Artificial Intelligence: Implications for Health Professions Education

OHSU Educators' Collaborative – October 18, 2023

William Hersh, MD
Professor
Department of Medical Informatics & Clinical Epidemiology
School of Medicine
Oregon Health & Science University
Portland, OR, USA
https://www.ohsu.edu/informatics
Email: hersh@ohsu.edu
Web: http://www.billhersh.info/
Blog: https://informaticsprofessor.blogspot.com/
Twitter: @williamhersh

References

Ali, S.R., Dobbs, T.D., Hutchings, H.A., Whitaker, I.S., 2023. Using ChatGPT to write patient clinic letters. Lancet Digit Health 5, e179–e181. https://doi.org/10.1016/S2589-7500(23)00048-1

Al-Zaiti, S.S., Martin-Gill, C., Zègre-Hemsey, J.K., Bouzid, Z., Faramand, Z., Alrawashdeh, M.O., Gregg, R.E., Helman, S., Riek, N.T., Kraevsky-Phillips, K., Clermont, G., Akcakaya, M., Sereika, S.M., Van Dam, P., Smith, S.W., Birnbaum, Y., Saba, S., Sejdic, E., Callaway, C.W., 2023. Machine learning for ECG diagnosis and risk stratification of occlusion myocardial infarction. Nat Med. https://doi.org/10.1038/s41591-023-02396-3

Antaki, F., Touma, S., Milad, D., El-Khoury, J., Duval, R., 2023. Evaluating the Performance of ChatGPT in Ophthalmology: An Analysis of Its Successes and Shortcomings. Ophthalmol Sci 3, 100324. https://doi.org/10.1016/j.xops.2023.100324

Attia, Z.I., Friedman, P.A., Noseworthy, P.A., Lopez-Jimenez, F., Ladewig, D.J., Satam, G., Pellikka, P.A., Munger, T.M., Asirvatham, S.J., Scott, C.G., Carter, R.E., Kapa, S., 2019. Age and Sex Estimation Using Artificial Intelligence From Standard 12-Lead ECGs. Circ Arrhythm Electrophysiol 12, e007284. https://doi.org/10.1161/CIRCEP.119.007284

Ayers, J.W., Poliak, A., Dredze, M., Leas, E.C., Zhu, Z., Kelley, J.B., Faix, D.J., Goodman, A.M., Longhurst, C.A., Hogarth, M., Smith, D.M., 2023a. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. JAMA Intern Med 183, 589–596. https://doi.org/10.1001/jamainternmed.2023.1838

Ayers, J.W., Zhu, Z., Poliak, A., Leas, E.C., Dredze, M., Hogarth, M., Smith, D.M., 2023b. Evaluating Artificial Intelligence Responses to Public Health Questions. JAMA Netw Open 6, e2317517. https://doi.org/10.1001/jamanetworkopen.2023.17517

Bair, H., Djulbegovic, M., Taylor Gonzalez, D., 2023. Opinion | We Use ChatGPT in Medical Practice -- And You Can Too [WWW Document]. URL https://www.medpagetoday.com/opinion/second-opinions/106085 (accessed 8.29.23).

Baker, M.N., Burruss, C.P., Wilson, C.L., Baker, M.N., Burruss, C.P., M.d, C.L.W., 2023. ChatGPT: A Supplemental Tool for Efficiency and Improved Communication in Rural Dermatology. Cureus 15. https://doi.org/10.7759/cureus.43812

Barnett, G.O., Cimino, J.J., Hupp, J.A., Hoffer, E.P., 1987. DXplain. An evolving diagnostic decision-support system. JAMA 258, 67–74. https://doi.org/10.1001/jama.258.1.67

Beam, K., Sharma, P., Kumar, B., Wang, C., Brodsky, D., Martin, C.R., Beam, A., 2023. Performance of a Large Language Model on Practice Questions for the Neonatal Board Examination. JAMA Pediatr e232373. https://doi.org/10.1001/jamapediatrics.2023.2373

Benoit, J.R.A., 2023. ChatGPT for Clinical Vignette Generation, Revision, and Evaluation. https://doi.org/10.1101/2023.02.04.23285478

Boscardin, C.K., Gin, B., Golde, P.B., Hauer, K.E., 2023. ChatGPT and Generative Artificial Intelligence for Medical Education: Potential Impact and Opportunity. Acad Med. https://doi.org/10.1097/ACM.0000000000005439

Brin, D., Sorin, V., Vaid, A., Soroush, A., Glicksberg, B.S., Charney, A.W., Nadkarni, G., Klang, E., 2023. Comparing ChatGPT and GPT-4 performance in USMLE soft skill assessments. Sci Rep 13, 16492. https://doi.org/10.1038/s41598-023-43436-9

Chang, W.L., Grady, N., 2019. NIST Big Data Interoperability Framework: Volume 1, Definitions.

Chen, S., Kann, B.H., Foote, M.B., Aerts, H.J.W.L., Savova, G.K., Mak, R.H., Bitterman, D.S., 2023. Use of Artificial Intelligence Chatbots for Cancer Treatment Information. JAMA Oncol e232954. https://doi.org/10.1001/jamaoncol.2023.2954

Chung, H.W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S., Webson, A., Gu, S.S., Dai, Z., Suzgun, M., Chen, X., Chowdhery, A., Castro-Ros, A., Pellat, M., Robinson, K., Valter, D., Narang, S., Mishra, G., Yu, A., Zhao, V., Huang, Y., Dai, A., Yu, H., Petrov, S., Chi, E.H., Dean, J., Devlin, J., Roberts, A., Zhou, D., Le, Q.V., Wei, J., 2022. Scaling Instruction-Finetuned Language Models. https://doi.org/10.48550/arXiv.2210.11416

Clancey, W.J., Shortliffe, E.H., 1984. Readings in medical artificial intelligence: the first decade. Addison-Wesley Longman Publishing Co., Inc., USA.

Clune, M.W., 2023. AI Means Professors Need to Raise Their Grading Standards [WWW Document]. The Chronicle of Higher Education. URL https://www.chronicle.com/article/ai-means-professors-need-to-raise-their-grading-standards (accessed 9.13.23).

