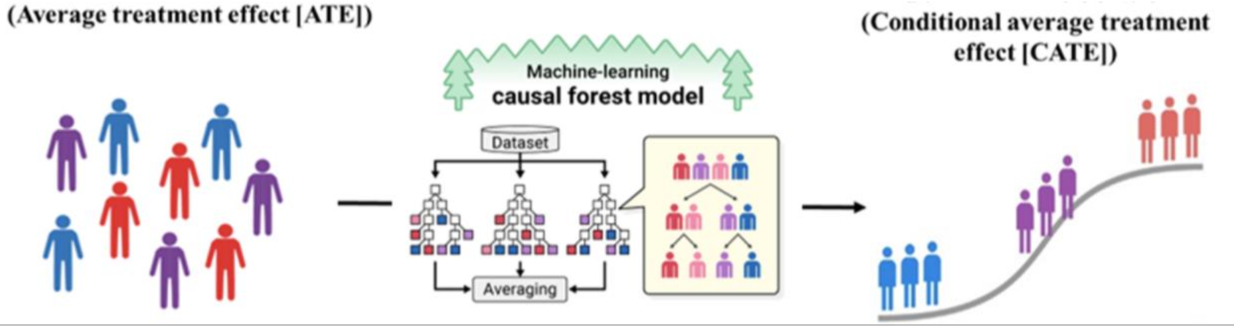
In 2008, a group of researchers ran the Oregon Health Insurance Experiment (OHIE), which evaluated the health effects of Medicaid, a health insurance program for low-income households.[1] The researchers took advantage of the randomized allocation of insurance enrollment through a lottery, and evaluated the effects of Medicaid coverage using a randomized controlled design. The OHIE received significant attention, as public health insurance programs for low-income households have been implemented in many countries, but its health effects had rarely been explored. The investigators found that Medicaid coverage significantly decreased the probability of a positive screening result for depression.[2]

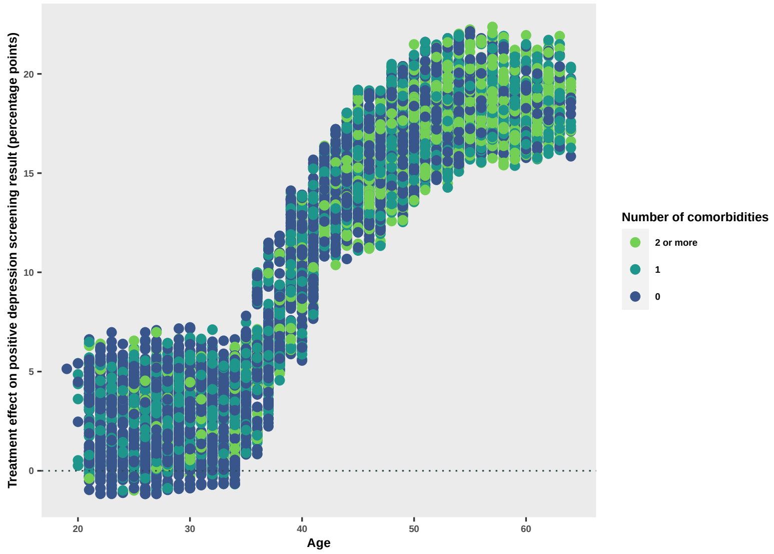
However, the average benefits of Medicaid expansion in treating depression seen in the OHIE may mask substantial heterogeneity, with some people benefitting much more than others. In this post hoc analysis of the OHIE, we assessed the degree of response heterogeneity and the extent to which it is predictable *ex ante*. By applying a novel machine learning method recently introduced in the econometrics literature, the causal forest,[3] we delineated the characteristics of individuals with high or low predicted benefit and evaluate both the health benefits and efficiency of an approach for targeting health insurance coverage on those most likely to benefit. Thus, in the original OHIE, investigators evaluated the average treatment effect (ATE) of Medicaid coverage, but in the present study, we evaluated the conditional average treatment effects (CATEs; alternatively, individual treatment effects) of Medicaid coverage (**Figure 1**). With this approach, we were able to evaluate the causal effects of Medicaid coverage at the individual level, conditional on individual level characteristics such as gender, age, educational level, race and ethnicity, whether the interview was conducted in English or not, diagnoses before the lottery, and previous histories of emergency department visits. To screen for depression an average of 25 months post-treatment, we used the Patient Health Questionnaire-8 (PHQ-8). Treatment effects were expressed as percentage point reduction in positive depression screening. A total of 10,068 low-income individuals were eligible for analysis.

