与损害结果之间的法律因果关系。因此，基本的责任承担模式按照造成损害发生的主体和原因不同可分为隐秘型缺陷致损、常识型缺陷致损和人机耦合型缺陷致损。如果人工智能辅助医疗诊断系统得出的诊断决策具有临床欺骗性，医务人员尽到合理注意义务也无法识别错误的情况下，属于隐秘型缺陷造成损害结果，由生产者承担产品责任；如果人工智能辅助医疗诊断系统的诊断决策的错误属于常识型错误，医务人员能够根据自身的医学知识作出判断却未能排除错误的情形，属于常识型缺陷造成损害结果，因其本质上是医务人员未尽到合理注意义务，由医疗机构承担责任。如果是人工智能辅助医疗诊断系统和医务人员共同侵权造成损害后果发生，那么根据二者的原因力分别确定责任，原因力大小的判断主要依据生产者是否尽到了说明义务，是否存在虚假宣传等使医务人员过度信赖系统决策的准确性，医务人员是否尽到了合理注意义务等因素。

目前我国《人工智能法（学者建议稿）》第四章明确了开发者和提供者的义务。我们可以借鉴欧盟《人工智能法案》《缺陷产品责任指令》以及美国FDA“SaMD监管框架”，引入合规免责和持续学习监管机制。形成我国人工智能辅助医疗诊断的法律规则体系和监管框架，防范于未然，更好地维护患者合法权益。

参考文献

- [1] 李润生. 人机协作的法律形态及过失认定——人类医生与医疗 AI 的共处之道[J]. 北京航空航天大学学报(社会科学版), 2026, 39(1): 48-61.
- [2] 杨立新. 人工智能产品责任的功能及规则调整[J]. 数字法治, 2023(4): 27-39.
- [3] 张新宝. 中国民法典释评·侵权责任编[M]. 北京: 中国人民大学出版社, 2020.
- [4] 郑志峰. 诊疗人工智能的医疗损害责任[J]. 中国法学, 2023(1): 203-221.
- [5] 林勇涛. 医疗人工智能、医疗决策与医疗说明义务——德国经验与中国借鉴[J]. 德国研究, 2025, 40(4): 134-159, 164.
- [6] 张新宝. 侵权责任法[M]. 第 5 版. 北京: 中国人民大学出版社, 2020.
- [7] 王莹. 算法侵害责任框架刍议[J]. 中国法学, 2022(3): 165-184.
- [8] 郑飞. 论人工智能法体系化的理论基础[J]. 中国政法大学学报, 2024(6): 236-250.
- [9] 陈嘉鑫, 李宝诚. 风险社会理论视域下生成式人工智能安全风险检视与应对[J]. 情报杂志, 2025, 44(1): 128-135, 171.