Cooper, A., Rodman, A., 2023. AI and Medical Education — A 21st-Century Pandora's Box. New England Journal of Medicine. https://doi.org/10.1056/NEJMp2304993

Coyner, A.S., Singh, P., Brown, J.M., Ostmo, S., Chan, R.V.P., Chiang, M.F., Kalpathy-Cramer, J., Campbell, J.P., Imaging and Informatics in Retinopathy of Prematurity Consortium, 2023. Association of Biomarker-Based Artificial Intelligence With Risk of Racial Bias in Retinal Images. JAMA Ophthalmol 141, 543–552. https://doi.org/10.1001/jamaophthalmol.2023.1310

Davenport, T.H., Harris, J., Abney, D., 2017. Competing on Analytics: The New Science of Winning; With a New Introduction, Revised Edition. ed. Harvard Business Review Press, Boston, Massachusetts.

DeCamp, M., Lindvall, C., 2023. Mitigating bias in AI at the point of care. Science 381, 150–152. https://doi.org/10.1126/science.adh2713

Decker, H., Trang, K., Ramirez, J., Colley, A., Pierce, L., Coleman, M., Bongiovanni, T., Melton, G.B., Wick, E., 2023. Large Language Model-Based Chatbot vs Surgeon-

Generated Informed Consent Documentation for Common Procedures. JAMA Netw Open 6, e2336997. https://doi.org/10.1001/jamanetworkopen.2023.36997
Dell'Acqua, F., McFowland, E., Mollick, E.R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., Lakhani, K.R., 2023. Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. https://doi.org/10.2139/ssrn.4573321
Donoho, D., 2017. 50 Years of Data Science. Journal of Computational and Graphical Statistics 26, 745–766. https://doi.org/10.1080/10618600.2017.1384734
Dorr, D.A., Adams, L., Embí, P., 2023. Harnessing the Promise of Artificial Intelligence Responsibly. JAMA 329, 1347–1348. https://doi.org/10.1001/jama.2023.2771
Gichoya, J.W., Banerjee, I., Bhimireddy, A.R., Burns, J.L., Celi, L.A., Chen, L.-C., Correa, R., Dullerud, N., Ghassemi, M., Huang, S.-C., Kuo, P.-C., Lungren, M.P., Palmer, L.J., Price, B.J., Purkayastha, S., Pyrros, A.T., Oakden-Rayner, L., Okechukwu, C., Seyyed-Kalantari, L., Trivedi, H., Wang, R., Zaiman, Z., Zhang, H., 2022. AI recognition of patient race in medical imaging: a modelling study. Lancet Digit Health 4, e406–e414. https://doi.org/10.1016/S2589-7500(22)00063-2
Greenes, R., Del Fiol, G. (Eds.), 2023. Clinical Decision Support and Beyond: Progress and Opportunities in Knowledge-Enhanced Health and Healthcare, 3rd edition. ed. Academic Press.
Han, C., Kim, D.W., Kim, S., You, S.C., Park, J.Y., Bae, S., Yoon, D., 2023. Evaluation Of GPT-4 for 10-Year Cardiovascular Risk Prediction: Insights from the UK Biobank and KoGES Data. https://doi.org/10.2139/ssrn.4583995
Han, R., Acosta, J.N., Shakeri, Z., Ioannidis, J., Topol, E., Rajpurkar, P., 2023. Randomized Controlled Trials Evaluating AI in Clinical Practice: A Scoping Evaluation. https://doi.org/10.1101/2023.09.12.23295381
Heaven, W.D., 2023. ChatGPT is going to change education, not destroy it [WWW Document]. MIT Technology Review. URL https://www.technologyreview.com/2023/04/06/1071059/chatgpt-change-not-destroy-education-openai/ (accessed 7.17.23).
Hersh, W., 2023. Physician and Medical Student Competence in AI Must Include Broader Competence in Clinical Informatics. Informatics Professor. URL https://informaticsprofessor.blogspot.com/2023/09/physician-and-medical-student.html (accessed 9.15.23).
Hersh, W., 2020. Information Retrieval: A Biomedical and Health Perspective, 4th ed, Health Informatics. Springer International Publishing.
Hersh, W., Ehrenfeld, J., 2020. Clinical Informatics, in: Health Systems Science, 2nd Edition. pp. 156–170.
Hersh, W.R., 2022. Health Informatics: Practical Guide, 8th Edition. Lulu.com.
Hersh, W.R., Gorman, P.N., Biagioli, F.E., Mohan, V., Gold, J.A., Mejicano, G.C., 2014. Beyond information retrieval and electronic health record use: competencies in clinical informatics for medical education. Adv Med Educ Pract 5, 205–212. https://doi.org/10.2147/AMEP.S63903
Heston, T.F., Khun, C., 2023. Prompt Engineering in Medical Education. International Medical Education 2, 198–205. https://doi.org/10.3390/ime2030019
Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. Neural Comput 18, 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

Holmstrom, L., Christensen, M., Yuan, N., Weston Hughes, J., Theurer, J., Jujjavarapu, M., Fatehi, P., Kwan, A., Sandhu, R.K., Ebinger, J., Cheng, S., Zou, J., Chugh, S.S., Ouyang, D., 2023. Deep learning-based electrocardiographic screening for chronic kidney disease. Commun Med (Lond) 3, 73. https://doi.org/10.1038/s43856-023-00278-w

Huang, J., Neill, L., Wittbrodt, M., Melnick, D., Klug, M., Thompson, M., Bailitz, J., Loftus, T., Malik, S., Phull, A., Weston, V., Heller, J.A., Etemadi, M., 2023. Generative Artificial Intelligence for Chest Radiograph Interpretation in the Emergency Department. JAMA Netw Open 6, e2336100. https://doi.org/10.1001/jamanetworkopen.2023.36100

Huston, J.C., Kaminski, N., 2023. A Picture Worth a Thousand Words, Created with One Sentence: Using Artificial Intelligence–created Art to Enhance Medical Education. ATS Scholar 4, 145–151. https://doi.org/10.34197/ats-scholar.2022-0141PS

James, C.A., Wachter, R.M., Woolliscroft, J.O., 2022. Preparing Clinicians for a Clinical World Influenced by Artificial Intelligence. JAMA 327, 1333–1334. https://doi.org/10.1001/jama.2022.3580

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kanjee, Z., Crowe, B., Rodman, A., 2023. Accuracy of a Generative Artificial Intelligence Model in a Complex Diagnostic Challenge. JAMA 330, 78–80. https://doi.org/10.1001/jama.2023.8288

Kumah-Crystal, Y., Mankowitz, S., Embi, P., Lehmann, C.U., 2023. ChatGPT and the clinical informatics board examination: the end of unproctored maintenance of certification? J Am Med Inform Assoc ocad104. https://doi.org/10.1093/jamia/ocad104

Kung, T.H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., Tseng, V., 2023. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digit Health 2, e0000198. https://doi.org/10.1371/journal.pdig.0000198