Figure 1. A comparison of average treatment effect (ATE) and conditional average treatment effect (CATE). By using the causal forest, we were able to evaluate the CATE (the causal effect at the individual level). Figure taken from an article (https://healthpolicyhealthecon.com/2023/04/06/causal-forest-blood-pressure-control/) by Dr. Yusuke Tsugawa, a co-author of this study.



We found that individuals with high predicted benefit were older and had more physical or mental health conditions at baseline (**Figure 2**), and there was substantial heterogeneity in the effects of Medicaid coverage on depression. Considering a scenario where Medicaid coverage cannot be expanded to the full population due to e.g., budget constraints, we evaluated the "high-benefit approach," where Medicaid coverage is expanded to only those with high predicted benefit. The cutoff for high benefit was defined as the median of the estimated CATEs. We found that compared to expanding coverage to the full sample of 10,068 low-income individuals ("population approach"), expanding coverage to 5,034 individuals with high predicted benefit ("high benefit approach") generated greater reduction in the proportion of individuals screening positive for depression (21.5 vs. 8.8 percentage point reduction; adjusted difference [95% confidence interval (CI)], +12.7 [+4.6, +20.8]; P=0.003), at substantially lower cost per case prevented ($16,627 vs. $36,048; adjusted difference [95%CI], -$18,598 [-$156,953, -$3,120]; P=0.04).

Figure 2. Conditional average treatment effects (individual treatment effects) across age (horizontal axis) and number of comorbidities (color-coded). Individuals with older age and greater number of comorbidities had greater conditional average treatment effects.



In this post hoc analysis of the OHIE using the machine learning causal forest, we found substantial – and predictable – heterogeneity in the effect of Medicaid coverage on depression. Those who experienced large improvements in depression were older and had more baseline physical or mental health conditions. We found that providing Medicaid coverage to individuals with high likelihood of benefit as predicted using *ex ante* information reduced depression by a three-fold greater margin than providing coverage to all low-income individuals. This approach was not only effective in reducing depression cases, but also more cost-effective than broader expansions as captured by the healthcare spending per case of depression averted. Taken together, our findings suggest that it is possible to use baseline information to prioritize coverage expansion to those who are likely to benefit the most, and demonstrate the promises of a novel machine learning-based approach to precision health and policy.

References

1. Finkelstein A, Taubman S, Wright B, et al. The Oregon Health Insurance Experiment: Evidence from the First Year. *The Quarterly Journal of Economics*. 2012;127(3):1057-1106.

2. Baicker K, Taubman SL, Allen HL, et al. The Oregon experiment—effects of Medicaid on clinical outcomes. *New England Journal of Medicine*. 2013;368(18):1713-1722.

3. Athey S, Tibshirani J, Wager S. Generalized random forests. *Annals of Statistics*. 2019;47(2):1148-1178.

4. Goto R, Inoue K, Osawa I, Baicker K, Fleming SL, Tsugawa Y. Machine learning detects heterogeneous effects of Medicaid coverage on depression. *American Journal of Epidemiology*. In press.

5. Goto R, Frodl T, Skokauskas N. Armed Conflict and Early Childhood Development in 12 Low- and Middle-Income Countries. *Pediatrics*. 2021;148(3):e2021050332.

6. Goto R, Guerrero APS, Speranza M, Fung D, Paul C, Skokauskas N. War is a public health emergency. *The Lancet*. 2022; 399(1302):1302.

7. Goto R, Nianogo R, Okubo Y, Inoue K. Evaluation of Obesity Trends Among US Adolescents by Socioeconomic Status, 1999-2018. *JAMA Pediatrics*. 2022;176(9):937.

8. Goto R, Okubo Y, Skokauskas N. Reasons and trends in youth's suicide rates during the COVID-19 pandemic. *The Lancet Regional Health-Western Pacific*. 2022;27:100567.

9. Goto R, Kawachi I, Kondo N, Inoue K. Contribution of Vaccinations to Reducing Socioeconomic Disparities in COVID-19 Deaths Across US counties. *Annals of Epidemiology*. 2023;86:65-71.e3.