Web3 Journal: ML in Health Science

ISSN 2942-8726

2024. Vol. 1(1)

Original Article

TxnHash: 0x30c063255095ee96b460e13bf9bc13948c83dc9f288ac4ccfa91ff8b4fea644d

DOI: 10.62487/vc54ms96

Confounders in Predictive Medical Models: Roles of Nationality and Immigrant Status

Yury Rusinovich[1], Volha Rusinovich[1]

Abstract.

Aim: The aim of this study was to assess the opinion of natural science specialists on the latest recommendations of official regulators regarding the prevention of causes of social disparities in artificial intelligence (AI) and machine learning (ML) models. **Materials and Methods:** An anonymous online survey was conducted using the Telegram platform, where participants were asked a single question: "Is the inclusion of predictors such as “nationality” and “immigrant status” in AI and ML medical models ethical and consistent with contemporary scientific standards?" Respondents were provided with two response options: "Yes" or "No." The survey was specifically targeted at international groups, focusing primarily on English and Russian-speaking clinicians and scientific researchers. **Results:** 180 unique individuals participated in the survey. The results revealed that one-third of the respondents (60 individuals) agreed that including predictors such as nationality and immigration status is inappropriate in the current ML and AI models. **Conclusion:** In conclusion, the fact that only one-third of respondents disagree with categorizing patients based on nationality is at odds with the standards set by official regulators. This discrepancy underscores the need for educational programs aimed at sensitizing the scientific community to prioritize biological predictors over data documented in passports or identity cards.

Keywords: Human-Centered AI, Human-AI Collaboration, Healthcare Survey

Background:

Confounders

A confounder is a variable that influences both the outcome and the predictor simultaneously[1]. A clear example of confounding in medicine is the mistaken belief that lower social competence is a predictor of schizophrenia[2,3]. In fact, the association may be inverse; schizophrenia is often more prevalent in populations with lower socioeconomic status[3]. This is not necessarily because poverty induces the disorder, but rather because schizophrenia itself can diminish social and vocational competencies, influencing the socioeconomic status of affected individuals.[3]

This research will examine confounders that not only introduce bias, as previously mentioned, but also raise ethical concerns, such as social disparities[4–9]. Examples of these “dirty” confounders include race, nationality, immigrant status, religion, and socioeconomic status.

[1]ML in Health Science, Leipzig, Germany

Corresponding author: Yury Rusinovich

Email: info@mlhs.ink

The scientific validity and statistical significance derived from these confounders are questionable because they can be influenced by a multitude of other factors associated with the identity of a particular population, such as environmental factors like climate, solar or geomagnetic activity, air pollution, natural radiation exposure, the region's gravitational field, etc[10–26]. Environmental factors can significantly impact our genome,[27] and disease patterns, more so than the superficial markers indicated on identity cards.

Aim:

The aim of this study was to assess the opinion of natural science practitioners on the latest recommendations of official regulators regarding the prevention of causes of social disparities in artificial intelligence (AI) and machine learning (ML) models[8,9], particularly those that could arise from nationality and immigrant status within healthcare settings.

Material and Methods:

An anonymous online survey was conducted using the Telegram platform, where participants were asked a single question: "Is the inclusion of predictors such as 'Nationality', 'Place of Birth', and 'Immigrant Status' in AI and ML medical models ethical and consistent with contemporary scientific standards?" Respondents were provided with two response options: "Yes" or "No."

The survey was specifically targeted at international groups, focusing primarily on Russian and English-speaking clinicians and scientific researchers.

Statistics:

The collected data were analyzed using descriptive statistics to summarize and interpret the responses.

Results:

The survey was conducted in January 2024 with 180 unique and verified individuals participating. The results revealed that 1/3 of the respondents (60 individuals) agreed that including predictors such as nationality and immigration status is inappropriate in the current ML and AI models.

The results of this survey are openly accessible on the official Telegram Channel of the Web3 Society: ML in Health Science, which can be visited at: https://t.me/MLinHS

Table 1 and **Figure 1** summarize the survey results:

Variable	Respondents
Yes	120
No	60
Total	180

Table 1: Survey Results: Is the inclusion of predictors such as Nationality, Place of Birth, and Immigrant Status in AI and ML medical models ethical and consistent with contemporary scientific standards?