Langlotz, C.P., 2019. Will Artificial Intelligence Replace Radiologists? Radiol Artif Intell 1, e190058. https://doi.org/10.1148/ryai.2019190058

Lea, A.S., 2023. Digitizing Diagnosis. Johns Hopkins University Press. https://doi.org/10.56021/9781421446813

Ledley, R.S., Lusted, L.B., 1959. Reasoning foundations of medical diagnosis; symbolic logic, probability, and value theory aid our understanding of how physicians reason. Science 130, 9–21. https://doi.org/10.1126/science.130.3366.9

Lee, B.K., Mayhew, E.J., Sanchez-Lengeling, B., Wei, J.N., Qian, W.W., Little, K.A., Andres, M., Nguyen, B.B., Moloy, T., Yasonik, J., Parker, J.K., Gerkin, R.C., Mainland, J.D., Wiltschko, A.B., 2023. A principal odor map unifies diverse tasks in olfactory perception. Science 381, 999–1006. https://doi.org/10.1126/science.ade4401

Levine, D.M., Tuwani, R., Kompa, B., Varma, A., Finlayson, S.G., Mehrotra, A., Beam, A., 2023. The Diagnostic and Triage Accuracy of the GPT-3 Artificial Intelligence Model. https://doi.org/10.1101/2023.01.30.23285067

Levkovich, I., Elyoseph, Z., 2023. Identifying depression and its determinants upon initiating treatment: ChatGPT versus primary care physicians. Fam Med Community Health 11, e002391. https://doi.org/10.1136/fmch-2023-002391
Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., Zou, J., 2023. GPT detectors are biased against non-native English writers. Patterns (N Y) 4, 100779. https://doi.org/10.1016/j.patter.2023.100779
Liaw, W., Kueper, J.K., Lin, S., Bazemore, A., Kakadiaris, I., 2022. Competencies for the Use of Artificial Intelligence in Primary Care. Ann Fam Med 20, 559–563. https://doi.org/10.1370/afm.2887
Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., Neubig, G., 2023. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. ACM Comput. Surv. 55, 195:1-195:35. https://doi.org/10.1145/3560815
McCarthy, J., Feigenbaum, E.A., 1990. In Memoriam: Arthur Samuel: Pioneer in Machine Learning. AIMag 11, 10–10. https://doi.org/10.1609/aimag.v11i3.840
Meskó, B., 2023. Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial. J Med Internet Res 25, e50638. https://doi.org/10.2196/50638
Miller, R.A., Pople, H.E., Myers, J.D., 1982. Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. N Engl J Med 307, 468–476. https://doi.org/10.1056/NEJM198208193070803
Mitsuyama, Y., Matsumoto, T., Tatekawa, H., Walston, S.L., Kimura, T., Yamamoto, A., Watanabe, T., Miki, Y., Ueda, D., 2023. Chest radiography as a biomarker of ageing: artificial intelligence-based, multi-institutional model development and validation in Japan. The Lancet Healthy Longevity 0. https://doi.org/10.1016/S2666-7568(23)00133-2
Mollick, E., 2023a. My class required AI. Here's what I've learned so far. [WWW Document]. One Useful Thing. URL https://www.oneusefulthing.org/p/my-class-required-ai-heres-what-ive (accessed 7.17.23).
Mollick, E., 2023b. The Homework Apocalypse [WWW Document]. One Useful Thing. URL https://www.oneusefulthing.org/p/the-homework-apocalypse (accessed 10.16.23).
Mollick, E.R., Mollick, L., 2023. Assigning AI: Seven Approaches for Students, with Prompts. https://doi.org/10.2139/ssrn.4475995
Moon, I., LoPiccolo, J., Baca, S.C., Sholl, L.M., Kehl, K.L., Hassett, M.J., Liu, D., Schrag, D., Gusev, A., 2023. Machine learning for genetics-based classification and treatment response prediction in cancer of unknown primary. Nat Med 29, 2057–2067. https://doi.org/10.1038/s41591-023-02482-6
Moskvichev, A., Odouard, V.V., Mitchell, M., 2023. The ConceptARC Benchmark: Evaluating Understanding and Generalization in the ARC Domain. https://doi.org/10.48550/arXiv.2305.07141
Mukherjee, P., Humbert-Droz, M., Chen, J.H., Gevaert, O., 2023. SCOPE: predicting future diagnoses in office visits using electronic health records. Sci Rep 13, 11005. https://doi.org/10.1038/s41598-023-38257-9
Odri, G.-A., Yun Yoon, D.J., 2023. Detecting generative artificial intelligence in scientific articles: evasion techniques and implications for scientific integrity. Orthop Traumatol Surg Res 103706. https://doi.org/10.1016/j.otsr.2023.103706
Omiye, J.A., Lester, J., Spichak, S., Rotemberg, V., Daneshjou, R., 2023. Beyond the hype: large language models propagate race-based medicine. https://doi.org/10.1101/2023.07.03.23292192

Plana, D., Shung, D.L., Grimshaw, A.A., Saraf, A., Sung, J.J.Y., Kann, B.H., 2022. Randomized Clinical Trials of Machine Learning Interventions in Health Care: A Systematic Review. JAMA Netw Open 5, e2233946. https://doi.org/10.1001/jamanetworkopen.2022.33946

Poplin, R., Varadarajan, A.V., Blumer, K., Liu, Y., McConnell, M.V., Corrado, G.S., Peng, L., Webster, D.R., 2018. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng 2, 158–164. https://doi.org/10.1038/s41551-018-0195-0

Pyrros, A., Borstelmann, S.M., Mantravadi, R., Zaiman, Z., Thomas, K., Price, B., Greenstein, E., Siddiqui, N., Willis, M., Shulhan, I., Hines-Shah, J., Horowitz, J.M., Nikolaidis, P., Lungren, M.P., Rodríguez-Fernández, J.M., Gichoya, J.W., Koyejo, S., Flanders, A.E., Khandwala, N., Gupta, A., Garrett, J.W., Cohen, J.P., Layden, B.T., Pickhardt, P.J., Galanter, W., 2023. Opportunistic detection of type 2 diabetes using deep learning from frontal chest radiographs. Nat Commun 14, 4039. https://doi.org/10.1038/s41467-023-39631-x

Rajkomar, A., Kannan, A., Chen, K., Vardoulakis, L., Chou, K., Cui, C., Dean, J., 2019. Automatically Charting Symptoms From Patient-Physician Conversations Using Machine Learning. JAMA Intern Med 179, 836–838. https://doi.org/10.1001/jamainternmed.2018.8558

