AI and Increased Medical Liability?

Extended Abstract

Alessandro Tacconelli

September 2023

1 Introduction

Daily, we are exposed to an increasing trend of incorporating the use of Artificial Intelligence (AI) systems in human-decision making processes, in several fields. Among those fields, the medical sector surely represents a peculiar one. Available findings show how AI medical technology is increasingly being adopted to provide patients with personalised medical treatments (Topol, 2019).

Given its high rate of regulatory restrictions, on the one hand, and the particular nature of the main good of the sector (that is, the human life), on the other one, scholars have been discussing about the use of AI and its repercussion on medical liability, yet without being able to reach an uniform opinion on the point. In Price and Cohen (2019), the authors posited that, given the applicable US medical liability regulations, legal systems may actually undermine the use of AI, thus depriving both physicians and patients of the potential value of those tools. Such a claim was challenged by the experimental work in Tobia et al. (2021). There, through the use of a vignette study, the authors tested the perceived medical liability by laypeople when exposed to different scenarios in which a hypothetical physician decided either to deviate (or not) from standard-care recommendations and to follow (or not) AI recommendations. Their results seemed to draw towards a radically opposite conclusion than the one of Price and Cohen (2019). Experimentally, the authors showed how a tort law system like the one of the US is unlikely to undermine the use of AI precision medicine tools and may even encourage the use of these tools. If those results shed light on the topic, they still are limited to a very peculiar cultural and legal circumstance: that is, the common-law system of the US, in which cases of medical malpractice are judged by laypeople as jurors. In other words, a doubt that is yet to be solved is whether those findings would hold even in civil-law jurisdiction.

Thus, the purpose of the present study is twofold: (i) first, we employ a replication of the study in Tobia et al. (2021), to test whether their original results hold – to do so, we recruit a nationally repre-

sentative sample from the US and we employ the same experimental manipulations performed in their research; and (*ii*) second, we further expand this investigation, by accessing a nationally representative sample from Germany and a peculiar 'professional' sample of physicians. For our extension, in other words, not only do we intend to check whether cultural factors could determine different results (US vs German 'laypeople' samples), but also investigate how professionalisation could impact the resulting liability assessment (German physicians' sample). The latter shall be of the utmost importance to disentangle the legally relevant differences (if any) that subsist between an emblematic common-law system like the US, and an emblematic civil-law system like Germany. In fact, as in many other civil-law jurisdictions, in Germany, a judgement on medical liability would not be taken by jurors, but rather professional judges. Those though, given the technical nature of the matter at hand, would defer to experts for their evaluation of each case and the subsequent decision on medical liability.

2 The Experiment

We collected a total of N' = 2,753 observations, divided as follows:

1. *n' = 1,188* observations for our US representative sample ('*US Laypeople*'),
2. *n' = 1,367* observations for our German representative sample ('*German Laypeople*'), and
3. *n' = 198* observations for our physicians' sample ('*Physicians*').

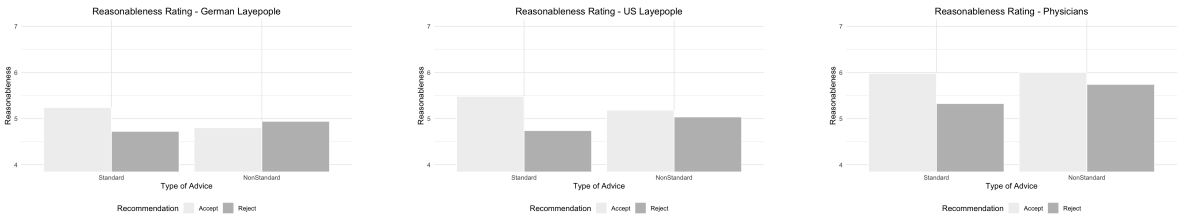
We employed a 2 x 2 between-participant factorial design, resulting in a total of four experimental groups. Each participant was assigned to one of those conditions, thus reading one and only one of the vignettes in which we elicited our four different experimental manipulations. Across the experimental groups, we varied two dimensions: (*i*) the nature of the decision suggested by the AI tool (*standard* vs *nonstandard* medical treatment), and (*ii*) the decision taken by the fictional physician described in our vignettes (*accept* vs *reject* the AI treatment recommendations). Specifically:

- in the ***Standard-Accept*** condition, after having received an AI recommendation to proceed with the *standard*-care procedure, the fictional physician decides to *accept* such advice;
- in the ***Standard-Reject*** condition, after having received the same AI recommendation to proceed with the *standard*-care procedure, the fictional physician decided to *reject* such advice;
- in the ***Non-Standard-Accept*** condition, after having received an AI recommendation to proceed with a *non-standard* care procedure this time, the fictional physician decides to *accept* the advice; and

- in the ***Non-Standard-Reject*** condition, after having received again an AI recommendation to proceed with a *non-standard* care procedure, the fictional physician decides to *reject* the advice.

In all four conditions, instead, we keep the type of harm resulting from the fictional physician's decision constant. We use reasonableness rating as the main dependent variable and a proxy of the perceived medical liability of the fictional physician described in each scenario (*i.e.*, the lower the declared perceived reasonableness of the physician's decision, the higher the likelihood of the physician to be held liable for medical malpractice).

3 Results



(a) German Laypeople (b) US laypeople (c) Physicians

Figure 1: Reasonableness Ratings per Sample

In Fig. 1 we graphically show the declared reasonableness rating in each of the four experimental groups resulting from our factorial for the three samples recruited.

We analyze our data through a robust set of parametric and non-parametric tests, employing multiple measurements and controlling for confounds. Results provide compelling evidence showing how:

1. the pattern observed in Tobia et al. (2021) replicates when analysing the results collected in the US sample – also in our case, when the medical advice provided is *standard*, a shielding effect subsists when physician decides to accept (rather than reject) the AI recommendation. Such an effect disappears, though, when the provided advice is a *nonstandard* one, so the risk of liability is not reduced by accepting or rejecting the advice;

2. interestingly, an almost identical pattern is also observable in a culturally different sample – that is, the recruited nationally representative sample of German 'laypeople,' although overall apparently more skeptical towards medical professionals (reasonableness rating being overall low compared to US laypeople across the 4 conditions), still behave and react in a way comparable to the one emerging from the US; and

3. even more interestingly, the same pattern observed for the US sample emerges among physicians, with an overall trend of the latter to be way more lenient (that is, a higher tendency not to hold the hypothetical physicians liable) in their judgement of medical liability.

4 Discussion

As in Tobia et al. (2021), our findings speak to the concerns expressed in Price and Cohen (2019). Again, by replicating the authors' results, our study provides yet another evidence of the fact that, even when varying the sample recruited, still a tort system on medical malpractice like the one implemented in several jurisdictions would not undermine or impede the medical use of AI. And this is true even when shifting from a common-law legal system (where medical liability assessment is left to jurors, that is, laypeople) to a civil-law system (in which, ultimately, the decision pertains to the evaluations of medical experts). Also in the latter systems, in other words, the concerns vocalised in Price and Cohen (2019) regarding the potential 'negative' effects of tort-law systems on the implementation of AI infrastructure in the medical sector do not seem to translate into concrete risks.

Taken together, our results provide also interesting information on what a potential modification of tort-law legislation in a civil-law system like the one in Germany could entail. In this 'projection' endeavour, from the findings collected, we can probably ascertain that the reduced pressure for malpractice litigation registered in Germany compared to the US (Sawyer and Sawyer, 1993) would be hampered by the (very) hypothetical decision of stripping medical experts of their decisive role in medical liability cases and assigning this role to jurors chosen among laypeople.