

NAME OF PRESENTING AUTHOR: Aysu Okbay

EMAIL ADDRESS OF PRESENTING AUTHOR: a.okbay@vu.nl

Resource Profile and User Guide of the Polygenic Index Repository

Joel Becker¹, Casper A.P. Burik², Grant Goldman³, Nancy Wang³, Hariharan Jayashankar³, Michael Bennett³, Daniel W. Belsky^{4,5}, Richard Karlsson Linnér², Rafael Ahlskog⁶, Aaron Kleinman⁷, David A. Hinds⁷, 23andMe Research Group⁷, Avshalom Caspi⁸⁻¹¹, David L. Corcoran¹⁰, Terrie E. Moffitt⁸⁻¹¹, Richie Poulton¹², Karen Sugden⁸, Benjamin S. Williams⁸, Kathleen Mullan Harris^{13,14}, Andrew Steptoe¹⁵, Olesya Ajnakina^{15,16}, Lili Milani¹⁷, Tõnu Esko^{17,18}, William G. Iacono¹⁹, Matt McGue¹⁹, Patrik K.E. Magnusson²⁰, Travis T. Mallard²¹, K. Paige Harden^{21,22}, Elliot M. Tucker-Drob^{21,22}, Pamela Herd²³, Jeremy Freese²⁴, Alexander Young^{25,26}, Jonathan P. Beauchamp²⁷, Philipp Koellinger^{2,28}, Sven Oskarsson⁶, Magnus Johannesson²⁹, Peter M. Visscher³⁰, Michelle N. Meyer³¹, David Laibson^{3,32}, David Cesarini^{1,3}, Daniel J. Benjamin^{3,25,26}, Patrick Turley^{33,34}, and Aysu Okbay²

¹ Department of Economics, New York University, New York, NY, USA.

² Department of Economics, School of Business and Economics, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands.

³ National Bureau of Economic Research, Cambridge, MA, USA.

⁴ Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA.

⁵ Robert N. Butler Columbia Aging Center, Columbia University, New York, NY, USA.

⁶ Department of Government, Uppsala University, Uppsala, Sweden.

⁷ 23andMe, Inc., Mountain View, CA, USA.

⁸ Department of Psychology and Neuroscience, Duke University, Durham, NC, USA.

⁹ Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom.

¹⁰ Center for Genomic and Computational Biology, Duke University, Durham, NC, USA.

¹¹ Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA.

¹² Dunedin Multidisciplinary Health and Development Research Unit, University of Otago, Dunedin, Otago, New Zealand.

¹³ Department of Sociology, University of North Carolina at Chapel Hill, Chapel Hill, NC USA.

¹⁴ Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC USA.

¹⁵ Department of Behavioural Science and Health, University College London, London, United Kingdom.

¹⁶ Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom.

¹⁷ Institute of Genomics, University of Tartu, Tartu, Estonia.

¹⁸ Broad Institute of MIT and Harvard, Cambridge, MA, USA.

¹⁹ Department of Psychology, University of Minnesota, Minneapolis, MN, USA.

²⁰ Swedish Twin Registry, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

- ²¹ Department of Psychology, The University of Texas at Austin, Austin, TX, USA.
- ²² Population Research Center, The University of Texas at Austin, Austin, TX, USA.
- ²³ McCourt School of Public Policy, Georgetown University, Washington, DC, USA.
- ²⁴ Department of Sociology, Stanford University, Stanford, CA, USA.
- ²⁵ UCLA Anderson School of Management, Los Angeles, CA, USA.
- ²⁶ Human Genetics Department, UCLA David Geffen School of Medicine, Los Angeles, CA, USA.
- ²⁷ Interdisciplinary Center for Economic Science and Department of Economics, George Mason University, Fairfax, Virginia, USA.
- ²⁸ Robert M. La Follette School of Public Affairs, University of Wisconsin-Madison, Madison, WI, USA.
- ²⁹ Department of Economics, Stockholm School of Economics, Stockholm, Sweden.
- ³⁰ Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia.
- ³¹ Center for Translational Bioethics and Health Care Policy, Geisinger Health System, Danville, PA, USA.
- ³² Department of Economics, Harvard University, Cambridge, MA, USA.
- ³³ Center for Economic and Social Research, University of Southern California, Los Angeles, CA, USA.
- ³⁴ Department of Economics, University of Southern California, Los Angeles, California, USA.

KEYWORDS: Polygenic score, polygenic index, repository, measurement error

ABSTRACT: Polygenic indexes (PGIs) are DNA-based predictors. Their value for research in many scientific disciplines is rapidly growing. As a resource for researchers, we used a consistent methodology to construct PGIs for 47 phenotypes in 11 datasets. To maximize the PGIs' prediction accuracies, we constructed them using genome-wide association studies—some not previously published—from multiple data sources, including 23andMe and UK Biobank. We present a theoretical framework to help interpret analyses involving PGIs. A key insight is that a PGI can be understood as an unbiased but noisy measure of a latent variable we call the “additive SNP factor.” Regressions in which the true regressor is the additive SNP factor but the PGI is used as its proxy therefore suffer from errors-in-variables bias. We derive an estimator that corrects for the bias, illustrate the correction, and make a Python tool for implementing it publicly available.

GRANT SUPPORT: The study was supported by funding from the Ragnar Söderberg Foundation (E42/15, D.C.);, the Swedish Research Council (421-2013-1061, M.J.; 2019-00244, S.O.);, an ERC Consolidator Grant (647648 EdGe, P.K.);, the Pershing Square Fund of the Foundations of Human Behavior (D.L.);, Open Philanthropy (010623-00001, D.J.B., P.T., M.N.M.);, Riksbankens Jubileumsfond P18-0782:1 (S.O.);, Netherlands Organisation for Scientific Research VENI grant 016.Veni.198.058 (A.O.);, and the NIA/NIH through grants R24-AG065184 (D.J.B.) and R01-AG042568 (D.J.B.) to the University of California Los Angeles, ; K99-AG062787-01 (P.T.) to Massachusetts General Hospital, and R56-AG058726 (Titus Galama) to the University of Southern California; the NIH/NICHD through grants R01-HD086313 (E.M.T.D.) and R01-HD092548 (K.P.H.) to the University of Texas at Austin; the NIA/NIMH through grants 1R01-MH101244-02 (P.T.; PI: Benjamin M. Neale) and 5U01-MH109539-02 (P.T.; PI: B.M.N.) to the Broad Institute at Harvard and MIT; R56-AG058726 (Titus Galama) to the University of Southern California; the Government of Canada through Genome Canada and the Ontario Genomics Institute (OGI-152) (J.P.B.); the Social Sciences and Humanities Research Council of Canada (J.P.B.); the European Union through grant MP1GI18418R (T.E.); the Estonian Research Council through grant PRG1291 (T.E.);, the National Health and Medical Research Council through grant GNT113400 (P.M.V.); and the Australian Research Council (P.M.V.).