Rajkomar, A., Oren, E., Chen, K., Dai, A.M., Hajaj, N., Hardt, M., Liu, P.J., Liu, X., Marcus, J., Sun, M., Sundberg, P., Yee, H., Zhang, K., Zhang, Y., Flores, G., Duggan, G.E., Irvine, J., Le, Q., Litsch, K., Mossin, A., Tansuwan, J., Wang, D., Wexler, J., Wilson, J., Ludwig, D., Volchenboum, S.L., Chou, K., Pearson, M., Madabushi, S., Shah, N.H., Butte, A.J., Howell, M.D., Cui, C., Corrado, G.S., Dean, J., 2018. Scalable and accurate deep learning with electronic health records. npj Digital Medicine 1, 1–10. https://doi.org/10.1038/s41746-018-0029-1

Rajpurkar, P., Chen, E., Banerjee, O., Topol, E.J., 2022. AI in health and medicine. Nat Med 1–8. https://doi.org/10.1038/s41591-021-01614-0

Rajpurkar, P., Lungren, M.P., 2023. The Current and Future State of AI Interpretation of Medical Images. N Engl J Med 388, 1981–1990. https://doi.org/10.1056/NEJMra2301725

Rao, A., Pang, M., Kim, J., Kamineni, M., Lie, W., Prasad, A.K., Landman, A., Dreyer, K., Succi, M.D., 2023. Assessing the Utility of ChatGPT Throughout the Entire Clinical Workflow: Development and Usability Study. J Med Internet Res 25, e48659. https://doi.org/10.2196/48659

Raschka, S., 2023. Understanding Encoder And Decoder LLMs. Ahead of AI. URL https://magazine.sebastianraschka.com/p/understanding-encoder-and-decoder (accessed 9.6.23).

Ratliff, M., Sya'ban, S.N., Wazir, A., Haidar, S., Keeth, S., 2023. AI and ChatGPT in Health Professions Education [WWW Document]. Lecturio. URL https://www.lecturio.com/pulse/using-chatgpt-in-medical-education-for-virtual-patient-and-cases/ (accessed 7.26.23).

Roberts, G., 2022. AI Training Datasets: the Books1+Books2 that Big AI eats for breakfast - Musings of Freedom. Musings of Freedom. URL https://gregoreite.com/drilling-down-details-on-the-ai-training-datasets/ (accessed 9.6.23).

Robin, C., 2023. The End of the Take-Home Essay? [WWW Document]. The Chronicle of Higher Education. URL https://www.chronicle.com/article/the-end-of-the-take-home-essay (accessed 8.25.23).

Russell, R.G., Lovett Novak, L., Patel, M., Garvey, K.V., Craig, K.J.T., Jackson, G.P., Moore, D., Miller, B.M., 2023. Competencies for the Use of Artificial Intelligence-Based Tools by Health Care Professionals. Acad Med 98, 348–356. https://doi.org/10.1097/ACM.0000000000004963

Sadasivan, V.S., Kumar, A., Balasubramanian, S., Wang, W., Feizi, S., 2023. Can AI-Generated Text be Reliably Detected? https://doi.org/10.48550/arXiv.2303.11156

Sahni, N.R., Carrus, B., 2023. Artificial Intelligence in U.S. Health Care Delivery. N Engl J Med 389, 348–358. https://doi.org/10.1056/NEJMra2204673

Sangha, V., Nargesi, A.A., Dhingra, L.S., Khunte, A., Mortazavi, B.J., Ribeiro, A.H., Banina, E., Adeola, O., Garg, N., Brandt, C.A., Miller, E.J., Ribeiro, A.L.J., Velazquez, E.J., Giatti, L., Barreto, S.M., Foppa, M., Yuan, N., Ouyang, D., Krumholz, H.M., Khera, R., 2023. Detection of Left Ventricular Systolic Dysfunction From Electrocardiographic Images. Circulation. https://doi.org/10.1161/CIRCULATIONAHA.122.062646

Sarraju, A., Bruemmer, D., Van Iterson, E., Cho, L., Rodriguez, F., Laffin, L., 2023. Appropriateness of Cardiovascular Disease Prevention Recommendations Obtained From a Popular Online Chat-Based Artificial Intelligence Model. JAMA. https://doi.org/10.1001/jama.2023.1044

Shah, C., 2022. A Hands-On Introduction to Machine Learning. Cambridge University Press.

Shortliffe, E.H., Davis, R., Axline, S.G., Buchanan, B.G., Green, C.C., Cohen, S.N., 1975. Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system. Comput Biomed Res 8, 303–320. https://doi.org/10.1016/0010-4809(75)90009-9

Spitale, G., Biller-Andorno, N., Germani, F., 2023. AI model GPT-3 (dis)informs us better than humans. Sci Adv 9, eadh1850. https://doi.org/10.1126/sciadv.adh1850

Tang, J., LeBel, A., Jain, S., Huth, A.G., 2023. Semantic reconstruction of continuous language from non-invasive brain recordings. Nat Neurosci 26, 858–866. https://doi.org/10.1038/s41593-023-01304-9

Terry, O.K., 2023. Opinion | I'm a Student. You Have No Idea How Much We're Using ChatGPT. [WWW Document]. The Chronicle of Higher Education. URL https://www.chronicle.com/article/im-a-student-you-have-no-idea-how-much-were-using-chatgpt (accessed 8.25.23).

Topol, E., 2019. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again, Illustrated Edition. ed. Basic Books, New York.

Topol, E.J., 2023. As artificial intelligence goes multimodal, medical applications multiply. Science 381, adk6139. https://doi.org/10.1126/science.adk6139

Ueda, D., Matsumoto, T., Ehara, S., Yamamoto, A., Walston, S.L., Ito, A., Shimono, T., Shiba, M., Takeshita, T., Fukuda, D., Miki, Y., 2023. Artificial intelligence-based model to classify cardiac functions from chest radiographs: a multi-institutional, retrospective model development and validation study. Lancet Digit Health S2589-7500(23)00107–3. https://doi.org/10.1016/S2589-7500(23)00107-3

Walters, W.H., Wilder, E.I., 2023. Fabrication and errors in the bibliographic citations generated by ChatGPT. Sci Rep 13, 14045. https://doi.org/10.1038/s41598-023-41032-5

Warner, H.R., Toronto, A.F., Veasey, L.G., Stephenson, R., 1961. A mathematical approach to medical diagnosis. Application to congenital heart disease. JAMA 177, 177–183. https://doi.org/10.1001/jama.1961.03040290005002