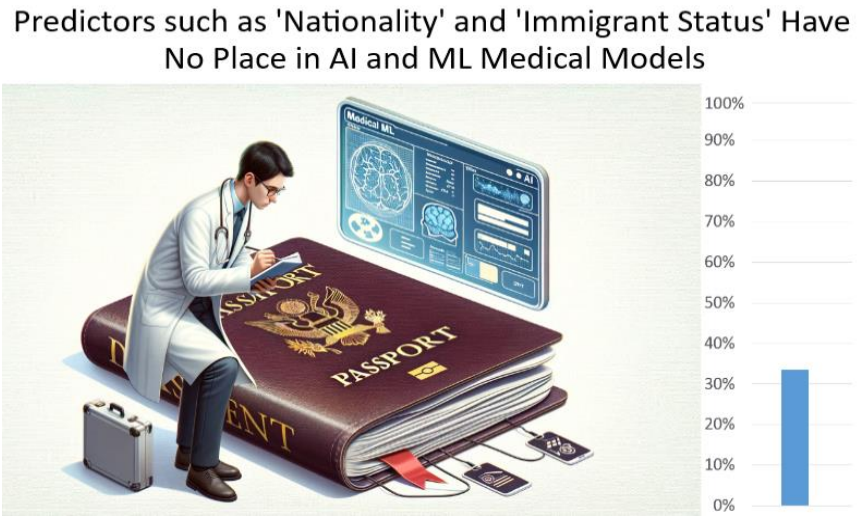

Figure 1: Survey Results. DALL°E

Discussion:

Practical standpoint

Our research indicates that a mere one-third of healthcare practitioners and researchers participating in the survey disagree with the categorization of individuals based on nationality in scientific and medical contexts. This finding emphasizes a concerning disconnect with the norms set by official regulators[8,9] and underscores the need for further research and possibly educational programs showcasing the prioritization of biological, natural, and environmental patterns over the data written in identity cards.

Limitations

The study's reliance on an anonymous methodology may introduce some uncertainty regarding the purity of the cohort and participant characteristics, which could impact the robustness and generalizability of the findings.

Conclusion

In conclusion, the fact that only one-third of respondents disagree with categorizing patients based on nationality is at odds with the standards set by official regulators. This discrepancy underscores the need for educational programs aimed at sensitizing the scientific community to prioritize biological predictors over data documented in passports or identity cards.

Conflict of Interest: YR and VR state that no conflict of interest exists.

Authorship: YR: Concept, data analysis, original draft, survey. YR, VR: Review and editing.

References

1 **Wikipedia. Confounding. https://en.wikipedia.org/wiki/Confounding#cite_note-Pearl_2009-1. Accessed February 1, 2024.**

2 **Cooper B. Schizophrenia, social class and immigrant status: the epidemiological evidence. *Epidemiol Psichiatr Soc*. 2005;14(3):137-144. doi:10.1017/s1121189x00006382.**

3 **Andrade C. Confounding. *Indian J Psychiatry*. 2007;49(2):129-131. doi:10.4103/0019-5545.33263.**

4 **Huang J, Galal G, Etemadi M, Vaidyanathan M. Evaluation and Mitigation of Racial Bias in Clinical Machine Learning Models: Scoping Review. *JMIR Med Inform*. 2022;10(5):e36388. doi:10.2196/36388.**

5 **Tipton K, Leas B, Flores E, et al. Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare. *Effective Health Care Program*. https://effectivehealthcare.ahrq.gov/products/racial-disparities-health-healthcare/research#field_report_title_3. Published December 8, 2023. Accessed February 1, 2024.**

6 **Allen A, Mataraso S, Siefkas A, et al. A Racially Unbiased, Machine Learning Approach to Prediction of Mortality: Algorithm Development Study. *JMIR Public Health Surveill*. 2020;6(4):e22400. doi:10.2196/22400.**

7 **Cary MP, Zink A, Wei S, et al. Mitigating Racial And Ethnic Bias And Advancing Health Equity In Clinical Algorithms: A Scoping Review. *Health Aff (Millwood)*. 2023;42(10):1359-1368. doi:10.1377/hlthaff.2023.00553.**

8 **European Parliamentary Research Service. Artificial intelligence act. https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf.**

9 **World Health Organization, ed. *Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models*; 2024.**

10 **Alabdulgader A, McCraty R, Atkinson M, et al. Long-Term Study of Heart Rate Variability Responses to Changes in the Solar and Geomagnetic Environment. *Sci Rep*. 2018;8(1):2663. doi:10.1038/s41598-018-20932-x.**

11 Davis GE, Lowell WE. Solar cycles and their relationship to human disease and adaptability. *Med Hypotheses*. 2006;67(3):447-461. doi:10.1016/j.mehy.2006.03.011.