Weller, O., Marone, M., Weir, N., Lawrie, D., Khashabi, D., Van Durme, B., 2023. “According to ...” Prompting Language Models Improves Quoting from Pre-Training Data. https://doi.org/10.48550/arXiv.2305.13252

Xu, S., Yang, L., Kelly, C., Sieniek, M., Kohlberger, T., Ma, M., Weng, W.-H., Kiraly, A., Kazemzadeh, S., Melamed, Z., Park, J., Strachan, P., Liu, Y., Lau, C., Singh, P., Chen, C., Etemadi, M., Kalidindi, S.R., Matias, Y., Chou, K., Corrado, G.S., Shetty, S., Tse, D., Prabhakara, S., Golden, D., Pilgrim, R., Eswaran, K., Sellergren, A., 2023. ELIXR: Towards a general purpose X-ray artificial intelligence system through alignment of large language models and radiology vision encoders [WWW Document]. arXiv.org. URL https://arxiv.org/abs/2308.01317v2 (accessed 9.26.23).

Zhou, H., 2023. Generative AI, ChatGPT, and Google Bard: Evaluating the Impact and Opportunities for Scholarly Publishing [WWW Document]. The Scholarly Kitchen. URL https://scholarlykitchen.sspnet.org/2023/08/17/generative-ai-chatgpt-and-google-bard-evaluating-the-impact-and-opportunities-for-scholarly-publishing/ (accessed 10.4.23).

Zhou, Q., Chen, Z.-H., Cao, Y.-H., Peng, S., 2021. Clinical impact and quality of randomized controlled trials involving interventions evaluating artificial intelligence prediction tools: a systematic review. NPJ Digit Med 4, 154. https://doi.org/10.1038/s41746-021-00524-2

Zimmerman, J., 2023. Opinion | Here’s my AI policy for students: I don’t have one. Washington Post.

Artificial Intelligence: Implications for Health Professions Education

William Hersh, MD
Professor
Department of Medical Informatics & Clinical Epidemiology
Oregon Health & Science University
Portland, OR, USA

1

Objectives

- After this talk, you will be able to
  - Define the terminology and major types of AI
  - Discuss key issues for AI in health professions education
  - Describe competencies in AI for health professions learners
  - Have pathway for further learning from (annotated) bibliography

2

If you cannot stay the whole hour, here is tl;dr version of talk

- All health professionals must be competent in the uses and limitations of AI technology they are likely to apply in their clinical practice or other work
- All educators of future health professionals must understand the benefits and challenges for AI technologies in education

3

Definitions and terminology related to artificial intelligence (AI)

- AI – "information systems and algorithms capable of performing tasks associated with human intelligence" (Rajpurkar, 2022; Sahni, 2023)
  - Predictive AI – use of data and algorithms to predict some output (e.g., diagnosis, treatment recommendation, prognosis, etc.)
  - Generative AI – generates new output based on prompts (e.g., text, images, etc.)
- A large part of modern success of AI due to machine learning – "computer programs that learn without being explicitly programmed" (McCarthy, 1990, attributed to Samuel, 1959; Shah, 2023)
  - Most success with deep learning, based on many-layered neural networks
- Other terms
  - Data science – science of learning from data (Donoho, 2017)
  - Data analytics – use of data and statistical analysis to build explanatory and predictive models and drive decisions and actions (Davenport, 2017)
  - Big Data – data characterized by large volume, velocity, variety and variability (Chang, 2019)

4

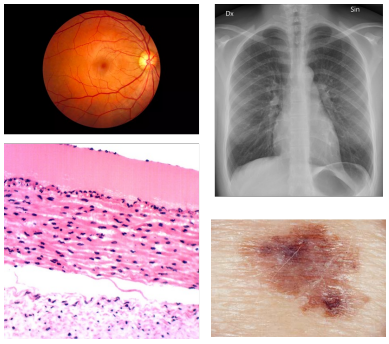
History of AI – first era in mid-20th century

- Earliest paper related to AI and biomedical informatics attributed to Ledley and Lusted (1959) aiming to model physician reasoning through symbolic logic and probability
- Warner (1961) developed mathematical model for diagnosing congenital heart disease
- In 1960s-1970s, emergence of "expert systems" – computer programs aiming to mimic human expertise (historical overview – Lea, 2023)
  - Rule-based systems – PhD dissertation of Shortliffe (1975) and subsequent work (Clancey, 1984)
  - Disease profiles and scoring algorithms – INTERNIST-1 (Miller, 1982) and DxPlain (Barnett, 1987)
- Limited by approach of manual construction and maintenance of knowledge
  - Not scalable or sustainable
  - Led to "AI winter" between 1990-2010
  - Main remnant is clinical decision support (CDS) for electronic health records (EHRs) that emerged in 1990s for electronic health records (Greenes, 2023)

5

Re-emergence of AI in 21st century

- "Predictive AI" driven by advances in machine learning, increasing availability of data, and more powerful computers and networks (Topol, 2019; Rajpurkar, 2022)
  - Deep learning in imaging breakthroughs by Hinton (2006)
- Most success in image interpretation (Rajpurkar, 2023); examples include
  - Radiology – chest x-rays for diagnosis of pneumonia and tuberculosis
  - Ophthalmology – retinal images for diagnosis of diabetic retinopathy
  - Dermatology – skin lesions for diagnosis of cancer
  - Pathology – breast cancer slides to predict metastasis



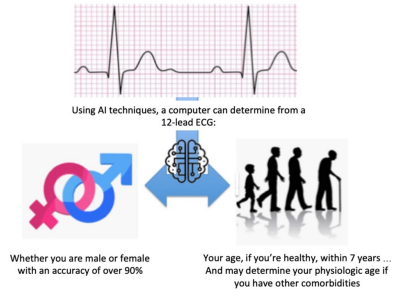
6

Predictive AI not limited to imaging

- Adverse events in hospitalizations from EHR data (Rajkomar, 2018)
- Generating clinical notes from patient and physician verbal interaction (Rajkomar, 2019)
- Protein folding from amino acid sequences (Jumper, 2021)
- Next-generation sequencing (NGS) data on >36K tumors able to predict cancer of unknown primary with high confidence for 41% of tumors, leading to improved survival for those patients (Moon, 2023)
- ML model based on past ICD-10 codes and lab results to predict future diagnoses in office visits (Mukherjee, 2023)
- Semantic reconstruction of continuous language from fMRI brain recordings (Tang, 2023)
- Map chemicals to odors perceived by humans (Lee, 2023)

Also success in "seeing" where humans cannot (Topol, 2023)

- Retinal images
  - Age, biological sex, and cardiovascular risk determination from retinal images (Poplin, 2018)
  - Race (Coyner, 2023)
- Electrocardiograms (ECGs)
  - Age and biological sex determination (Attia, 2019)
  - Diagnosis and risk stratification in occlusive myocardial infarction (Al-Zaiti, 2023)
  - Chronic kidney disease (Holmstrom, 2023)
  - Left ventricular systolic dysfunction from ECG images (Sangha, 2023)
- Chest x-rays
  - Race (Gichoya, 2022)
  - Cardiac function and valvular heart diseases (Ueda, 2023)
  - Diabetes (Pyrros, 2023)
  - Correlation with chronological age in healthy cohorts and, for various chronic diseases, difference between estimated age and chronological age (Mitsuyama, 2023)



And now, "generative AI"

- Introduction of ChatGPT in November, 2022 brought new type of AI into focus: generative AI
- Based on large language models (LLMs) processed by deep neural networks using large amounts of training data and tuned for specific tasks
  - Trained on massive amounts of text and other content, e.g., large Web crawls, books, Wikipedia, and more for GPT (Roberts, 2022)
  - Use transformer models that predict words in sequence from billions/trillions of words and add measure of importance to "attention" words (Raschka, 2023)
  - Fine-tuned for specific tasks (Chung, 2022)
  - Activated by (and importance of) prompting (Liu, 2023)

Results of ChatGPT and other LLMs

- Medical board exams
  - USMLE "arms race," starting with (Kung, 2023)
    - Claimed best – https://www.openevidence.com/blog/openevidence-ai-first-ai-score-above-90-percent-on-the-usmle
    - Even on "soft skills" (e.g., communication skills, ethics, empathy, and professionalism) questions (Brin, 2023)
  - Passing level on some board exams (clinical informatics – Kumah-Crystal, 2023; radiology – Bhayana, 2023) but not others (neonatology – Beam, 2023)
- Answering questions
  - Vary by subject domain and type, but sometimes wrong and/or incomplete (e.g., Antaki, 2023; Chen, 2023; Goodman, 2023)
- Solving clinical cases
  - Comparable to but not better than expert humans (e.g., Levine, 2023; Kanjee, 2023; Rao, 2023; Benoit, 2023; Lekovich, 2023)

Results of ChatGPT and other LLMs (cont.)

- Communicating with patients
  - Answer questions in public forums (Sarraju, 2023; Ayers, 2023)
  - Write letters with comparable or better empathy (Ali, 2023, Ayers, 2023)
  - Generating surgical consent forms better than surgeons (Decker, 2023)
- Use of predictive AI (closing the AI loop)
  - Classifying CXR findings based on previous images and reports (Xu, 2023)
  - Generating CXR reports from new images in ED from prior images and reports (Huang, 2023)
  - Predicting cardiovascular risk comparable to Framingham models (Han, 2023)

11

But there are downsides to generative AI

- Equally compelling disinformation – humans cannot distinguish between true and false tweets generated by GPT-3 and written by real Twitter users (Spitale, 2023)
- Fabrication and errors in the bibliographic citations – asked to produce short literature reviews on 42 multidisciplinary topics (Walters, 2023)
  - 55% of GPT-3.5 citations and 18% of GPT-4 citations fabricated
  - 43% of real (non-fabricated) GPT-3.5 citations and 24% of real GPT-4 citations include substantive errors
- 8 dermatology questions asked of 4 LLMs recapitulated “harmful, race-based medicine” (Omiye, 2023)
- Performs worse than humans in abstraction and analogy problems (Moskvichev, 2023)
- Automated GPT detectors do not work well (Sadasivan, 2023; Odri, 2023)
  - More likely to classify non-native English writing as AI-generated (Liang, 2023)
  - Humans not able to discern AI writing either (Dell'Acqua, 2023)

12

Will AI help or hinder healthcare?

- Real-world use and evidence base still modest – systematic reviews of clinical trials of predictive AI systems (Zhou, 2021; Plana, 2022; Han, 2023) show
  - Small numbers of trials – especially relative to predictive model papers)
  - Suboptimal methodologies leading to risk of bias
  - Mix of positive/negative results
- "AI won't replace radiologists, but radiologists who use AI will replace radiologists who don't," (Langlotz, 2019)
  - (Plug in your health profession)
- Must address bias in data and algorithms
  - AI may compromise care if not used properly (DeCamp, 2023)
  - Must be implemented in responsible (Dorr, 2023) and fair (Chen, 2023) ways

13

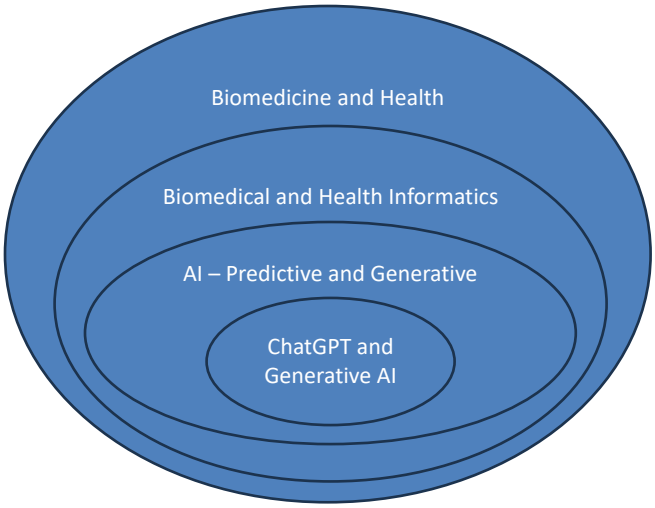
AI and health professions education

- Mostly physician-based but applies to all health professions
- Before generative AI there was recognition of need for competencies in clinical informatics for medical education (Hersh, 2014; Hersh 2020; Hersh, 2023)
- Others note
  - Clinicians must be prepared to practice in a world of AI (James, 2022)
  - Medical schools face dual challenges of needing to teach about AI in practice but also adapt to its use by learners and faculty (Cooper, 2023)
- New AI-competency frameworks (Russell, 2023; Liaw, 2023)

1. Find, search, and apply knowledge-based information to patient care and other clinical tasks
2. Effectively read from, and write to, the electronic health record (EHR) for patient care and other clinical activities
3. Use and guide implementation of clinical decision support (CDS)
4. Provide care using population health management approaches
5. Protect patient privacy and security
6. Use information technology to improve patient safety
7. Engage in quality measurement selection and improvement
8. Use health information exchange (HIE) to identify and access patient information across clinical settings
9. Engage patients to improve their health and care delivery though personal health records and patient portals
10. Maintain professionalism in use of information technology tools, including social media
11. Provide clinical care via telemedicine and refer patients as indicated
12. Apply personalized/precision medicine
13. Participate in practice-based clinical and translational research
14. Use and critique artificial intelligence (AI) applications in clinical care

14

Where does AI fit in health professions education?