12 Anand K, Vieira CLZ, Garshick E, et al. Solar and geomagnetic activity reduces pulmonary function and enhances particulate pollution effects. *Sci Total Environ*. 2022;838(Pt 3):156434. doi:10.1016/j.scitotenv.2022.156434.

13 Zilli Vieira CL, Alvares D, Blomberg A, et al. Geomagnetic disturbances driven by solar activity enhance total and cardiovascular mortality risk in 263 U.S. cities. *Environ Health*. 2019;18(1):83. doi:10.1186/s12940-019-0516-0.

14 Zakharov IG, Tyrnov OF. The effect of solar activity on ill and healthy people under conditions of nervous (correction of neurous) and emotional stresses. *Adv Space Res*. 2001;28(4):685-690. doi:10.1016/s0273-1177(01)00379-9.

15 David E, Wolfson M, Fraifeld VE. Background radiation impacts human longevity and cancer mortality: reconsidering the linear no-threshold paradigm. *Biogerontology*. 2021;22(2):189-195. doi:10.1007/s10522-020-09909-4.

16 Zlobina A, Farkhutdinov I, Carvalho FP, et al. Impact of Environmental Radiation on the Incidence of Cancer and Birth Defects in Regions with High Natural Radioactivity. *Int J Environ Res Public Health*. 2022;19(14). doi:10.3390/ijerph19148643.

17 Autsavapromporn N, Krandrod C, Klunklin P, et al. Health Effects of Natural Environmental Radiation during Burning Season in Chiang Mai, Thailand. *Life (Basel)*. 2022;12(6). doi:10.3390/life12060853.

18 World Health Organization. Climate change. https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health. Accessed May 8, 2024.

19 Patz JA, Frumkin H, Holloway T, Vimont DJ, Haines A. Climate change: challenges and opportunities for global health. *JAMA*. 2014;312(15):1565-1580. doi:10.1001/jama.2014.13186.

20 Ohashi Y, Ihara T, Oka K, Takane Y, Kikegawa Y. Machine learning analysis and risk prediction of weather-sensitive mortality related to cardiovascular disease during summer in Tokyo, Japan. *Sci Rep*. 2023;13(1):17020. doi:10.1038/s41598-023-44181-9.

21 Semenza JC, Paz S. Climate change and infectious disease in Europe: Impact, projection and adaptation. *Lancet Reg Health Eur*. 2021;9:100230. doi:10.1016/j.lanepe.2021.100230.

22 Bartlett VL, Doernberg H, Mooghali M, et al. Published research on the human health implications of climate change between 2012 and 2021: cross sectional study. *BMJ Med*. 2024;3(1):e000627. doi:10.1136/bmjmed-2023-000627.

23 Rusinovich Y, Rusinovich V. Earth's gravity field and prevalence of varicose veins and chronic venous disease: Systematic review. *Phlebology*. 2022;37(7):486-495. doi:10.1177/02683555221090054.

24 Berrang-Ford L, Sietsma AJ, Callaghan M, et al. Systematic mapping of global research on climate and health: a machine learning review. *Lancet Planet Health*. 2021;5(8):e514-e525. doi:10.1016/S2542-5196(21)00179-0.

25 Schiff JE, Vieira CLZ, Garshick E, et al. The role of solar and geomagnetic activity in endothelial activation and inflammation in the NAS cohort. *PLoS One*. 2022;17(7):e0268700. doi:10.1371/journal.pone.0268700.

26 Missimer TM, Teaf C, Maliva RG, Danley-Thomson A, Covert D, Hegy M. Natural Radiation in the Rocks, Soils, and Groundwater of Southern Florida with a Discussion on Potential Health Impacts. *Int J Environ Res Public Health*. 2019;16(10). doi:10.3390/ijerph16101793.

27 Overbey EG, Da Silveira WA, Stanbouly S, et al. Space-flight influences gene expression, photoreceptor integrity, and oxidative stress-related damage in the murine retina. *Sci Rep*. 2019;9(1):13304. doi:10.1038/s41598-019-49453-x.