15

Generative AI in education – it's already here

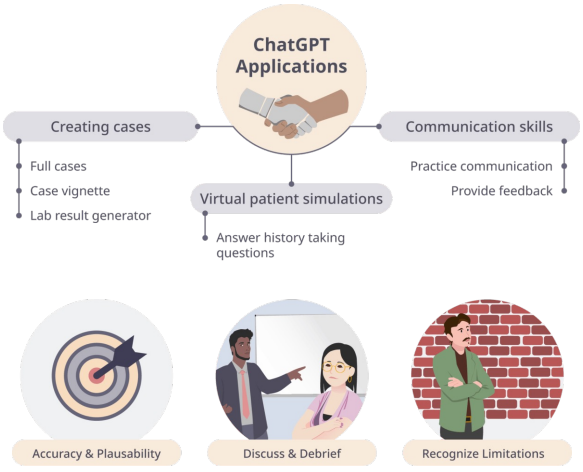
- "My class required AI. Here's what I've learned so far." (Mollick, 2023)
- "I'm a student. You have no idea how much we're using ChatGPT. No professor or software could ever pick up on it." (Terry, 2023)
- "The end of the take-home essay? How ChatGPT changed my plans for the fall." (Robin, 2023)
- "Homework as we know it is over." (Mollick, 2023)
- "Here's my AI policy for students: I don't have one." (Zimmerman, 2023)
- "ChatGPT has transformed the problem of grade inflation from a minor corruption to an enterprise-destroying blight." (Clune, 2023)
- "Assigning AI" (Mollick, 2023 – https://www.oneusefulthing.org/)

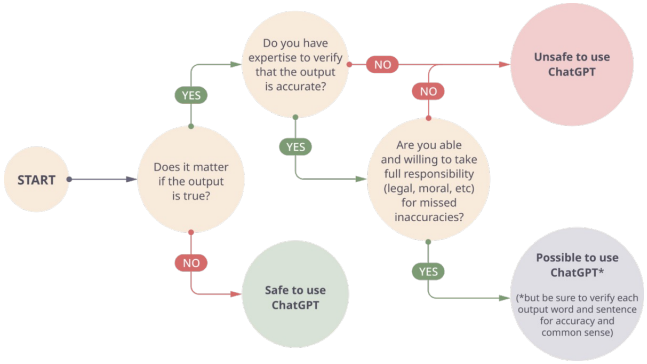
16

Some advocated that ChatGPT will change education, not destroy it (Heaven, 2023)

- Assessment already broken: “if ChatGPT makes it easy to cheat on an assignment, teachers should throw out the assignment rather than ban the chatbot”
- Change focus: use ChatGPT to generate an argument and then annotate it according to how effective argument was for a specific audience; then turn in rewrite based on their criticism
- Overcome misinformation and bias: ask students to use ChatGPT to generate text on a topic and then point out flaws
- Interact with ChatGPT to debate and generate counterarguments

ChatGPT in medical education (Ratliff, 2023)





Recommendations for medical faculty and institutions (Boscardin, 2023)

Educators

- Increase AI knowledge
- Understand the current landscape of AI use in medical education
- Review strategies for successful AI integration into education
- Become stewards of ethical use of AI

Institutions

- Review and revise school policies (and create new policies as needed) regarding use of generative AI
- Support faculty development about AI and provide resources for teaching
- Offer information-checking tools for originality and plagiarism to faculty

Prompt engineering

- Best practices (Heston, 2023)
  - Structure with context, general request, how the generative AI is to act, and output format
  - Utilize iterative prompts
  - Avoid vague, misleading, and inappropriate prompts
  - Mathematical, logical, and academic citation prompts can be challenging
  - Clarify what constitutes "cheating"
- Tutorial (Meskó, 2023)
- Adding "according to sources" to ChatGPT prompts improves quoting of pre-trained data (Weller, 2023)
- "Cheat sheets"
  - https://www.aifire.co/c/ai-learning-resources

Uses and risks of "assigning AI" (Mollick, 2023)

AI USE	ROLE	PEDAGOGICAL BENEFIT	PEDAGOGICAL RISK
MENTOR	Providing feedback	Frequent feedback improves learning outcomes, even if all advice is not taken.	Not critically examining feedback, which may contain errors.
TUTOR	Direct instruction	Personalized direct instruction is very effective.	Uneven knowledge base of AI. Serious confabulation risks.
COACH	Prompt metacognition	Opportunities for reflection and regulation, which improve learning outcomes.	Tone or style of coaching may not match student. Risks of incorrect advice.
TEAMMATE	Increase team performance	Provide alternate viewpoints, help learning teams function better.	Confabulation and errors. "Personality" conflicts with other team members.
STUDENT	Receive explanations	Teaching others is a powerful learning technique.	Confabulation and argumentation may derail the benefits of teaching.
SIMULATOR	Deliberate practice	Practicing and applying knowledge aids transfer.	Inappropriate fidelity.
TOOL	Accomplish tasks	Helps students accomplish more within the same time frame.	Outsourcing thinking, rather than work.

Risks:

- Confabulation
- Bias – from training content
- Privacy – policies not always clear
- Instructional – student over-reliance

21

Other uses of ChatGPT

- Using ChatGPT in medical practice, e.g.,
  - Developing patient handouts in rural dermatology (Baker, 2023)
  - Create checklists for common presentations and generate templates for common clinical scenarios (Bair, 2023)
- Assist with aspects of scholarly publishing (Zhou, 2023)
- Plugins enhance functionality
  - BrowserPilot – access live Web sites
  - ScholarAI – search PubMed and other research databases
  - SmartSlides – generate (short) Powerpoint presentations

22

Other uses of LLMs

- Creating medical art (Huston, 2023)
- Adding generative AI to search, e.g.,
  - Bing – with version of GPT-4
  - Google – with Bard/PaLM
- As someone whose research area includes information retrieval (Hersh, 2020), do not yet find generative AI ready to replace most types of searching, especially
  - Comprehensive literature review
  - Known-item searching

My (small) use of ChatGPT in informatics teaching so far (due next month)

- You need to have a conversation with ChatGPT, not just ask it a single question.
- There are some good recent papers on methods for prompting ChatGPT (Heston, 2023; Meskó, 2023).
- You conversation with ChatGPT should ask it about some aspect of informatics that interests you or is important to your job. You can further ask it about things you want to learn more about or point out things it says that are incorrect or are missing and should be there.
- You should include your own post-conversation analysis summarizing what ChatGPT told you, where it added new knowledge for you, what you believe it said that was not correct (i.e., confabulating), and what was missing. If you asked it to provide references, you should identify any that were incorrect.
- All of your conversation should be recorded in a Word document and then submitted as outlined in the syllabus. You can do this by copy and paste from ChatGPT to get your conversation into the Word document, and then add your post-conversation analysis after that. Your document should be the same 2-3 pages as required for the original assignment.

Uses of LLMs for which I am more ambivalent

- Online discussion forums – could be helpful for students with difficulty expressing themselves vs. disincentivizing original thinking
- Quizzes and exams – tend to prefer open-book testing of applying knowledge but preliminary analysis shows ChatGPT capable of passing
- Conclusion – we need discussion and policy

25

Competencies for use of AI-based tools by healthcare professionals (Russell, 2023)

Domain	Competency
Basic knowledge of AI	Explain what AI is and describe its healthcare applications
Social and ethical implications of AI	Explain how social, economic, and political systems influence AI-based tools and how these relationships impact justice, equity, and ethics
AI-enhanced clinical encounters	Carry out AI-enhanced clinical encounters that integrate diverse sources of information in creating patient-centered care plans
Evidence-based evaluation of AI-based tools	Evaluate the quality, accuracy, safety, contextual appropriateness, and biases of AI-based tools and their underlying datasets in providing care to patients and populations
Workflow analysis for AI-based tools	Analyze and adapt to changes in teams, roles, responsibilities, and workflows resulting from implementation of AI-based tools
Practice-based learning and improvement regarding AI-based tools	Participate in continuing professional development and practice-based improvement activities related to use of AI tools in healthcare

26

Competencies for use of AI in primary care (Liaw, 2023) – applicable to all health professions use

Domain	Bottom Line	Competency
Foundational knowledge	What is this tool?	Clinicians will explain the fundamentals of AI, how AI-based tools are created and evaluated, the critical regulatory and socio-legal issues of the AI-based tools, and the current and emerging roles of AI in health care.
Critical appraisal	Should I use this tool?	Clinicians will appraise the evidence behind AI-based tools and assess their appropriate uses via validated evaluation frameworks for health care AI.
Medical decision making	When should I use this tool?	Clinicians will identify the appropriate indications for and incorporate the outputs of AI-based tools into medical decision making such that effectiveness, value, equity, fairness, and justice are enhanced.
Technical use	How do I use this tool?	Clinicians will execute the tasks needed to operate AI-based tools in a manner that supports efficiency and builds mastery.
Patient communication	How should I communicate with patients regarding the use of the tool?	Clinicians will communicate what the tool is and why it is being used, answer questions about privacy and confidentiality, and engage in shared decision making, in a manner that preserves or augments the clinician-patient relationship.
Unintended consequences (cross-cutting)	What are the "side effects" of this tool?	Clinicians will anticipate and recognize the potential adverse effects of AI-based tools and take appropriate actions to mitigate or address unintended consequences.

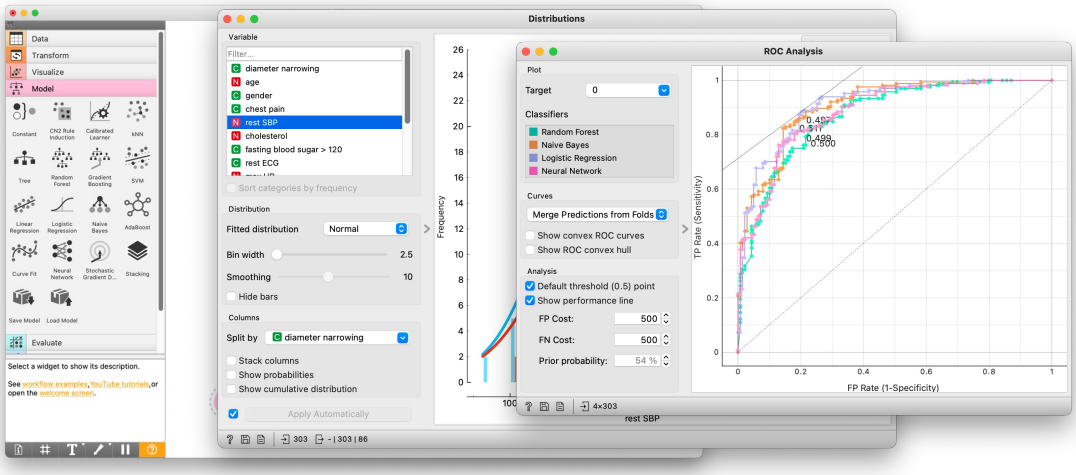
27

Teaching AI is more than ChatGPT/LLMs – some examples from DMICE R25 summer internship

- Predictive AI
  - Orange Data Mining
    - https://orangedatamining.com/
  - Open-source "no code" visual programming environment to teach data exploration and visualization, ML modeling, etc.
  - Excellent set of introductory YouTube videos
    - https://www.youtube.com/@OrangeDataMining
- Health Data Equity and Ethics Seminar
  - Student-led presentations and discussion about issues related to data and algorithm bias, fair and responsible use of AI, etc.

28

Visual programming, exploration, and modeling in Orange



Conclusions

- Predictive and generative AI will profoundly impact the practice and education of all health professions
  - Day-to-day impact, especially in clinical settings, small so far but likely to grow
  - Need real-world implementation and evaluation for safety and efficacy just like all other clinical interventions
- Clinical and informatics professionals must be able to understand, implement, and critique applications of AI in their work and in healthcare more broadly
- Health professions educators must adapt to generative AI for writing, examination, and other pedagogic tasks
  - Where do we draw line on appropriate use?

Questions?

William Hersh, M.D.
Professor
Department of Medical Informatics & Clinical Epidemiology
Oregon Health & Science University
Portland, OR, USA
Email: hersh@ohsu.edu
Web: http://www.billhersh.info
Blog: https://informaticsprofessor.blogspot.com/
Textbook: http://www.informaticsbook.info
What is Informatics?: http://informatics.health
Also on
Twitter – @williamhersh
Facebook
LinkedIn